HYDROLOGICAL PROCESSES
Hydrol. Process. **28**, 1032–1042 (2014)
Published online 17 December 2012 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.9645

Using the SWAT model to assess impacts of land use changes on runoff generation in headwaters

Guoqiang Wang,[1] Huicai Yang,[1] Lijing Wang,[2] Zongxue Xu[1] and Baolin Xue[3]*

[1] *College of Water Sciences, Beijing Normal University, Beijing 100875, China*
[2] *State Key Laboratory of Environmental Criteria and Risk Assessment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China*
[3] *State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China*

Abstract:

The upstream regions of the Three Gorges Reservoir (TGR) have undergone significant changes in land use during recent years, and these changes have strongly influenced runoff generation downstream. In this study, the relationships between land use changes and corresponding hydrological responses in the Dong and Puli River basins in the upstream region of the TGR were quantified using the runoff coefficient. Empirical regression equations between the runoff coefficient and the percentage of land use types were developed for the study area using partial least squares regression (PLSR). The Soil and Water Assessment Tool was used to simulate the runoff generation processes in the two basins, and land use maps developed using Landsat Thematic Mapper images from 2000, 2005, and 2010 were compared to extract information on changes in land use. The results showed that the total area of forest and pasture decreased over the 10-year study period, while paddy fields and upland increased in both basins. These land use changes dramatically affected hydrological processes. Evapotranspiration decreased by 2.13% and 2.41% between 2000 and 2010 in the Dong and Puli River basins, respectively, whereas quickflow, infiltration, and baseflow increased to varying degrees. The PLSR modeling results showed that upland had a negative effect on the runoff coefficient and was the most influential land use type in the study area. In contrast, a positive effect of forest on runoff generation was found in most of the regression models. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS land use; Three Gorges Reservoir; runoff coefficient; SWAT

Received 26 July 2012; Accepted 5 November 2012

INTRODUCTION

Quantification of the effect of land use changes on runoff dynamics in river basins has been a focus of research for hydrologists in recent years. However, it is not yet known whether a well-defined quantitative relationship exists between various land use types and runoff generation mechanisms (Hundecha and Bardossy, 2004). This lack of knowledge is mainly a consequence of the heterogeneity of land use characteristics in river basins, and the impact of the concurrent influence of climate, which is often difficult to distinguish (Sullivan *et al*., 2004; Wang *et al*., 2009). The construction of reservoirs is an important driving factor underlying land use changes. China's Three Gorges Reservoir (TGR) project, the largest water conservation project in the world, has greatly influenced land use patterns in the surrounding regions as a result of population migration and dam construction. Zhang *et al*. (2009a) found that cropland, woodland, and grassland areas in the regions surrounding the TGR decreased continuously during 1975–2005. The rapid and significant land use changes upstream of the TGR altered runoff generation mechanisms and changed hydrological responses downstream (Zhang *et al*., 2009a)

Different methods have been used to assess the impact of land use changes on runoff, but to date no generally applicable and credible model has been developed. In the past, the impact of land use changes on runoff was mainly investigated using catchment experiments. These produced variable and sometimes contradictory results (Hundecha and Bardossy, 2004). Hibbert (1967) reported a clear increase in water yield associated with a reduction in forest cover. After reviewing a number of studies, Hollis (1975) concluded that urbanization increases runoff. However, experimental methods for research on hydrological responses to land use change are time consuming and hydrological models are now being increasingly used to assess the impact of land use changes. Lørup *et al*. (1998) and Schreider *et al*. (2002) calibrated lumped models for a reference period with little change in land use and applied the calibrated models to a subsequent period over which changes had occurred. To investigate changes in catchment runoff that might arise from the land use changes they conducted trend analysis of the bias between modeled and observed runoff (Lørup *et al*., 1998; Schreider *et al*., 2002). Fohrer *et al*. (2001) used a physically based model to predict the impact of land use changes through a model sensitivity study of those changes. Wooldridge *et al*. (2001) regionalized the parameters of a simple model for forest and non-forest land use classification and different climate regions, with the aim of predicting the influence of land use changes on

*Correspondence to: Baolin Xue, State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China.
E-mail: xuebaolin@qq.com

the hydrological response of a basin. These studies have been a major step forward in improving our understanding of the impact of land use changes on runoff generation. However, to eliminate the influences of climate variability (particularly rainfall) on the runoff generation, the researchers had to select study periods for model calibration and validation with comparable total rainfall, or had to divide the catchment into several climate zones according to regional rainfall averages. This necessitated even longer time series for meteorological and discharge data, or complex spatial analysis of rainfall. A simple and easily applied method to constrain the influence of rainfall on runoff is therefore necessary.

In this study, we selected the runoff coefficient, the ratio between total runoff and total rainfall, to examine the impact of land use changes on runoff generation. The runoff coefficient can reflect the complex influences of land use on water storage in a given river basin (Merz and Bloschl, 2009). However, there have been few quantitative studies of the impact of land use changes on runoff generation using the runoff coefficient. Sriwongsitanon and Taesombat (2011) found that runoff coefficients were strongly influenced by land use in a northern Thailand basin. The presence of forest increased the coefficients significantly, whereas agriculture and disturbed forest decreased the runoff coefficients in most sub-basins. This work confirmed that the runoff coefficient is a good proxy to assess the impact of land use changes on runoff generation. However, the study conducted by Sriwongsitanon and Taesombat (2011) focused on peak flow rate rather than overall hydrography influenced by land use changes, and the researchers did not present a predictive method for correlating the runoff coefficients and land use changes (Sriwongsitanon and Taesombat, 2011).

The main aim of our study was to quantify the impact of land use changes on runoff generation in the region upstream of the TGR. The Soil and Water Assessment Tool (SWAT) was used to perform hydrological simulations in two small river basins upstream of the TGR. Land use maps derived from Landsat Thematic Mapper (TM) images acquired in 2000, 2005, and 2010 were used to analyze basin land use changes. The runoff coefficient was used to examine the relationship between land use changes and corresponding hydrological responses, and partial least squares regression (PLSR) was applied to build correlations between land use and runoff coefficients. Results were analyzed to determine which land use types had obvious effects on runoff generation in the region upstream of the TGR.

STUDY AREA

The TGR, in the middle reaches of the Yangtze River in China, is one of the largest reservoirs in the world. It was built to harness hydropower and to mitigate floods and droughts in the middle and lower reaches of the Yangtze River, and it connects lake basins and river tributaries. The study area was within a typical area upstream of the TGR, along the Pengxi River (30°50′–31°42′N, 107°56′–108°54′E). The length of the mainstream of the Pengxi River is about 182 km, and the Pengxi River basin covers an area of approximately 5172.5 km^2. Average annual precipitation over the basin is 1100–1500 mm, and average annual runoff is about 3.41 billion m^3. Based on the availability of discharge data, we focused on two sub-basins of the Pengxi River basin. As shown in Figure 1, the Dong and Puli Rivers are two major tributaries of the

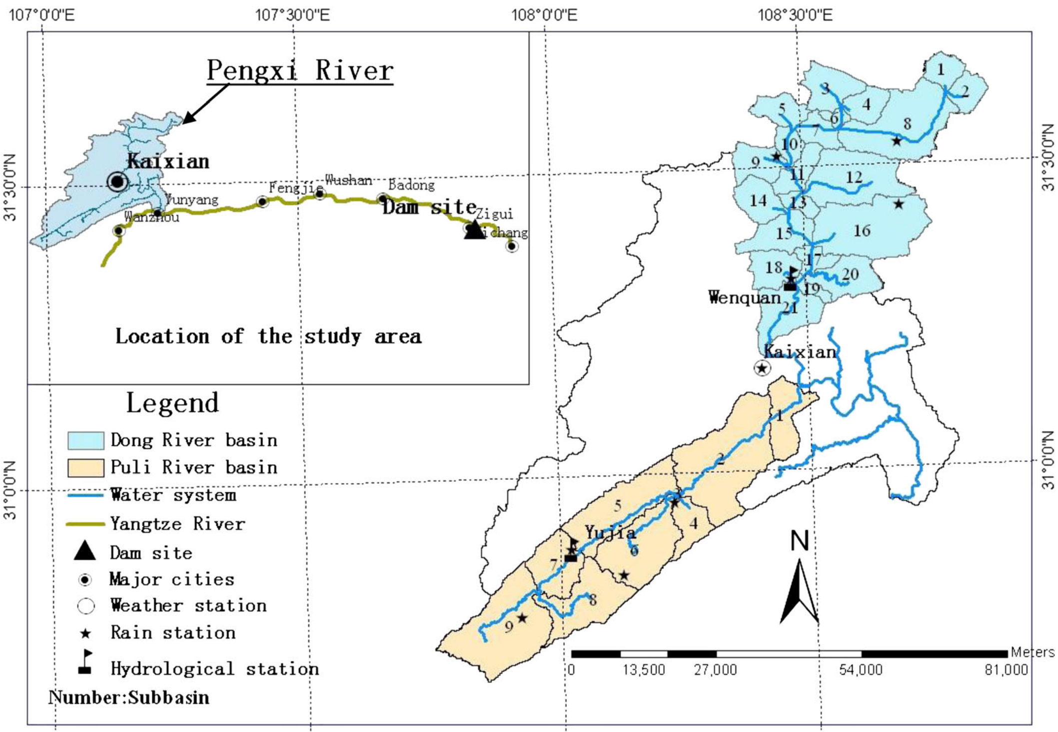

Figure 1. Location of the study area

Pengxi River in the region upstream of the TGR. The Dong River, in the northern part of the study area, has a mainstream length of about 106 km, and its basin has a drainage area of 1469 km^2, and an average slope of about 18.97 degree. The Puli River has a mainstream length of about 121 km, and its basin, located in the southern part of the Pengxi River basin, has a drainage area of 1151 km^2 and an average slope of about 11.02 degree. Average annual discharges are 43.31 m^3/s and 21.15 m^3/s for the Dong River and Puli River, respectively. About 221500 immigrants have settled along the two rivers since the commencement of the TGR project. Land use patterns in the basins upstream of the TGR were relatively stable in the 1980s, but they have changed significantly since the 1990s because of socio-economic development in neighboring cities and as a result of the construction of the TGR.

METHODOLOGY

SWAT model description

SWAT is a temporally continuous, physically based hydrological model that considers both upland and stream processes in a catchment (Arnold *et al*., 1998). SWAT subdivides a basin into sub-basins connected by a stream network and further delineates each sub-basin in to Hydrologic Response Units (HRUs) consisting of unique combinations of land use and soils. SWAT allows a number of different physical processes to be simulated in a basin. The hydrological routines within SWAT account for snowfall and melt, vadose zone processes (infiltration, evaporation, plant uptake, lateral flows, and percolation), and groundwater flows (Zhang *et al*., 2009b). The hydrological cycle simulated by SWAT is based on the water balance equation:

$$SW_t = SW_{t-1} + \sum_{i=1}^{t} (R_i - Q_i - ET_i - P_i - QR_i), \quad (1)$$

where SW_t (mm) is the final soil water content, SW_{t-1} (mm) is the initial soil water content on day i, t (days) is time, R_i (mm) is the precipitation amount on day i, Q_i (mm) is the amount of surface runoff on day i, ET_i (mm) is the evapotranspiration (ET) amount on day i, P_i (mm) is the amount of water entering the vadose zone from the soil profile on day i, and QR_i (mm) is the amount of return flow on day i.

In the SWAT model, there are numerous parameters to be calibrated for matching simulated and observed flows. For this study, ten parameters were identified as sensitive as suggested by Van Liew *et al*. (2007) and tested for the two basins (Table I). According to the sensitivity analysis, runoff was most sensitive to curve number *CN2* for both basins, but the ranking of sensitivity for the ten parameters varied between the two basins (Table I).

In this study, the SWAT model was calibrated and validated at daily time steps based on discharges at Wenquan and Yujia stations, at the basin outlets of the Dong and Puli Rivers, respectively (Figure 1). An automatic parameter estimation procedure, SWAT-CUP (SWAT-Calibration and Uncertainty Procedures), was used to estimate parameter values for the runoff simulations. To estimate the impact of land use change on runoff generation, the model was calibrated and validated using land use data from 2000. Land use data from the years 2005 and 2010 were then used in the model, with all other parameters remaining the same. The model performance was evaluated using the Nash–Sutcliffe model efficiency coefficient, E_{ns} (Nash and Sutcliffe, 1970), which is calculated as:

$$E_{ns} = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} \left(O_i - \overline{O}\right)^2}, \quad (2)$$

where S_i and O_i are simulated and observed discharge values, respectively, $\bar{O}$ is the average observed discharge value, and n is the number of discharge values. The higher the E_{ns}, the better the performance of a model, and a perfect fit has an E_{ns} equal to one.

PLSR

PLSR is regarded as the second generation of multivariate analysis. It combines features from Principal

Table I. List of sensitive parameters selected for calibration, and ranges and values of major parameters used[a]

Parameters	Description	Ranges	
		Dong River basin	Puli River basin
$CN2^{1,1}$	Initial SCS CN II value	40.63–87.5	34.12–69.77
$ALPHA\_BF^{2,10}$	Baseflow alpha factor [days]	0–0.08	0–0.08
$GW\_DELAY^{9,4}$	Groundwater delay [days]	34–45	34–45
$CH\_N2^{4,2}$	Manning's n value for main channel	0–0.08	0–0.08
$CH\_K2^{7,5}$	Effective hydraulic conductivity [mm/hr]	5–13	5–13
$ALPHA\_BNK^{5,3}$	Baseflow alpha factor for bank storage [days]	0–1	0–1
$SOL\_AWC^{5,7}$	Available water capacity (mm H^2O/mm soil)	0.14–0.22	0.00295–0.00472
$SOL\_K^{3,6}$	Saturated hydraulic conductivity (mm/hr)	3.49–44.17	1.15–14.54
$SOL\_BD^{8,9}$	Moist bulk density (Mg/m^3 or g/cm^3)	1.27–1.44	0.05–0.06
$SFTMP^{10,8}$	Snowfall temperature [^{o} C]	−0.3–1	−0.3–1

[a] Superscripts of each parameter are their corresponding rankings from the sensitivity analysis of runoff

Component Analysis (PCA) and Multiple Linear Regression (MLR) (Fornell, 1982; Wold *et al*., 1983; Barclay *et al*., 1995). PLSR has accuracy advantages over standard MLR through the application of cross validation, when independents have high correlations, and the number of independents is more than the number of sample points. In this study, the relationship between the land use and runoff coefficients was calculated using the PLSR method, using Minitab 15 statistical software. The general formula can be expressed as follows:

$$Y = b_0 + \sum_{i=1}^{m} b_i X_i. \tag{3}$$

The percentages of land use types by area for 2000, 2005, and 2010 in two basins were used for model fitting. Here, X_i ($i = 1, 2, \ldots 6$) and Y represent percentages of six land use types by area and runoff coefficients, respectively. m is the number of variables in the regressions, b_i represents the regression coefficients, and b_0 is constant in the regressions. Analyzing the above data, the PLSR model can be established using the following operation. First, a model combining all land use types is set up (one six-variable model). Water bodies and construction areas are then removed because of their small contribution to total area (<1%), and a model including the four remaining land use types is established (one four-variable model). Six further models are established by arbitrarily choosing two land use types from the remaining four (six two-variable models). Last, six models are established for each of the six land use types (six single-variable models). In all, 14 PLSR models are obtained.

Data preparation

Shuttle Radar Topography Mission Digital Elevation Model (DEM) data at a resolution of 90 m were used to extract the topographic information needed for hydrological modeling, such as slope and channel width. Land use maps were developed from Landsat 5 TM data (Figures 2 and 3). Overall, the predominant land use type was forest. Soils types were identified from the 1:1 000 000-scale Chinese national soil map (Figure 4). Original digital spatial data from the DEM, the land use map, the soil type map, and the river drainage map were converted to the same spatial resolution. Simulations were performed at a spatial resolution of 100 m, but because of the precision of DEM data, and the impacts of human activity, the river network generated by the DEM may differ slightly from the real stream structure. Consequently, we revised the river network derived from the DEM data using a 1:250 000-scale drainage map, by removing the vector lines of stream banks and lake shorelines. All input spatial data were processed into a uniform geographic coordinate and projection format, to meet the requirements of the SWAT simulation using the Albers equivalent conical projection.

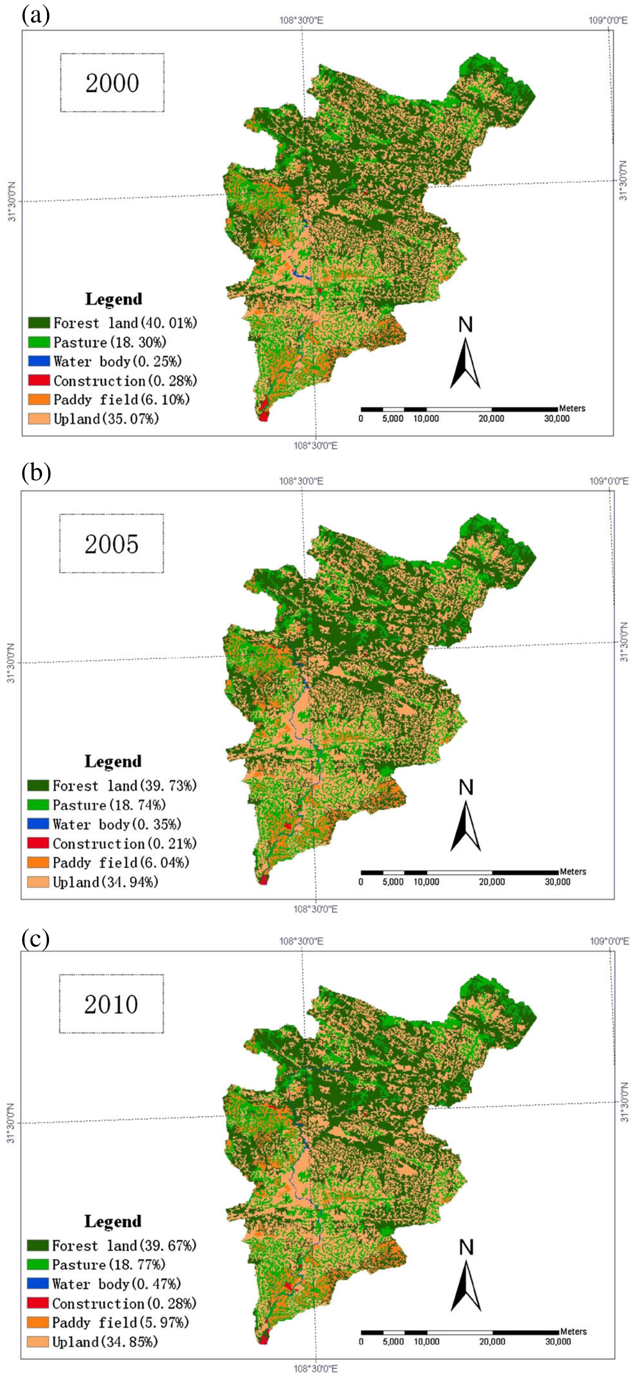

Figure 2. Land use map of Dong River basin in 2000 (a), 2005 (b), and 2010 (c)

Meteorological and hydrological data used to construct the SWAT model were obtained from the State Meteorological Administration and Hydrological Statistical Yearbook. Daily climate inputs to the model were precipitation, minimum and maximum air temperature, and wind speed. These input data were collected from meteorological stations within the study area (Figure 1). Additional climate variables such as solar radiation and dew-point temperature were produced from a weather generator using values from the nearest standardized weather station. Daily discharge data were collected from gauging stations in the two basins (Figure 1).

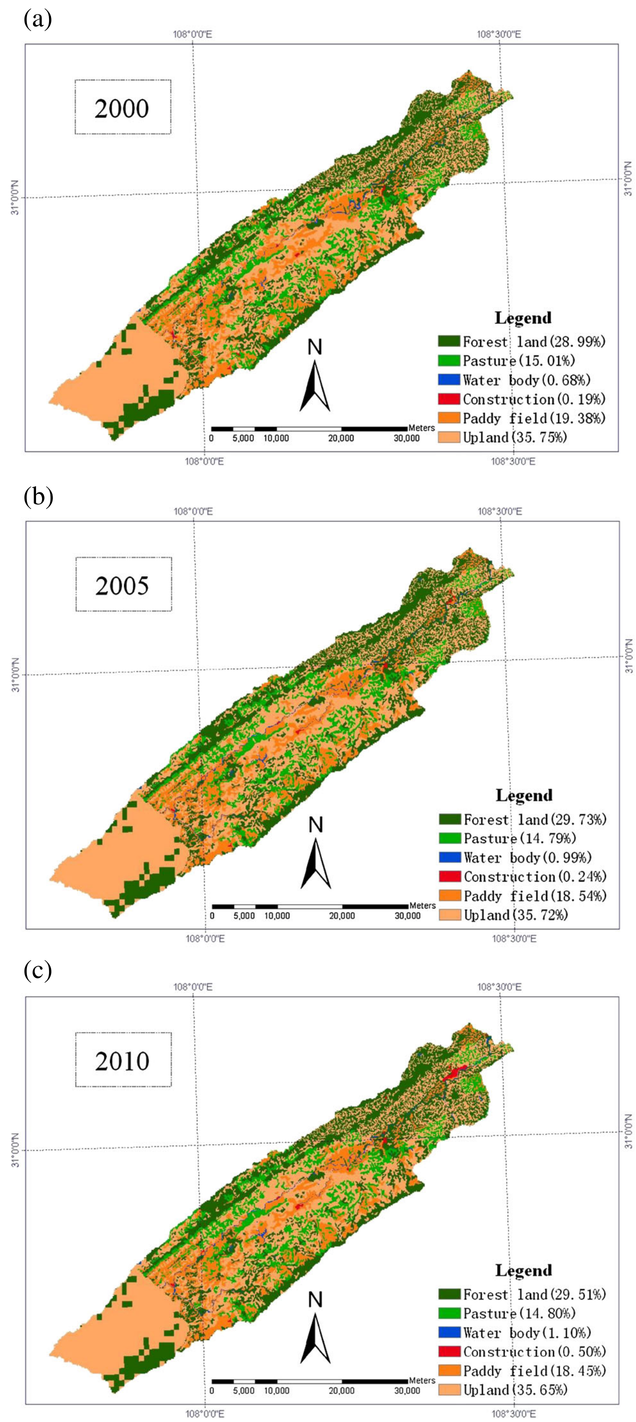

Figure 3. Land use map of Puli River basin in 2000 (a), 2005 (b), and 2010 (c)

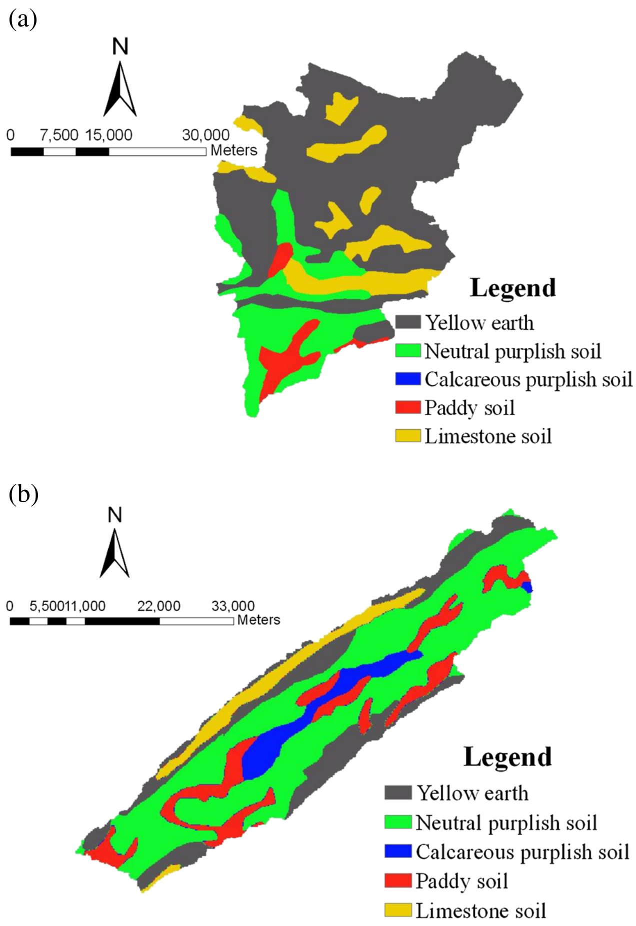

Figure 4. Soil maps of Dong River basin (a), and Puli River basin (b)

Table II. Results of simulation of runoff for calibration and validation periods in the two basins

Station	Simulation period	E_{ns}	r^2
Wenquan station (Dong River basin)	Calibration for 2002–2006 Validation for 2007–2010	0.94 0.98	0.94 0.99
Yujia station (Puli River basin)	Calibration for 2001–2002 Validation for 2003–2005, 2010	0.80 0.93	0.81 0.87

RESULTS AND DISCUSSION

Model calibration and validation

Based on the natural river network, basin topography, and rainfall station distribution, the Dong and Puli River basins were divided into 30 sub-basins, as shown in Figure 1. These basins were further divided into 181 HRUs according to land use, soil properties, and slope. Initial values of parameters for model calibration were obtained based on input maps and the database. Table II shows the calibration and validation results for the two basins. The E_{ns} and determination coefficients between observed and simulated values (r^2) for the runoff simulation at Wenquan in the Dong River basin were 0.94 and 0.94 in the calibration period, and 0.98 and 0.99 in the validation period. Corresponding values at Yujia in the Puli River basin were 0.80 and 0.81 in the calibration period, and 0.93 and 0.87 in the validation period.

Figure 5 shows the scatterplot and time series discharge hydrographs generated during the validation period at Wenquan and Yujia stations for the Dong and Puli River basins, respectively, demonstrating the SWAT model's ability to simulate the temporal variations in daily runoff generation. For the validation period, simulation results at Wenquan were better than those at Yujia. Overall, simulated results in both the Dong and Puli River basins agreed well with observed data.

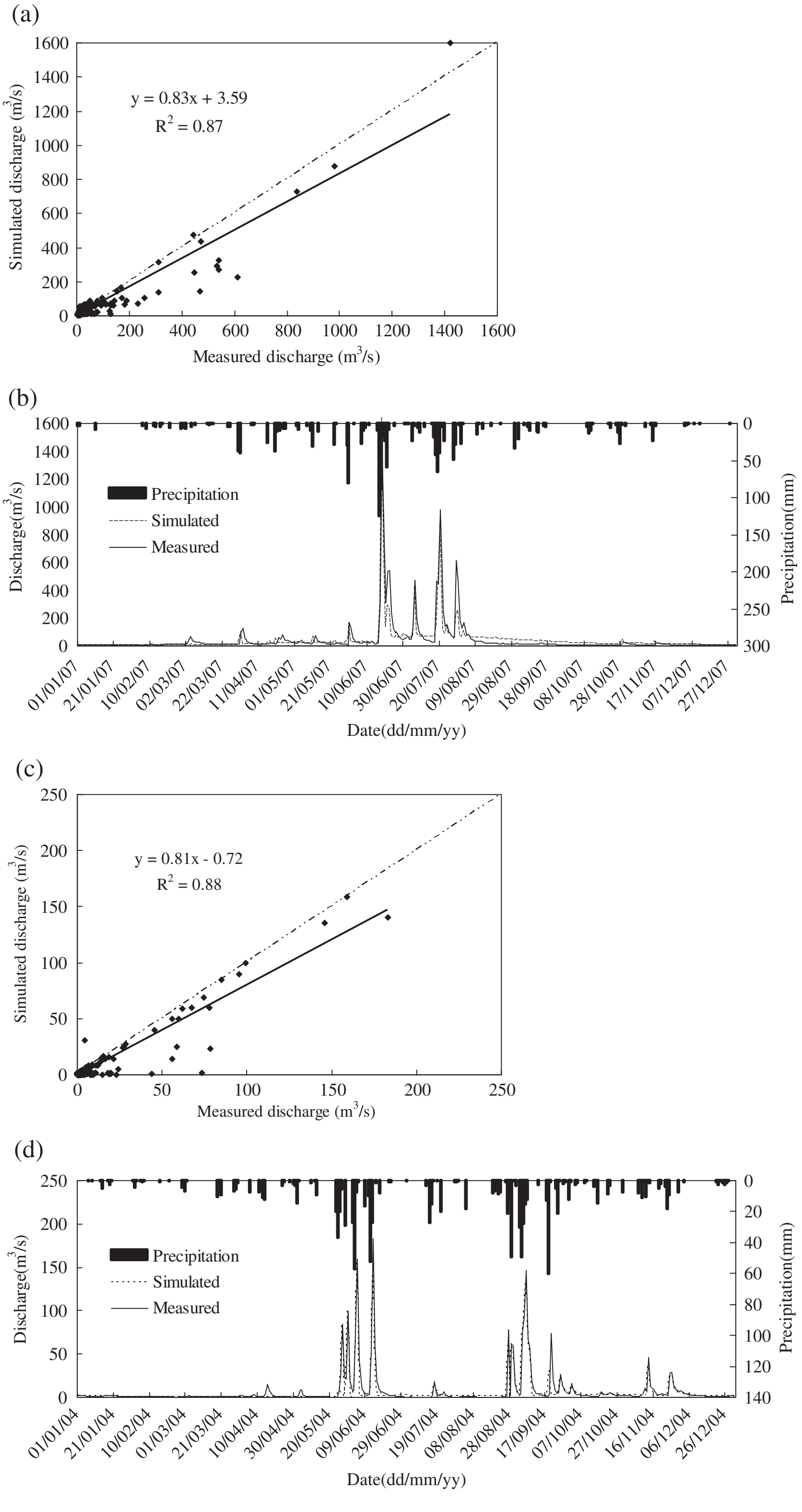

Figure 5. Scatterplot (a), and time series (b) of simulated and observed daily discharge during 2007 at Wenquan station in the Dong River basin; scatterplot (c), and time series (d) of simulated and observed daily discharge during 2004 at Yujia station in the Puli River basin

Land use changes

The predominant land use types in the Dong River basin in 2000 were forest (40.0%), upland (35.1%), and pasture (18.3%). For the Puli River basin, predominant land uses were upland (35.8%), forest (29.0%), and paddy field (19.4%). A detailed description of changes in land use in the study area from 2000 to 2010 is presented in Table III.

During the 10-year period from 2000 to 2010, land use patterns in the Dong and Puli River basins changed greatly, and the proportions of the previously predominant land use types (paddy field, upland, and forests) changed as well. Pasture increased in the Dong River basin and in the Puli River basin, but there was a decline in paddy field, upland, and pasture areas. However, the

Table III. Land use area and percentage changes relative to 2000

Land use type		2005		2010	
		Dong River basin	Puli River basin	Dong River basin	Puli River basin
Paddy field	Area (km^2)	82.97	176.96	81.93	176.12
	%	−0.89	−4.34	−2.13	−4.79
Upland	Area (km^2)	479.86	340.90	478.63	340.25
	%	−0.37	−0.11	−0.63	−0.30
Forest	Area (km^2)	545.65	283.78	544.90	281.69
	%	−0.70	+2.55	−0.84	+1.80
Pasture	Area (km^2)	257.44	141.18	257.76	141.24
	%	+2.43	−1.43	+2.56	−1.39
Water body	Area (km^2)	4.80	9.43	6.52	10.46
	%	+38.57	+44.83	+88.25	+60.59
Construction	Area (km^2)	2.83	2.25	3.80	4.74
	%	−26.50	+25.70	−1.22	+164.66

proportion of forest in the Puli River basin increased slightly (1.8%). Although water bodies and urban or construction areas accounted for only a small proportion of the total study area, these two land use types underwent the largest changes during the study period. Water body area increased by 88.3% and 60.6% in the Dong and Puli River basins, respectively, and urban area increased by 164.7% in the Puli River basin (Table III). These changes were clearly a consequence of the TGR. Its construction took place from 1993 to 2009, with a final water level of 175 m above sea level and a total storage capacity of 39.3 billion m^3. In June 2003, the water level reached 135 m above sea level, and with its continued rise, 1.13 million inhabitants had to be relocated, many to the upstream river basins. From 2000 to 2004, a further 96 000 residents from the region along the upper reaches of the TGR were relocated. As a result of the reservoir and the shifting population, the study area has undergone a continuous increase in the area of water bodies and urbanization (Zhang *et al*., 2009a).

The land use change matrix from 2000 to 2010 for the Dong River basin is presented in Table IV. As also shown in Table III, the total area of upland in the basin fell, while the area of pasture increased. The main land use change from 2000 to 2010 in this basin was from upland to pasture. About 13 km^2 of upland was transformed into pasture, corresponding to an increase in pasture area of 2.56% by 2010. In the Puli River basin, there were three main land use changes over this period. One was the conversion to construction or urban land use from other land use types, and the others were the conversion from paddy field to upland and water bodies, and from upland to forest (Table V). Overall, land use was strongly influenced by socio-economic development and the TGR project in the basins upstream of TGR.

Table IV. Land use area change matrix from 2000 to 2010 for the Dong River basin (km^2)

Land use type	Pasture	Water body	Paddy field	Upland	Forest	Construction	2000 total	Ratio (%)
Pasture	243.20	1.44	0.55	6.12	0.22	0.23	251.75	18.30
Water body	0.39	1.99	0.10	0.78	0.06	0.00	3.33	0.24
Paddy field	0.81	0.44	80.06	0.30	0.57	1.31	83.50	6.07
Upland	12.75	1.24	0.04	454.80	11.24	0.40	480.46	34.92
Forest	0.62	1.09	0.92	14.40	535.79	0.09	552.91	40.19
Construction	0.33	0.17	0.02	1.26	0.29	1.72	3.80	0.28
2010 total	258.10	6.36	81.69	477.67	548.18	3.75	1375.75	100.00
Ratio (%)	18.76	0.46	5.94	34.72	39.85	0.27	100.00	

Table V. Land use area change matrix from 2000 to 2010 for the Puli River basin (km^2)

Land use type	Pasture	Water body	Paddy field	Upland	Forest	Construction	2000 total	Ratio (%)
Pasture	139.90	0.38	0.13	0.69	2.01	0.00	143.12	12.27
Water body	0.05	3.72	1.58	0.30	0.64	0.08	6.38	0.55
Paddy field	0.64	3.11	171.80	6.23	2.19	1.30	185.26	15.88
Upland	0.27	1.32	1.90	497.90	5.06	0.59	507.04	43.47
Forest	0.13	1.63	0.77	0.88	317.76	1.77	322.94	27.68
Construction	0.17	0.10	0.33	0.02	0.12	1.01	1.76	0.15
2010 total	141.16	10.26	176.51	506.03	327.79	4.74	1166.48	100.00
Ratio (%)	12.10	0.88	15.13	43.38	28.10	0.41	100.00	

Water balance changes between 2000 and 2010

Table VI summarizes the water balance in the two basins according to simulation results for the calibration and validation periods. Each scenario (2000, 2005, and 2010) was simulated on a daily basis, and an annual mean calculated for the entire simulation period from 2000 to 2010. Table VI shows the mean water balance changes in the two basins, relative to land use in 2000. To identify differences in hydrological processes with various types of land use, we summarized the water balance associated with land use types using annual means (Table VII). Construction areas and water bodies were excluded from these simulations because their proportion across the entire study area was minimal, with only a small effect on water balance. The average ET in 2010 was 600 mm and 581 mm for the Dong and Puli River basins, respectively (Table VI). Although the annual ET in the basins did not change much, there was a gradual decrease (Table VI). Compared with the land usage in 2000, by 2005 mean annual ET had decreased by 0.70% and 1.91% in the Dong and Puli River basins, respectively. By 2010, ET had decreased by 2.13% and 2.41% in the Dong and Puli River basins, respectively (Table VI). There were larger ET changes in the Puli River basin for both the 2005 and 2010 land use scenarios. Precipitation and land use are the two dominant factors controlling ET. Precipitation did not change significantly between period 1 (2000–2005) and period 2 (2005–2010), which indicates that a change in land use may be the reason for the decline in ET.

Table VI. Water balances for the two basins in 2010, and changes (%) relative to the 2000 land use conditions

Water balance		Basin: Dong River basin	Basin: Puli River basin
ET (mm)		600	581
ET changes (%)	2005	−0.7	−1.91
	2010	−2.13	−2.41
Quickflow (mm)		434	227
Quickflow changes (%)	2005	7.86	18.28
	2010	5.37	15.1
Baseflow (mm)		450	262
Baseflow changes (%)	2005	5	11.95
	2010	2.68	14.47
Infiltration (mm)		490	290
Infiltration changes (%)	2005	4.59	9.52
	2010	3.68	11.8

Quickflow (surface runoff plus subsurface runoff) was 434 mm and 227 mm in the Dong and Puli River basins, respectively (Table VI). The change in quickflow in the Puli River basin was more obvious than in the Dong River basin. In the Puli River basin, annual quickflow increased by 18.25% and 15.1% under 2005 and 2010 land use conditions, relative to 2000. Corresponding values for the Dong River basin were 7.86% and 5.37%.

Changes in infiltration were similar in the both basins. Infiltration in the Dong River basin increased by 4.59% and 3.68% under 2005 and 2010 land use conditions, respectively, and by 9.52% and 11.8% in the Puli River basin. Construction decreased in the Dong River basin and increased in the Puli River basin, but this did not produce a difference in infiltration trends (Tables II and VI).

Baseflow from the shallow aquifer contributes to the main river channel and river reaches. In the SWAT model, baseflow is programmed to enter the reaches only if the amount of water stored in the aquifer exceeds a threshold value specified by users. Baseflow in both basins clearly increased under 2005 and 2010 land use conditions, and the trend for baseflow was similar to infiltration. The land use types producing the greatest baseflow were forest and pasture, which was also the case for infiltration (Table VII). Mean annual baseflow depends on mean annual infiltration to some great extent; therefore, variations in baseflow mostly originated from variations in infiltration, caused by the land use changes.

According to Table III, upland occupied large percentages of land in both basins, but the area decreased substantially from 2000 to 2010. This may have contributed to the decrease in ET. Paddy fields are thought to have a positive effect on runoff generation because of the absence of runoff detention (e.g. Wang *et al.*, 2011). However, the paddy field area in the study area was much smaller than the area of upland and forest (Table III). Therefore, the increase in runoff generation by upland might be exaggerated relative to the decrease related to the area of paddy fields. A thorough investigation of the effect of paddy fields on runoff

Table VII. Simulated results of annual average water balance for different land uses from 2000 to 2010

Water balance	Paddy field: Dong River basin	Paddy field: Puli River basin	Upland: Dong River basin	Upland: Puli River basin	Forest: Dong River basin	Forest: Puli River basin	Pasture: Dong River basin	Pasture: Puli River basin
ET (mm)	564.8	535.6	565.1	534.4	612.0	590.4	651.6	630.4
Infiltration (mm)	417.7	261.8	463.8	273.5	529.2	289.1	479.5	275.8
Surface runoff (mm)	393.8	237.8	396.4	213.3	197.4	116.8	283.9	155.4
Groundwater flow (mm)	382.4	236.0	425.8	261.6	487.3	246.9	440.9	249.1
Subsurface lateral flow (mm)	34.6	32.7	97.5	47.5	192.0	98.9	98.1	82.4
Transmission losses (mm)	0.9	0.8	0.8	0.8	0.3	0.4	0.6	0.5

generation is required to clarify this. Previous studies have shown that forest and pasture tend to decrease runoff, because of precipitation interception, high ET, high infiltration capacity, and soil porosity (Liu *et al.*, 2006; Wang *et al.*, 2010; Sriwongsitanon and Taesombat, 2011). In the Dong River basin, although the decrease in the percentage of forest area was small, its absolute area was large (Table III). This may also partly explain the decrease in ET, and the related increase in runoff. Despite the increase in pasture within the Dong River basin (associated with an increase in ET), the overall decrease in ET suggested a predominant role of paddy fields, upland, and forest in runoff generation. For the Puli River basin, the increase in forest area was small, as was its absolute area. Therefore, paddy fields, upland, and pasture seemed to play a more important role. Although the percentage of water bodies and construction increased significantly during the study period in both basins, the absolute area of these two land use types was much smaller than the others, and they had a smaller impact on the runoff generation. On the other hand, since these two land use types could contribute to an increase in runoff generation by shortening flood retention time, if they consistently increase in the future as a consequence of human activity, as they have over the past 10 years, their influence might become more significant.

The above analyses demonstrate the effect of land use changes on the hydrological processes of basins upstream of the TGR. ET, quickflow, infiltration, and baseflow changed greatly between 2000 and 2010, in both upstream basins. With regards to water balance, the absolute values of hydrological components in the Dong River basin, with its larger basin area and higher mean annual precipitation, were larger than in the Puli River basin (Table VI). However, despite the slight difference in land use changes between the basins over the study decade, the water balance in the Puli River basin showed larger changes than in the Dong River basin (Table VI).

Quantitative analysis of land use effect on runoff coefficients

Our analysis shows the qualitative effect of land use on runoff generation in the study area. However, to explicitly assess the impact of land use changes on runoff generation, the impact of rainfall variation should first be excluded. The runoff coefficient is used here as an indicator for hydrological cycles in these basins. The runoff coefficient is widely used as a diagnostic variable to represent runoff generation in a catchment, and as an important parameter in hydrological design. In this study, the runoff coefficient was calculated as the ratio of annual runoff to annual rainfall. Standardized regressions of the runoff coefficient models derived using PLSR are presented in Table VIII. Non-standardized regression

Table VIII. Fitted models and model performance of runoff coefficient

Number of variables (m)	Predicator variables	Standardized regression coefficient	F statistic	P-value	R^2	Models	
1	Paddy field(X_1)	−0.727	4.48	0.1	0.53	$Y = 0.641 - 1.357X_1$	(1)
	Upland(X_2)	−0.754	5.26	0.08	0.57	$Y = 8.734 - 23.385\ X_1$	(2)
	Forest(X_3)	0.718	4.25	0.11	0.52	$Y = -0.097 + 1.645\ X_1$	(3)
	Pasture(X_4)	0.693	3.7	0.13	0.48	$Y = -0.264 + 4.404\ X_1$	(4)
	Water (X_5)	−0.363	4.61	0.12	0.52	$Y = 0.561 - 13.751\ X_1$	(5)
	Construction(X_6)	0.116	4.55	0.12	0.51	$Y = 0.435 + 13.442\ X_1$	(6)
2	Paddy field(X_1)	−0.366	4.94	0.09	0.55	$Y = 4.720 - 0.684X_1 - 11.781X_2$	(7)
	Upland(X_2)	−0.38					
	Paddy field(X_1)	−0.364	4.37	0.11	0.52	$Y = 0.272 - 0.679\ X_1 + 0.823\ X_2$	(8)
	Forest(X_2)	0.359					
	Paddy field(X_1)	−0.365	4.11	0.11	0.51	$Y = 0.187 - 0.681X_1 + 2.210\ X_2$	(9)
	Pasture(X_2)	0.35					
	Upland(X_1)	−0.381	4.84	0.09	0.55	$Y = 4.360 - 11.816\ X_1 + 0.831\ X_2$	(10)
	Forest(X_2)	0.363					
	Upland(X_1)	−0.38	4.48	0.1	0.53	$Y = 4.262 - 11.777\ X_1 + 2.218\ X_2$	(11)
	Pasture(X_2)	0.349					
	Forest(X_1)	0.361	4	0.12	0.50	$Y = -0.183 + 0.826\ X_1 + 2.212\ X_2$	(12)
	Pasture(X_2)	0.348					
4	Paddy field(X_1)	−0.183	4.47	0.1	0.53	$Y = 2.269 - 0.342X_1 - 5.897\ X_2 + 0.414\ X_3 + 1.110x_4$	(13)
	Upland(X_2)	−0.19					
	Forest(X_3)	−0.18					
	Pasture(X_4)	0.175					
6	Paddy field(X_1)	−0.162	3.96	0.12	0.50	$Y = -168.172 - 7.488X1 + 428.750X2 - 6.619\ X_3 + 113.978\ X_4 + 197.592\ X_5 + 21.698X_6$	(14)
	Upland(X_2)	−0.168					
	Forest(X_3)	0.16					
	Pasture(X_4)	0.154					
	Water(X_5)	−0.081					
	Construction(X_6)	0.026					

coefficients for respective variables, and standardized regression coefficients, are also shown. The standardized coefficients can be used to indicate the sign and magnitude of predictor variables on response variables. Using standardized regression coefficients, we can directly compare the magnitude or importance of relationships between predictors and responses, whether or not the predicators are of the same scale.

All 14 models show a negative effect of upland and paddy field on the runoff coefficient. Models (2), (7), and (10) were relatively significant, with $P < 0.1$. The R^2 for model (2) and (7) was 0.57 and 0.55, respectively. All these models have upland as the predictor, demonstrating its importance to runoff generation in the study area. Paddy fields are thought to shorten flood detention time and to positively affect the runoff coefficient. Our model results seem to contradict this. In fact, irrigation of paddy fields in the study area primarily used water from the river systems and therefore directly reduced runoff.

Forest and pasture were found to have a positive effect on the runoff coefficient. The only exception was model (13), in which forest had a negative effect. This result appears to contrast with the findings of Sriwongsitanon and Taesombat (2011). They did find forest area to have a negative effect in almost all sub-basins for smaller flood conditions. However, they also concluded that forest effects on runoff generation could be positive if the basin had high antecedent soil moisture or was saturated. Since forest tends to retain more water because of its large soil water capacity, it can produce more runoff compared to non-forest areas for large flood conditions. However, Sriwongsitanon and Taesombat (2011) focused on the peak flow of flood rather than overall runoff. In this study area, precipitation was continuously high during the summer monsoon season, possibly causing high antecedent soil moisture. Therefore, forest in the area had a positive effect on runoff generation. This explanation is also applicable to pasture, because it is mixed with forest in the study area.

Statistically, the models did not yield significant results, with p-values exceeding 0.05. The R^2 was less than 0.6, indicating a poor predictive ability of the model (Table VIII). This poor performance may arise from the PLSR method, in spite of its advantage in solving multicollinearity. In model applications, the ordinary least squares technique performances are better than PLSR when there is no multicollinearity issue. Moreover, the poor model performance may be partly a consequence of uncertainty in estimating surface runoff using a hydrological model.

CONCLUSIONS

Long-term land use changes have been shown to have large effects on hydrological processes in river basins. In this study, three land use scenarios (2000, 2005, and 2010) were used to evaluate the impact of land use changes on hydrological processes in basins upstream of the TGR, using a calibrated and validated version of the SWAT model. Between 2000 and 2010, there was a large redistribution of the human population in the TGR area. As a consequence, land use patterns in the basins upstream of the TGR changed greatly. Areas of forest and pasture increased during the period, but paddy fields and upland decreased in both basins. Hydrological processes were markedly affected by land use changes in the basins upstream of the TGR. From 2000 to 2010, ET decreased by 2.13% and 2.41% in the Dong and Puli River basins, respectively. Quickflow, infiltration, and baseflow increased to varying degrees over the same period. Results from this study indicate that the water balance in the Puli River basin, with a smaller area, showed greater changes compared to the Dong River basin, demonstrating different sensitivity to land use changes for the two river basins. However, average slope, major soil type, and the initial proportion of land use types by area were different for the two river basins. Further study is therefore necessary to investigate which are the primary factors underlying the sensitivity to land use changes.

The PLSR modeling of runoff coefficient showed that upland was the land use type that had the most impact on runoff generation in the study area. The regression models also verified the negative effect of paddy fields on runoff generation in the study area, indicating irrigation water use effects on river runoff. Forest area was found to have a positive effect in most of the models. As far as the authors know, this is the first study to investigate the relationships between land use type and runoff coefficients in the river basins upstream of the TGR. Despite of the absence of physical evidence, these empirical relationships could provide useful predictive information for land use policy makers and water resource management. On the other hand, the results modeled by the PLSR method were not significant, indicating a relatively poor predictive ability. The regression model uncertainties might arise in part from systematic errors in the land use type classification generated by the TM data and the runoff values generated by the SWAT model. Further research should focus on improvements in hydrological models, incorporating land use changes and climate variables determined using multi-source satellite data. This detailed information could result in more accurate data on land use changes and runoff and could improve the predictive ability of regression models. Regression models could also be improved by using a hierarchical approach known as Hierarchical Cluster-based PLSR (HC-PLSR). Tøndel *et al.* (2011) found that, in terms of explained variance and prediction accuracy, the HC-PLSR method was more accurate than PLSR.

ACKNOWLEDGEMENT

This research was supported by the Chinese National Special Science and Technology Program of Water Pollution Control and Treatment (No. 2009ZX07528-003).

REFERENCES

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modeling and assessment. Part 1: Model development. *Journal of the American Water Resources Association* **34**(1): 73–89.

Barclay D, Higgins C, Thompson R, 1995. The partial least squares (PLS) approach to causal modeling: personal Computer Adoption and use as an illustration. *Technology Studies* **2**(2): 285–309.

Fohrer N, Haverkamp S, Eckhardt K, Frede HG. 2001. Hydrologic response to land use changes on the catchment scale. *Physics and Chemistry of the Earth Part B* **26**: 577–582.

Fornell C. 1982. A Second Generation of Multivariate Analysis. In *Methods*, vol. 1. Praeger: New York.

Hibbert AR. 1967. Forest treatment effects on water yield. In *International Symposium on Forest Hydrology*, Sopper WE, Lull HW (eds). Pergamon: Oxford; 813.

Hollis GE. 1975. The effect of urbanization on flood of different recurrence interval. *Water Resources Research* **11**(3): 431–435.

Hundecha Y, Bardossy A. 2004. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *Journal of Hydrology* **292**: 281–295.

Liu YB, Gebremeskel S, Smedt FD, Hoffmann L, Pfister L. 2006. Predicting storm runoff from different land-use classes using a geographical information system-based distributed model. *Hydrological Processes* **20**(3): 533–548.

Lørup JK, Refsgaard JC, Mazvimavi D. 1998. Assessing the effect of land use change on catchment runoff by combined use of statistical tests and hydrological modeling: case studies from Zimbabwe. *Journal of Hydrology* **205**: 147–163.

Merz R, Bloschl G. 2009. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research* **45**: W01405. DOI: 10.1029/2008WR007163

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

Schreider SY, Jakeman AJ, Letcher RA, Nathan RJ, Neal BP, Beavis SG. 2002. Detecting changes in streamflow response to changes in non-climatic catchment conditions: farm dam development in the Murray–Darling basin, Australia. *Journal of Hydrology* **262**: 84–98.

Sriwongsitanon N, Taesombat W. 2011. Effects of land cover on runoff coefficient. *Journal of Hydrology* **410**: 226–238.

Sullivan A, Ternan JL,Williams AG. 2004. Land use change and hydrological response in the Camel catchment, Cornwall. *Applied Geography* **24**: 119–137.

Tøndel K, Indahl UG, Gjuvsland AB, Vik JO, Hunter P, Omholt SW, Martens H. 2011. Hierarchical Cluster-based Partial Least Squares Regression (HC-PLSR) is an efficient tool for metamodelling of nonlinear dynamic models. *BMC Systems Biology*. DOI: 10.1186/1752-0509-5-90

Van Liew MW, Veith TL, Bosch DD, Arnold DD. 2007. Suitability of SWAT for the Conservation Effects Assessment Project: Comparison on USDA Agricultural Research Service Watersheds. *Journal of Hydrologic Engineering* **12**(2): 173–189.

Wang GQ, Hapuarachchi P, Ishidaira H, Kiem AS, Takeuchi K. 2009. Estimation of soil erosion and sediment yield during individual rainstorms at catchment scale. *Water Resources Management* **23**: 1447–1465.

Wang GQ, Jiang H, Xu ZX, Wang LJ, Yue WF. 2011. Evaluating the effect of land use changes on soil erosion and sediment yield using a grid-based distributed modelling approach. *Hydrological Processes* DOI: 10.1002/hyp.9193

Wang GQ, Yu JS, Shrestha S, Ishidaira H, Takeuchi K. 2010. Application of a distributed erosion model for the assessment of spatial erosion patterns in the Lushi catchment, China. *Environmental Earth Science* **61**: 787–797.

Wold S, Albano C, Dunn M, Esbensen K, Hellberg S, Johansson E, Sjöström M. 1983. *Pattern Regression Finding and Using Regularities in Multivariate Data*. London: Analysis Applied Science Publication.

Wooldridge S, Kalma J, Kuczera G. 2001. Parameterisation of a simple semi-distributed model for assessing the impact of land use on hydrologic response. *Journal of Hydrology* **254**: 16–32.

Zhang JX, Liu ZJ, Sun XX. 2009a. Changing landscape in the Three Gorges Reservoir Area of Yangtze River from 1977 to 2005: Land use/land cover, vegetation cover changes estimated using multi-source satellite data. *International Journal of Applied Earth Observation and Geoinformation* **11**: 403–412.

Zhang XS, Srinivasan R, Bosch D. 2009b. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *Journal of Hydrology* **374**(3–4): 307–317.

APPENDIX

Nomenclature used in this paper

E_{ns}	Nash–Sutcliffe efficiency coefficient
ET	Evapotranspiration (mm)
HC-PLSR	Hierarchical Cluster-based Partial Least Squares Regression
HRU	Hydrologic Response Units
MLR	Multiple Linear Regression
PCA	Principal Component Analysis
PLSR	Partial Least Squares Regression
SWAT	Soil and Water Assessment Tool
SWAT-CUP	SWAT Calibration and Uncertainty Programs
TM	Thematic Mapper
TGR	Three Gorges Reservoir