Journal of Environmental Sciences Vol.16, No.4, pp.646—650, 2004

ISSN 1001—0742
CN11—2629/X

Article ID: 1001-0742(2004)04-0646-05 **CLC number**: X11 **Document code**: A

A distributed non-point source pollution model: calibration and validation in the Yellow River Basin

HAO Fang-hua, ZHANG Xue-song, YANG Zhi-feng*

(Institute of Environmental Sciences, Beijing Normal University, Beijing 100875, China. E-mail: zfy@bnu.edu.cn)

Abstract: The applicability of a non-point source pollution model—SWAT(soil and water assessment tools) in a large river basin with high sediment runoff modulus(770 t/km^2 in the Yellow River) was examined. The basic database, which includes DEM, soil and landuse map, weather data, and land management data, was established for the study area using GIS. A two-stage "Brute Force" optimization method was used to calibrate the parameters with the observed monthly flow and sediment data from 1992 to 1997. In the process of calibration automated digital filter technique was used to separate direct runoff and base flow. The direct runoff was firstly calibrated, and the base flow, then the total runoff was matched. The sediment yield was calibrated to match well. Keeping input parameters set during the calibration process unchanged, the model was validated with 1998—1999's observed monthly flow and sediment. The evaluation coefficients for simulated and observed flow and sediment showed that SWAT was successfully applied in the study area: relative error was within 20%, coefficient of determination and Nash-Suttcliffe simulation efficiency were all equal to or above 0.70 during calibration and validation period.

Keywords: Yellow River; SWAT; sediment; flow; non-point source pollution

Introduction

Water resources depletion, including not only water quantity scarcity but also water quality deterioration, is a serious issue in the Yellow River Basin(Zhu, 1999; Xi, 1996). Xia *et al*. found decreasing ratio of point source pollution to non-point source pollution through the analysis of water quality data from 1990 to 1999(Xia, 2001), pointing to an increasing trend of non-point source pollution in the Yellow River Basin. With water pollution control being mainly concentrated on industrial point sources, there is a lack of attention and study on non-point source pollution in the Yellow River Basin. And there was no non-point source pollution study that has been conducted using physically-based distributed parameter model. SWAT(soil and water assessment tools)—a distributed non-point source pollution model has been used in several projects by EPA, NOAA, NRCS and others to estimate the off-site impacts of climate and management on water use, non-point source loadings, and has been extensively validated across United States for stream flow and sediment yields(Arnold, 1998). Therefore SWAT was selected as a tool to study non-point source pollution in the Yellow River Basin.

Luohe River is the biggest tributary of Yellow River between Xiaolangdi and Huayuankou, and contributes greatly to the water quality pollution of the mainstream. With the point source pollution being controlled, non-point source pollution would be the main contributor to water pollution. It is necessary to study non-point source pollution in this area. Flow and sediment modeling is the basis of non-point source pollution prediction. The soil erosion is serious in the Yellow River Basin. For instance, in the upstream watershed of Luohe River the rate of soil erosion is about 2000 t/(km^2·a) and sediment runoff modulus reaches 770 t/(km^2·a). While SWAT was often used in watersheds with low sediment runoff modulus, for example, Santhi *et al*. applied SWAT in the 4277 km^2 Bosque River Watershed(Santhi, 2001), whose sediment runoff modulus is normally below 100 t/(km^2·a). The objective of this study is to examine whether SWAT model could be successfully validated for flow and sediment in the Yellow River Basin, and the upstream watershed of Luohe River that is upper than Lushi Hydrological Station, which was called Lushi Watershed below, was selected as the study area.

1 Materials and methods

1.1 SWAT model description

SWAT is a hydrologic/water quality model developed by United States Department of Agriculture—Agricultural Research Service (USDA-ARS) (Santhi, 2001). It is a continuous time model that operates on a daily time step. The objective in model development was to predict the impact of management on water, sediment, and agricultural chemical yields in large ungaged basins. To satisfy the objective, the model (a) is physically based (calibration is possible on ungaged basins); (b) uses readily available inputs; (c) is computationally efficient to operate on large basins in a reasonable time; and(d) is continuous in time and capable of simulating long periods for computing the effects of management changes. The subbasin/subwatershed

Foundation item: The National High-Tech R&D Program(973) of China(No. G1999043605); * Corresponding author components of SWAT can be placed into eight major components: hydrology, weather, erosion/sedimentation, soil temperature, plant growth, nutrients, pesticides, and land management. SWAT uses a modified version of the SCS CN method for predicting surface runoff yield(USDA-SCS):

$$Q = \frac{(R-0.2S)^2}{(R+0.8S)}, \qquad R > 0.2S \tag{1}$$

$$Q = 0, \qquad R \leqslant 0.2S \tag{2}$$

where Q is the daily surface runoff(mm), R is the daily rainfall(mm), and S is a retention parameter. The retention parameter, S, varies(a) among watersheds because soils, landuse, management, and slope all differ from each other and(b) with time because of changes in soil water content. The parameter S is related to curve number(CN) by the SCS equation:

$$S = 25.4\left(\frac{1000}{CN} - 10\right). \tag{3}$$

Peak runoff rate is estimated by using the modified rational formula. Estimation of percolation is conducted by using a storage routing technique. This is based on the assumption that percolation occurs when the field capacity of the soil is exceeded and if the layer below is unsaturated. The contribution of groundwater to stream flow is simulated by creating shallow aquifer.

Erosion and sediment yield are estimated for each subbasin with the Modified Universal Soil Loss Equation (MUSLE)(Williams, 1975):

$$Y = 11.8(Vq_p)^{0.56}(K)(C)(PE)(LS), \tag{4}$$

where Y is the sediment yield from the subbasin, V is the surface runoff column for the subbasin in m^3, q_p is the peak flow rate for the subbasin in m^3/s, K is the soil erodibility factor, C is the crop management factor, PE is the erosion control practice factor, and LS is the slope length and steepness factor.

Channel routing consists of flood and sediment routing. The flood routing model uses a variable storage coefficient method developed by Williams(Williams, 1969). Channel inputs include the reach length, channel slope, bankfull width and depth, channel side slope, flood plain slope, and Manning's n for channel and floodplain. Flow rate and average velocity are calculated using Manning's equation and travel time is computed by dividing channel length by velocity. Outflow from a channel is also adjusted for transmission losses, evaporation, diversions, and return flow. The channel sediment routing equation uses a modification of Bagnold's sediment transport equation that estimates the transport concentration capacity as a function of velocity(Bagnold, 1997):

$$CY_u = SPCON \times V^{SPEXP}, \tag{5}$$

where, CY_u is the sediment transport concentration capacity in g/m^3; $SPCON$ is the concentration capacity in g/m^3 at a velocity of 1 m/s; V is the flow velocity in m/s; and SPEXP is a constant in Bagnold's equation. The SWAT model either deposits excess sediment or reentrains sediments through channel erosion depending on the sediment load entering the channel.

1.2 Watershed description

Lushi watershed with an area of 4623 km^2 is characterized by mountainous landscape (Fig. 1). The Qinling Mountain is located to the south of the watershed, and the Huashan Mountain and the Yaoshan Mountain to the north. This area belongs to the warm temperate climate zone, and the annually average precipitation is between 800 mm and 1000 mm.

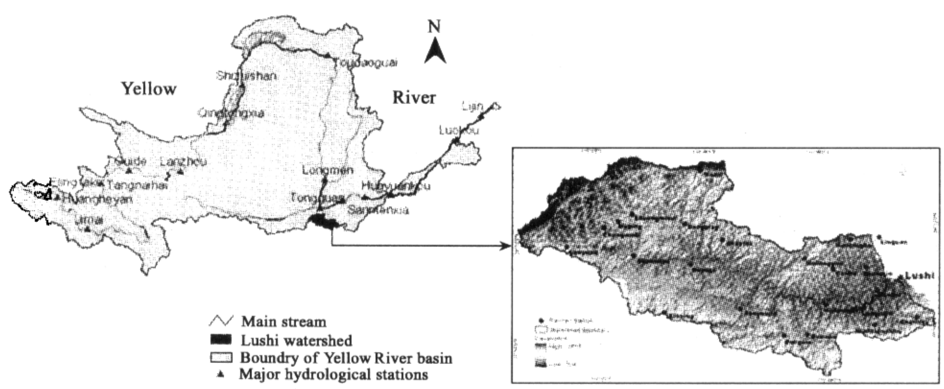

Fig. 1 Geographic position of the study area

Landuses in this watershed are mostly forest and cropland in the upper portion while cropland and pasture are wide spread in the lower portion. The major soil series are zonal from the low elevation to the high elevation area, accordingly the soil series change from Calcic Cinnamon Soils, Typic Cinnamon Soils, Typic Burozems to Clay Pan Yellow-brown Earths, the percentage of which in the entire watershed are about 27, 34, 38 and 1 respectively.

1.3 Establishment of basic database for Lushi watershed

With the support of ArcView GIS, the basic database for Lushi watershed was established, which mainly include topography, soil and landuse map, weather and land management data (Table 1). Initially, the watershed was delineated into subbasins using the digital elevation map. The delineated subbasin map, landuses, soils maps were overlaid. SWAT simulates different landuses in each subbasin. Winter wheat and summer maize in rotation was simulated on the cropland.

Table 1 Model input data sources for Lushi watershed

Data type	Scale	Data description/properties
Topography	1:250000	Elevation, overland, and channel slopes, lengths
Soil	1:4 000 000	Soil classifications and physical properties like bulk density, texture, saturated conductivity, etc.
Landuse	1:1 000 000	Landuse classifications such as cropland, pasture, forest, etc.
Weather	–	Daily precipitation, air temperature, relative humidity, solar radiation and wind speed, etc.
Land management information	–	Tillage, planting and harvesting dates for different crops

1.4 Evaluation of model prediction

Mean, relative error(Re), coefficient of determination (R^2), nash-suttcliffe simulation efficiency(E_{ns}) were used to evaluate model prediction (Nash, 1970). The R^2 is an indicator of strength of relationship between the observed and simulated values. E_{ns} indicates how well the plot of the observed value versus the simulated value fits the 1:1 line. If the R^2 and E_{ns} values are less than or very close to zero, the model prediction is considered "unacceptable or poor". If the values are equal to one, then the model prediction is "perfect".

1.5 Model calibration

It is important to understand that SWAT is not a "parametric model" with a formal optimization procedure (as part of the calibration process) to fit any data. Instead, a few important variables that are not well defined physically such as runoff curve number and universal soil loss equation's cover and management factor, or C factor may be adjusted to provide a better fit. A two-stage "brute force" optimization procedure described by Allred and Haan was used to find the optimum parameter values (Allred, 1996). This "brute force" optimization procedure, although being computationally less efficient than other methods, has the advantage of not being sensitive to local minimums in the objective function.

The procedure for calibrating the SWAT model for flow, sediment is shown in Fig. 2 (Santhi, 2001). Initially, base flow was separated from surface flow for both observed and simulated stream flows using an automated digital filter technique (Arnold, 1999). Calibration parameters for various model outputs were constrained within the ranges shown in Table 2. Model outputs were calibrated to fall within a percentage of average measured values and then monthly regression statistics (R^2 and E_{ns}) were evaluated. If all parameters were pushed to the limit of their ranges for a model output (i.e., flow or sediment) and the calibration criteria were still not met, then calibration would be stopped for that output.

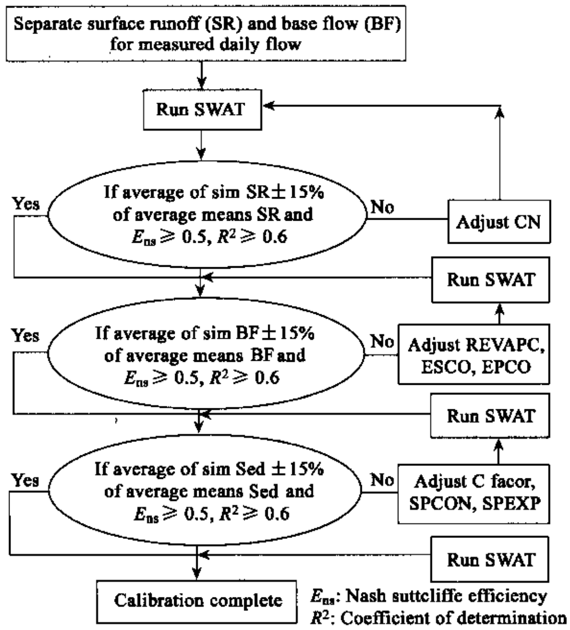

Fig. 2 Calibration procedure for flow and sediment in SWAT model

1.5.1 Flow

Stream flows from the Water Conservancy Committee of the Yellow River's Monitoring Station (Lushi Hydrological Station), and SWAT simulation were calibrated (Fig. 2) for the period from 1992 to 1997. For surface runoff calibration, the runoff curve number (CN_2) were adjusted within ± 8 from the tabulated curve numbers to reflect conservation tillage practices and soil residue cover conditions of the watershed (Table 2). The initial area-weighted CN_2 value is 69.2. For base flow calibration, related model parameters such as re-evaporation coefficient (*REVAPC*) for ground water that represents the water that moves from the shallow aquifer back to the soil profile/root zone and plant uptake from deep roots, soil evaporation compensation factor (*ESCO*), and plant evaporation compensation factor (*EPCO*) were adjusted from SWAT initial estimates to match the simulated and observed base flows (Table 2). At last, in order to match the stream flow, minimum melt factor for snow (*SMFMN*) was adjusted for snow melt months. The simulation was started from 1991 to reduce errors in initial estimates of state variables such as soil water content and surface residue.

1.5.2 Sediment

The cover, or C factor, of the Universal Soil Loss Equation was adjusted to match observed and simulated sediment loadings (Fig. 2). The C factor was adjusted (Table 2) to represent the surface better. Channel sediment routing variables such as the linear factor (*SPCON*) and the exponential factor (*SPEXP*) for calculating the maximum amount of sediment reentrained during channel sediment routing were also adjusted (Table 2) in the process of sediment calibration. These two variables were adjusted to represent the cohesive nature of the channels in this watershed.

Table 2 Inputs used in model calibration

Variable name	Model processes	Description	Model range	Actual value/change used
CN_2	Flow	Curve number	± 8	+ 2
REVAPC	Flow	Ground water revap coefficient	0.00 to 1.00	0.10
ESCO	Flow	Soil evaporation compensation factor	0.00 to 1.00	0.4
EPCO	Flow	Plant uptake compensation factor	0.00 to 1.00	0.2
SMFMN	Flow	Melt factor for snow on December 21	0 to 10	5.5
C factor	Sediment	Cover or management factor	0.003 to 0.45	Pasture: 0.009 Forest: 0.08 Cropland: 0.20
SPCON	Sediment	Linear factor for channel sediment routing	0.0001 to 0.01	0.0006
SPEXP	Sediment	Exponential factor for channel sediment routing	1.0 to 1.5	1.2

1.6 Model validation

In the validation process, the model was operated with input parameters set in the process of calibration without any change and the results are compared with the remaining observational data (from January 1998 to December 1999) to evaluate the model prediction. The same statistical measures were used to assess the model prediction.

2 Results and discussion

2.1 Calibration

2.1.1 Flow

Monthly measured and simulated flows at Lushi Hydrological Station match well. The estimated proportion of base flow from the observed flows was 30 percent from the filter technique which was 26 percent for the same location for SWAT simulated flows. Means of the observed and simulated stream flow was within a difference of 15 percent (Table 3). Further agreement between observed and simulated flows is shown by the coefficient of determinations and E_{ns} higher than 0.8 (Table 3). These results for surface runoff and base flow for observed and simulated flows revealed that hydrologic processes in SWAT are modeled realistically.

2.1.2 Sediment

Means of observed and simulated sediment is within a difference of 20 percent (Table 3). The values of R^2 and E_{ns} are both 0.70 (Table 3) which indicate that the simulated sediment is closer to the observed sediment and this model is able to predict sediment loadings well.

Table 3 Calibration results at Lushi Hydrological Station for period 1992 to 1997

Variable (units)	Annual mean		*Re*, %	R^2	E_{ns}
	Observed	Simulated			
Flow volume, m^3/s	13.15	15.04	14.4	0.87	0.87
Sediment, 10^4 ton	96.6	106.9	7	0.72	0.72

2.2 Validation

2.2.1 Flow

Observed and simulated flows at Lushi Hydrological Station matched well. The proportion of base flow is 28 percent for measured flow and 26 percent for simulated. *Re* is 14.6 percent, R^2 and E_{ns} are all greater than 0.80. The model overpredicted the flow in some months such as September 1998, from April to August 1999, and slightly underpredicted in May, August, and December 1998. The difference might stem from the spatial variability of precipitation. However, the prediction statistics were acceptable (Table 4).

Table 4 Validation Results at Lushi Hydrological Station for period 1998 to 1999

Variable (units)	Annual mean		*Re*, %	R^2	E_{ns}
	Observed	Simulated			
Flow volume, m^3/s	4.87	5.58	14.6	0.84	0.81
Sediment, 10^4 ton	189.24	158.9	−16.9	0.98	0.94

2.2.2 Sediment

Observed and simulated sediment loading matched well except that August 1998 when the sediment was underpredicted and May 1998 and September 1999 when the sediment was overpredicted. The values of R^2 and E_{ns} are both above 0.9 which indicated that the model is able to predict sediment reasonable. The reason for high values of R^2 and E_{ns} maybe that in 1998 the sediment yield was much greater than sediment yield in 1999. As the "goodness of fit" of observed and simulated data in 1998 was good (R^2 and E_{ns} are 0.989 and 0.944, respectively), the results were acceptable even though in 1999 the data did not match well (R^2 and E_{ns} are 0.50 and 0.51, respectively). Another reason was that the observed data were not accurate due to the difficulties to measure sediment. The results to some extent indicated that SWAT model is more suitable for high flow year (the precipitation in 1998 was at 10% frequency) than low flow year (the precipitation in 1999 was at 75% frequency).

3 Conclusions

It was the first time that a physically-based distributed non-point source model was validated in the Yellow River Basin. Using GIS technology, the basic database for non-point source pollution modeling was established for the study area. And the results showed that in most instances, simulated flow and sediment were close to the measured values during the calibration period and validation. In general, SWAT predictions were acceptable for watershed with high sediment runoff modulus. Meanwhile it should be noted that SWAT model is more suitable for the high flow year than the low flow year. With its spatial analytical capability, SWAT model based on GIS could be used to simulate non-point source pollutant loadings in such watershed with high sediment loadings, and be used as a useful tool for water resources planning and management in the Yellow River basin.

Acknowledgements: The authors are very grateful to Professor Datong Ning and Jialin Xu at the Institute of Environmental Sciences, Beijing Normal University, for providing detailed advice that contributed to improvements and refinements in the paper.

References:

Allred B, Haan C T, 1996. Small watershed monthly hydrologic modeling system [M]. Users manual. Biosystems and Agricultural Engineering Department, Oklahoma State University, Stillwater, OK.

Arnold J G, R Srinivasan, R S Muttiah *et al*., 1998. Large area hydrologic modelling and assessment. Part I: Model development [J]. Journal of American Water Resources Association, 34(1):73—89.

Arnold J G, P M Allen, 1999. Automated methods for estimating baseflow and grounderwater recharge from stream flow records [J]. Journal of American Water Resources Association, 35(2):411—424.

Bagnold R A, 1997. Bedload transport in natural rivers [J]. Water Resources Research, 13(2):303—312.

Nash J E, J V Suttcliffe, 1970. River flow forecasting through conceptual models. Part I: A discussion of principles [J]. Journal of Hydrology, 10(3):282—290.

Santhi C, J G Arnold, J R Williams *et al*., 2001. Validation of the SWAT model on large river basin with point and nonpoint sources [J]. Journal of the American Water Resources Association, 37(5):1169—1188.

USDA-SCS, 1972. National engineering handbook: hydrology. Section 4. Chap. 4—10 [M]. Washington, DC, USA : US Dept. of Agriculture, Soil Conservation Service.

Williams J R, 1975. Sediment routing for agricultural watersheds [J]. Water Resources Bulletin, 11(5):965—974.

Williams J R, 1969. Flood routing with variable travel time or variable storage coefficients [J]. Trans ASAE, 12(1):100—103.

Xi J Z, 1996. Water resources in Yellow River Basin [M]. Zhengzhou, China: Yellow River Publ House of the Water Resources.

Xia X H, Zhou J S, Yang Z F *et al*., 2001. An analysis of nitrogen pollution in river water of Yellow River Basin [J]. Journal of Environmental Sciences, 13 (5): 563—568.

Zhu X Y, Zhang X C, 1999. Change of water resources in Yellow River [M]. Zhengzhou, China: Yellow River Publ House of the Water Resources.

(Received for review July 29, 2003. Accepted September 1, 2003)