An Integrated water system model considering hydrological and biogeochemical processes at basin scale with application to Shaying River Catchment

Y.Y. Zhang[1,*], Q.X. Shao[2,*], A.Z. Ye[3] and H.T. Xing[1,4]

[1]Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China.

[2]CSIRO Digital Productivity and Services Flagship, Leeuwin Centre, 65 Brockway Road, Floreat Park, WA 6014, Australia.

[3]College of Global Change and Earth System Science, Beijing Normal University, 100875, Beijing, China.

[4] CSIRO Agriculture Flagship, GPO BOX 1666, Canberra, ACT 2601, Australia

Correspondence to: Y.Y. Zhang (zhangyy003@gmail.com) and

Q.X. Shao (quanxi.shao@cisro.au)

Abstract

Integrated water system modeling is a reasonable approach to provide scientific understanding and possible solutions to tackle the severe water crisis faced all over the world and to promote the implementation of integrated river basin management. Such a modeling practice becomes more feasible nowadays due to better computing facilities and available data sources. In this study, an integrated water system model (HE^{X}M) was developed by coupling multiple water related processes including hydrology, biogeochemistry, environment and ecology, as well as the interference of human activities. The model was tested in the Shaying River Catchment, the largest, highly regulated and heavily polluted tributary of Huai River Basin in China. The results showed that: HE^{X}M was well integrated with good simulation performance on the key water related components in the complex catchments (e.g. runoff for hydrology, water quality and nonpoint source pollutant load for water environment, crop yield for ecology). The simulated daily runoff series at most regulated and less-regulated stations matches observations well. The average values of correlation

coefficient and coefficient of efficiency were 0.81 and 0.63, respectively. The model performance of low and high flow events was improved when the dam regulation was considered although the low flow simulation was still not satisfactory. The observed daily ammonia-nitrogen (NH_4-N) concentration, as a key index to assess water quality in China, was well captured with the average correlation coefficient of 0.66. Furthermore, the nonpoint source NH_4-N load and grain yield were simulated for each administrative region and the results were reasonable in comparison with the data from the official report and the statistical yearbooks, respectively. The simulation performances were also compared with the SWAT's results from a previous study to test its practicability, and our model showed better simulation performances in both calibration and validation, particularly for daily runoff and water quality. This study is expected to give an improved water system modeling in complex basins, and provide a scientific support for the implementation of integrated river basin management.

1. Introduction

Severe water crisis hinders sustainable development in many regions all over the world nowadays, including flooding (Milly *et al.*, 2002. Schiermeier *et al.*, 2011), water shortages (Pimentel *et al.*, 2004; Wilhite *et al.*, 2005), water pollution (Jordan *et al.*, 2014; Zhou *et al.*, 2014) and ecological degradation (Revenga *et al.*, 2000; Vörösmarty *et al.*, 2010) . It is impossible to address these water problems using only the traditional single disciplinary approach (viz., hydrology, environmental sciences or ecology) because of the interconnections between water and the other related environment in the complicated water system (Kindler, 2000). The integrated river basin management might be one of the most sensible frameworks to comprehensively tackle these problems at basin scale. Thereinto, the integrated water system model is a reasonable practice to simultaneously simulate water related elements (flow regimes, nutrient loss, sediment and water pollution) (Kirchner, 2006), and also an effective tool to support water resource allocation, environment flow management, and river ecological restoration (Arthington, 2012).

The integrated water system model has been proposed and promoted due to the rapid development of water related sciences, computer sciences and earth observation technologies in the last decade. The hydrological cycle has been widely accepted as a

critical linkage among physical (e.g. runoff), biogeochemical (e.g. nutrient, water quality) and ecological processes (e.g. plant growth), energy fluxes at basin scale (Wigmosta *et al.*, 1994; Singh and Woolhiser, 2002; Burt and Pinay 2005). For example, the physiological and ecological processes of vegetation affect evapotranspiration, soil moisture distribution and infiltration, and nutrients absorption and movement. On the contrary, soil moisture and nutrient content directly affects crop growth. Overland flow is a carrier of the pollutant load discharge to water body (Li *et al.*, 1992; Wigmosta *et al.*, 1994; Burt *et al.*, 2005; Lohse *et al.*, 2009). Therefore, it is more reasonable to analyze variation patterns of water related elements and their causes at the basin scale by coupling all these processes and capturing the interactions and feedbacks between individual cycles. Furthermore, multidisciplinary research provides an effective way to make the possible breakthrough in water system modeling. Mature basic theories of water-related disciplines have been formalized for over a century (Singh and Woolhiser, 2002) and more observation data can be obtained including high resolution of spatial information data (DEM, land use and crop distribution), chemical and isotopic data from field experiment (Kirchner, 2006).

Several models have been developed based on the mature models of different disciplines since the 1980s (Singh and Woolhiser, 2002). Most of existing models focus on one or two major processes in simulation (e.g. hydrology, water quality, biogeochemistry). Table 1 gives the components considered in several famous models. Due to the complexity of the integrated system, it is rare that each process is simulated with the same detail. Thus one gets best performance for processes that are described with the most detail and only approximate results for others outside of the model's focus (Mantovan and Todini, 2006; Mantovan *et al.* 2007). The hydrological models emphasize on the rainfall-runoff relationship and link with several dominating water environmental and biogeochemical processes. As a result, the hydrological models usually have satisfactory performance in hydrological process. Examples of widely accepted models include HSPF (Bicknell *et al.*, 1993), ANSWERS (Bouraoui and Dillaha, 1996) and AnnAGNPS (Bingner and Theurer, 2001). The water quality models depict the detailed migration and transformation processes of pollutants in receiving water bodies using complicated numerical solutions of multi-dimensional dynamics equations. Thus the models are subject to computational instability and time consuming. The typical models are WASP (Di Toro *et al.*, 1983), QUAL2K (Brown

and Barnwell 1987), EFDC (Hamrick, 1992). The biogeochemistry models have advantages to simulate physiological and ecological processes of vegetation, the vertical movement of nutrients and water in soil layers at the field scale or experimental catchment scale, but lack the accurate hydrological features (Deng *et al.*, 2011). Thus it is hard to simulate the longitudinal movement of water, nutrients and their loss along flow path in the basin. The examples are EPIC (Sharpley and Williams, 1990), DNDC (Li *et al.*, 1992).

So far, SWAT is a typical integrated water system model, which simulates most of water related processes over long time periods at large scales (Arnold *et al.*, 1998). Its model structure and powerful functions are considered as a landmark in the field of water system modeling. However, not all of water related processes could be well captured in practice, such as daily flow and extreme events (Borah and Bera, 2004), soil nitrogen and carbon (Gassman *et al.*, 2007), the performance in regulated basins (Zhang *et al.*, 2012). The probable reasons were the applicability and inaccurate descriptions of some modules (Neitsch *et al.*, 2011). In particular, two methods are adopted in SWAT to estimate surface runoff, viz., Soil Conservation Service (SCS) curve number method and Green-Ampt infiltration model. The SCS equation is usually given priority, but it is developed for rural watersheds in the United States and the applicability of curve number to other regions is questioned (Rallison and Miller 1981). The Green-Ampt infiltration model is usually limited to simulate flow events at micro scales (time: hours or minutes, space: fields or 10^{-1} to 10 km^2 watersheds) (Brakensiek, 1977; King *et al.*, 1999). Furthermore, it is much more difficult for SWAT to capture the complicated dynamic processes of soil nitrogen and carbon accurately compared with other biochemistry models, such as DNDC (Li *et al.*, 1992; Gassman *et al.*, 2007).

In practice, the integrated river basin management is replacing the traditional water resource management gradually all over the world by comprehensively considering runoff, water quality and ecological responses, as well as human water requirements (Gleick 1998). However, several challenges are now becoming obvious for the practitioners, such as the complicated interaction mechanism of water, geochemistry, ecology (Kirchner 2003), the multiple scale problem (McDonnell *et al.*, 2007), and the trade-off in allocating water resources among the living, production and ecology (Letcher *et al.*, 2007). The models mentioned above, which only concentrate on one or

two processes (e.g. flooding control, drought relief and pollution improvement), are difficult to account for these challenges. The integrated water system models should be proposed by considering multiple processes and interactions in more detail to provide the scientific and technical support for integrated river basin management.

The objective of this study is to develop an integrated water system model considering hydrological and biogeochemical processes, with the aim to simulate the spatial and temporal variations of several key elements (e.g. evapotranspiration, soil water, flow regimes, nonpoint source pools of nitrogen (N), phosphorus (P) and carbon (C), water quality variables in water body, crop yield and greenhouse gas emissions) in complex basins. The model framework is put forward based on the interchange and balancing processes of water and nutrient which are depicted by several robust models. The parameter analysis module is also included in our programming codes. The model performance is illustrated by a case study in China and compared with the results from the existing SWAT to test its practicability .

The paper is organized as follows. The model framework and individual modules are introduced in Section 2, followed by the case study, including data pre-processing, the calibration criteria, and model performance. Conclusions and perspective discussions are drawn in Section 4.

2. Model Framework

In this study, an integrated water system model is developed to capture the more accurate processes for hydrological cycle and detailed processes for soil biogeochemistry in complex basin. The proposed model is named HE^{X}M where **H** indicates **H**ydrological modules and **X** is used to indicate **E**cological, water **E**nvironmental modules with possible future extension to **E**conomic modules (water consumption processes in the social economy system). There are seven major modules in HE^{X}M, named as hydrological cycle module (HCM), soil biochemical module (SBM), crop growth module (CGM), soil erosion module (SEM), matter migration module (MMM) and water quality module (WQM), as well as dams regulation module (DRM). Parameter analysis tool (PAT) is a useful tool for HE^{X}M calibration and is independent from other modules. The exterior exchange elements connecting different modules are given in Figure 1. The interior elements of each

module have not been listed in order to better present the model structure and the relationship of these modules.

The motivation driving the development of our proposed model is to improve the model simulation results through the following aspects. The proposed model is constructed based on the hypothesis that the cycles of water and nutrients (N, P and C) are inseparable and act as the critical linkages among all the modules. The model takes full advantages of powerful interconnection and simulation functions of hydrological model at large spatial scale, elaboration of matter vertical movement in soil layers of ecological model at field scale and matter longitudinal movement in river networks of environmental model. Firstly, several key hydrological elements are critical linkages among all the modules. Specifically, evaportranspiration changes the vertical distribution of matter in the soil layers (SBM); soil water content is a key constraint for soil biogeochemical processes (SBM) and plant growth (CGM); flow is the critical driving force to the transport of sediment and pollutants in both lands (MMM) and streams (WQM) (Section 2.1). Secondly, soil biochemical processes determine the nutrient amount absorbed in the crop growth process (CGM) and migrated into water bodies as the nonpoint pollutant source (MMM and WQM). The complicated nutrient processes in the soil profiles, including decomposition, mineralization, immobilization, nitrification, denitrification, plant uptake and leaching, are described in more detail in order to improve the simulation of water environmental processes in responding to agricultural management (Section 2.2). Thirdly, the hydrological module also provides the spatial relation function among different calculation cells, which is helpful to simulate the longitudinal movement of water and nutrient in the land surface and river networks (Section 2.1 and 2.3).

Although some modules of the proposed model stem from the same existing models with SWAT (e.g. CGM from EPIC, WQM from QUAL-2E), it has several key differences from SWAT. Firstly, the proposed model adopts Time Variant Gain Model (TVGM) to calculate surface runoff yield because of its strong theoretical basis and easy applicability (Xia *et al.*, 2005). Secondly, the soil biogeochemical processes are depicted in more detail by the DNDC model (Li *et al.*, 1992). Thirdly, the linkages among modules were different from SWAT because water and nutrient cycles were considered as two critical linkages, rather than water cycle in SWAT. Seven forms of organic nitrogen and three forms of inorganic nitrogen were simulated. However, only

three forms of organic nitrogen and two forms of inorganic nitrogen were considered in SWAT. Fourthly, the dams regulation module is extended to further approximate the actual flow regulation rules of dams and sluices (Zhang *et al.*, 2013). Fifthly, the three levels of spatial calculation cell are designed in HE^{X}M according to the different scopes of these modules, i.e. subbasin cell, landuse cell and crop cell from largest to smallest. All the levels of spatial cell are visible, more reasonable and flexible than the virtual hydrologic response unit of SWAT (Neitsch *et al.*, 2011).

More detailed description of each module and its interactions with other modules are given in the following individual sub-sections. The main equations of each process are deferred to the appendix and supplementary material section for readers who are interested in the mathematical details.

2.1 Hydrological cycle module (HCM)

Hydrological process controls physiological and ecological processes of vegetation, oxidation-reduction reaction and anaerobic reaction of matters in soil layers, spatial and temporal distribution of water and pollutant in the basin. A flowchart is given in Figure 2, from which it can be seen that shallow soil water from HCM is one of the major factors connecting CGM (to control crop growth) and SBM (to control vertical migration and reaction of matters in soil profiles). Plant transpiration is also linked to SBM (to provide energy for vertical migration of matters in soil profiles). The overland flow including surface runoff and soil runoff is linked to SEM and MMM (to drive longitude migration of matter and sediment along flow paths), to WQM for runoff routing in water bodies (rivers and lakes). Moreover, HCM calculates inflow of dams or sluices for DRM.

Surface runoff yield calculation is the core of hydrological simulation and has close relationships with many other processes. Time Variant Gain Model (TVGM) (Xia *et al.*, 2005) is applied to calculate surface runoff yield because of its strong theoretical basis. In TVGM, the rainfall-runoff relationship is nonlinear with surface runoff coefficient varying and being affected significantly by antecedent soil moisture (Xia *et al.*, 1991). TVGM is based on the Volterra function and has satisfactory performance in many different scale basins, especially in arid and semiarid regions (Xia *et al.*, 2005; Wang *et al.*, 2009). In HE^{X}M, surface runoff yields of different landuse fields are calculated separately including forest, grassland, water, urban,

unused land, paddy land and dry land.

The potential evapotranspiration is calculated using Hargreaves method (Hargreaves and Samani, 1982) because it only uses the daily maximum and minimum temperature data which are widely available. The actual plant transpiration is expressed as a function of potential evapotranspiration and leaf area index while the soil evaporation is expressed as a function of potential evapotranspiration and surface soil residues. The interflow and baseflow are considered as linear storage-outflow relationships (Wang *et al.*, 2009). The infiltration from the upper to lower soil layer is calculated using storage routing methodology (Neitsch *et al.*, 2011). The Muskingum method or kinetic wave equation is used for river flow routing.

2.2 Ecological process modules

Ecosystem is one of the decisive components to the hydrological cycle and the material migration and transportation. The main feature of HE^{X}M is that water cycle, nutrient cycles and crops growth, as well as their key linkages are incorporated. The ecological processes contain SBM and CGM modules.

2.2.1 Soil biochemical module (SBM)

SBM simulates the key processes of C, N and P dynamics in the soil profiles, including decomposition, mineralization, immobilization, nitrification, denitrification, plant uptake and leaching. C constrains the decomposition and denitrification of N and P. Soluble nutrient including nitrate nitrogen (NO_3-N), nitrite (NO_2-N), NH_4-N and Soluble P outputted from SBM are connected to CGM as nutrient constraint of crop growth, and to MMM as main sources of pollutant to water bodies as well as insoluble matters (Organic P and N) (Figure 3a). The daily step decomposition and denitrification submodels in DNDC (Li *et al.*, 1992) are adopted to simulate biogeochemical processes of C and N in the soil profile at field scale (Li *et al.*, 1992). The major processes of soil P cycle are simulated based on the studies of Horst *et al.* (2001). The soil profile is divided into three layers, viz. surface (0-10cm), and user defined upper and lower layer.

Soil C and N cycle: In the aerobic state, the decomposition and other oxidation processes, such as nitrification, mineralization and immobilization, are the dominant microbial processes. The denitrification process is activated by rainfall or irrigation

events. As the oxygen availability is limited, a series of N oxides is used to replace oxygen as the terminal electron acceptor during soil oxidation-reduction reaction.

- *Decomposition:* There are three conceptual organic C pools: the decomposable residue pool, microbial biomass pool and a stable pool (humus). Every pool contains resistant and labile components. Additionally, the residue pool contains very labile component. Decomposition of each C pool is treated as the first-order decay process with the individual decomposition being modified by soil temperature and moisture, clay content and the C: N ratio. Carbon dioxide (CO_2), released from soil organic carbon (SOC), is calculated as a constant fraction of the C undergoing decomposition of three C pools. When precipitation and/or irrigation happen, the decomposition process will pause and the denitrification process will start until the soil water file pore space (WFPS) in the surface soil layer reaches a threshold (e.g. 40%) or the substrates are used up. Then the decomposition will restart. The details of SOC pool structure are described in Li *et al.* (1992).

- *Nitrogen transformation during decomposition*. The major simulated processes with decomposition under aerobic condition are mineralization, immobilization, ammonia (NH_3) volatilization and nitrification. Ammonium (NH_4^+) is mineralized from organic N pool when SOC flows from lower C: N ratio C pools into higher C: N ratio C pools. During immobilization, if the mineral N (NH_4^+ and NO_3-N) is not enough, SOC decomposition will reduce to an allowable level. NH_3 volatilization is controlled by NH_4^+ concentration, clay content, pH, soil moisture and temperature. NH_4^+ is microbial oxidized to NO_3-N and nitrous oxide (N_2O) which emit into the air as a gaseous intermediate during nitrification. The proportion of N_2O is small and is controlled by NH_4^+ concentration, pH, temperature, moisture, etc in the soil layer.

- *Denitrification:* The denitrification process works during rainfall or irrigation events when WFPS is greater than the threshold. The general recognized reduction sequence in denitrification is $NO_3 \rightarrow NO_2 \rightarrow NO \rightarrow N_2O \rightarrow N_2$. The denitrification rate correlates with denitrifier biomass, moisture, pH, temperature, and NO_3-N concentration in the soil layer. The denitrifier biomass is estimated with the growth and dead rate of denitrifier which is controlled by soluble soil C and soil moisture, temperature. The C and N from dead cells are added to the pools of immobilized C and N which no longer participate in the dynamic processes. The consumption rate of soluble C depends on the biomass, relative growth rate, and the maintenance

coefficients of the denitrifier populations. The daily emissions of N_2O and N_2 are calculated as a proportion of total production of N_2O and N_2 which is related to the adsorption coefficients of gases in soils and the air filled porosity of the soil. But the emission is neglected because of the low diffusion rates in soil water during the rainfall events.

Soil P cycle: Four major forms of P in soils are considered, viz., stable organic P, active organic P for plant uptake, fresh organic P associated with plant residue, microbial biomass and soluble mineral P as the consequence of mineralization, decomposition and sorption (Horst *et al*, 2001). The P dynamics processes are considered in Horst *et al.* (2001) and Neitsch *et al.* (2011), through modeling the P release from fertilizer, manure, residue, microbial biomass and humic substances, and plant uptake and transport by soil erosion.

2.2.2 Crop Growth Module (CGM)

CGM is developed based on EPIC crop growth model (Hamrick, 1992), which applies the concept of daily accumulated heat units on phenological crop development, with Monteith's approach for potential biomass, harvest index for partitioning grain yield, stress adjustments for water, temperature, and N availability in the root zone of the soil profile. It simulates total dry matter, leaf area index, root depth and density distribution, harvest index, and N uptake, etc (Willians *et al.*, 1989; Sharpley and Williams, 1990). The crop respiration and photosynthesis drive the vertical movement of water and nutrient, and transpiration. In CGM, the output of leaf area index is the main factor connecting HCM (to control the transpiration), and the crop residues left in the fields is the main source of organic matters (N, P and C) connecting to SBM for soil biochemical degradation, to MMM for overland migration, and to SEM as one of the five constraint factors (Figure 3b).

2.3 Water environmental modules

The water environmental modules are to simulate the matter (e.g., different forms of nutrient, chemical oxygen demand) migration and transformation with the movement of surface water and sediment. The main modules are SEM for simulating sediment yield, MMM for matter migration to water bodies (rivers or lakes) with overland flow and sediment, WQM for the matter migration and transformation in water bodies.

2.3.1 Soil erosion module (SEM)

The soil erosion by precipitation is estimated using the improved ULSE equation (Onstad and Foster 1975) based on runoff outputted from HCM, crop management factor outputted from CGM. SEM simulates sediment load for MMM to provide the carrier for the migration of insoluble organic matters along overland transport paths and water bodies (Figure 4a).

2.3.2 Matter migration module (MMM)

The eroded matters are divided into insoluble and soluble forms and the main sources are the erosion of soil and urban area, the sewage discharge of rural living and livestock breeding. The soil erosion, as the primary source in most catchment, is estimated using DNDC (Li *et al.*, 1992) and the other sources are estimated using the export coefficient method (Johnes,1996). The overland migration processes contain the soluble matter migration with overland flow, the insoluble matter migration with sediment, and the loss during the migration. All of these processes take place along the overland transport paths. This module calculates the matter load discharged into rivers for WQM (Figure 4b).

2.3.3 Water quality module (WQM)

Two water quality modules are designed for different types of water bodies at the subbasin scale, viz., the in-stream water quality module and the water quality module of water impounding (reservoir or lake). The enhanced stream water quality model (QUAL-2E) (Brown and Barnwell 1987), as a comprehensive and versatile stream model, is adopted to simulate the longitudinal movement and transformation of water quality constituents in the branch stream systems. The model is centered at dissolve oxygen (DO) and can simulate up to 15 water quality constituents including temperature, DO, sediment, different forms of nutrient (N and P), chemical oxygen demand (COD) (Neitsch *et al.*, 2011) (Figure 4c). In order to avoid the computational instability and improve the calculating efficiency, the model is solved at the subbasin scale rather than the fine grid scale for numerical solution. The mass balance is used to determine the constituents' concentration based on the mass fluxes into and out of each computational unit and the degradation of the constituents themselves. The water quality outputs are linked to DRM to provide upper water quality boundary of dams or sluices. The water quality module of water impounding assumes that water body is

at the steady state and focuses on the vertical interaction of constituents. The main processes are the constituent's degradation, settlement, resuspension and decay in the sediment.

2.4 Dams regulation module (DRM)

The dams or sluices highly disturb flow regimes and associated environmental processes in most river networks (Zhang *et al.*, 2013). The regulation of dams or sluices should be considered in water system models. DRM provides hydrological boundaries (e.g. water storage, runoff) regulated by dams or sluices to HCM for flow routing and to WQM for matter migration (Figure 1).

In our system, three methods are proposed for calculating water storage and outflow of dams or sluices, i.e., measured outflow, controlled outflow with target water storage, and the relationship between outflow and water storage volume (Zhang *et al.*, 2013). The first methods require users to provide the measured outflow series during the simulation period. The second method simplifies the regulation rule of dam or sluice for the long-term analysis, that is, it assumes that water is stored according to the usable water level during the non-flooding season and the flood control level during the flooding season and that the redundant water is discharged. The method requires the characteristic parameters of dam or sluice including water storage capacities of dead, usable, flood control and maximum flood levels and the corresponding water surface areas. The third method is proposed based on relationship among water level, water surface area, storage volume and outflow according to the actual situation of China.

2.5 Parameter analysis tool (PAT)

Parameter sensitivity analysis and auto-calibration are critical steps for the applications of highly parameterized models, especially the integrated water system models (McDonnell *et al.*, 2007). Several parameter analysis methods are coupled in HE^{X}M, including parameter sensitivity method (Latin Hypercube One factor At a Time: LH-OAT) (van Griensven *et al.*, 2006), auto-optimization methods such as Particle Swarm Optimization (PSO) (Kennedy *et al.*,1995), Genetic Algorithm (GA) (Goldberg 1989) and Shuffled Complex Evolution (SCE-UA) (Duan *et al.*,1994) as well as uncertainty analysis method (Bayesian approach). Five indices are provided to

evaluate model performance including bias (*bias*), relative error (*re*), root mean square error (*RMSE*), correlation coefficient (*r*) and coefficient of efficiency (*NS*). The *RMSE* is used to calculate the likelihood function.

2.6 Multi-scale solution

Generally, the hydrological, water environmental, and ecological processes take place at different spatial and temporal scales due to the data spatial heterogeneities and reaction time (Blöschl and Sivapalan, 1995; Singh and Woolhiser, 2002; Lohse *et al.*, 2009) (Figure 5). For the spatial scale, the hydrological and water environmental processes are at catchment scale while the ecological processes are at site or field scale (Xia *et al.*, 1991; Li *et al.*, 1992). For the temporal scale, the main hydrological processes are at daily scale except the long-term baseflow; the main water environmental, soil biogeochemical and ecological processes are at daily or sub-daily scale including nitrification and denitrification, mineralization, adsorption and photosynthesis. The accumulation and loss of matter and biomass are at the daily or larger (months, years) scale (Kirchner, 2006; Lohse *et al.*, 2009).

In HE^{X}M, three levels of spatial calculation cell are designed, i.e. subbasin cell, landuse cell and crop cell from largest to smallest. The subbasin is divided on the basis of DEM, the position of gauges, dams and sluices, and used for HCM (e.g. flow routing in both land and instream), MMM, WQM and DRM. Seven specific landuse cells of each subbasin are partitioned related to the landuse classification including forest, grassland, water, urban, unused land, paddy land and dry land. The related modules are HCM (e.g., water yield, infiltration, interception and evapotranspiration) and SEM. Moreover, in these specific landuse cells, paddy land and dry land, forest, grassland are considered in the agricultural management. In current version of HE^{X}M, ten specific categories of crop cells are divided for these four landuse cells, i.e. fallow for all these landuse cells; grass for grassland cell; fruit tree and non-economic tree for forest cell; early rice and late rice for paddy cell; spring wheat, winter wheat, corn, and mixed dry crop for dry land cell. The crop cell category of a certain landuse is variable, which depends on crop cultivation structure and timing. The related modules are SBM and CGM. All the outputs of crop cell are summarized at landuse cell scale, or subbasin scale based on the area percentage of different cells, respectively.

For the multiple temporal scale, because most simulated processes are considered to

take place at daily scale (Kirchner, 2006; Lohse *et al.*, 2009), the time step of HE^{X}M is one day. Several linear or nonlinear aggregating functions are used to transform different time scales to the daily scale, such as exponential relation, accumulative relation (Vinogradov *et al.*,2011).

2.7 Basic datasets and spatial delineation

The indispensable spatial and temporal datasets of HE^{X}M are GIS data (DEM, soil physical and chemical properties, land use and crop types), daily meteorological data (precipitation, maximum and minimum air temperature), social-economic data (populations in urban and rural area, breeding stock of large animals and livestock, chemical fertilizer amount and cultivation methods, water withdrawal and point source pollutant load), dams characteristic data (water storage capacities of dead, usable, flood control and maximum flood levels and the corresponding water surface areas). Several monitoring data series are also needed for model calibration, such as runoff or water quality series at river sections, soil water content and crop yield at the field scale. All the datasets and their usages in HE^{X}M are given in Table 2.

The hydrological toolset of Arc GIS platform are used to delineate all the spatial simulation cells, river system and flow routing relationship between subbasins based on DEM, landuse data.

3. Model Testing

As an example, HE^{X}M is applied in a highly regulated and heavily polluted river basin of China in order to test the model performance. The simulated components contain daily runoff and water quality concentration at several river cross-sections, spatial patterns of nonpoint source pollutant load and crop yield at subbasin scale. Furthermore, the simulation results are compared with the existing studies calculated by another widely-used model (SWAT) in the same area (see Zhang *et al.*, 2013).

3.1 Study area

Shaying River Catchment (112°45′~113°15′E, 34°20′~34°34′N), as the largest subbasin of Huai River Basin in China, is selected as our study area (Figure 6). It has the drainage area of 36,651 km^2 and the mainstream of 620 km long. The average

annual population (2003-2008) is 32.42 million including 23.70 million rural population. The average annual stocks of big animals and livestock are 8.30 million and 178.42 million, respectively. The average annual amount of chemical fertilizer is 1.55 million ton. The basin is located in the typical warm temperate, semi-humid continental climate zone. The annual average temperature and rainfall are 14-16^{o}C and 769.5 mm, respectively. Meanwhile, Shaying River is the most serious polluted tributary with pollutant load contributing over 40% of the whole Huai River and is usually known as the water environment barometer of Huai River mainstream. In order to reduce flood or drought disasters, 24 reservoirs and 13 sluices have been constructed and fragment river into several impounding pools which control over 50% of the total annual runoff.

3.2 Model setup and evaluation

Shaying River Catchment was divided into 46 subbasins and the main land use types were dry land (84.04%), forest (7.66%), urban (3.27%), grassland (2.68%), water (1.43%), paddy (0.91%) and unused land (0.01%) according to the landuse classification standard of China (CNS,2007). The soil input parameters (the contents of sand, clay and organic matter) were calculated based on the percentage of soil types in each subbasin.

The daily data series at 65 precipitation stations and six temperature stations were interpolated to each subbasin from 2003 to 2008, using the inverse distance weighting method and the nearest-neighbor interpolation method, respectively. The social economic data (i.e., populations in urban and rural area, breeding stock of large animals and livestock, chemical fertilizer amount) were also interpolated into each subbasin based on the area percentage (Figure 7). The main crops of Shaying River Catchment were the early rice, late rice, winter wheat and corn, and their agricultural management schemes were summarized by field investigation and referring to Zhai *et al.* (2014) and Wang *et al.*, (2008) (Table 3). Moreover, 23 major dams and sluices and over 200 pollutant outlets were considered in the model according to the geographical positions.

To reduce the dimensions of the calibration problem, we restricted SCE-UA to calibrate only the sensitive parameters as defined by LH-OAT. The model calibration was conducted step-by-step as follows. Hydrological parameters were calibrated

firstly according to observed runoff series at each station from upstream to downstream, and then water quality parameters according to observed NH_4-N concentration series. The calibration and validation periods were from 2003 to 2005 and from 2006 to 2008, respectively. The *bias*, *r* and *NS* were used to evaluate model performance in the case study as a demonstration of our proposed model. However, *NS* was sensitive to extreme value, outlier and number of data points and was not commonly applied in the environmental sciences (Ritter and Muñoz-Carpena, 2013). Thus it was not used to evaluate the NH_4-N concentration simulation. Weighting method was usually used to comprehensively handle different objectives (Efstratiadis and Koutsoyiannis, 2010). In this study, these objective functions were simply aggregated to a single objective (f_{runoff} and $f_{NH4\text{-}N}$) because the case study was only a demonstration of our model performance.

$$\begin{cases} f_{runoff} = \min[(|bias| + 2 - r - NS)/3] \\ f_{NH_4-N} = \min[(|bias| + 1 - r)/2] \end{cases} \quad (1)$$

Moreover, because of the high regulation in most rivers, it is necessary to consider the impact of dam's regulation in the integrated water system models. The regulation of dams and sluices usually disturbs the intra-annual distribution of flow events, e.g. flattening high flow and increasing low flow. The simulation performances of high and low flow were evaluated separately, and the effectiveness of the DRM was tested by comparing the simulation with and without considering regulation. The high and low flow was determined by flow duration curves and the threshold was 50% for easy presentation. That was, the flow was high flow if its percentile was greater than the threshold, whereas it was low flow.

3.3 Results

3.3.1 Parameter sensitivity analysis

Over 200 parameters (93 lumped and 112 distributed) control the hydrological, ecological and environmental processes of HE^{X}M according to the degree of spatial heterogeneity. The 112 distributed parameters are divided into 48 overland parameters, 18 stream parameters and 46 water project parameters (only for the subbasin owning water project) according to their spatial distribution. These parameter values are determined by the properties of overland landscape and soil, stream patterns and water

projects, respectively. Different subbasins share many common parameter values because of the same properties above. LH-OAT was used to test the sensitive parameters in HE^{X}M.

Nine most sensitive parameters were detected for runoff simulation (Table 4) including soil related parameters *WMc* (field capacity), *WM* (saturated moisture capacity), k_r (interflow yield coefficient) and f_c (steady state infiltration rate); TVGM parameters g_1 (basic surface runoff coefficient) and g_2 (influence coefficient of soil moisture) for surface water calculation; ground water recharge parameters k_g (baseflow yield coefficient) and T_g (delay time for aquifer recharge); and adjusted factor KET_p of evapotranspiration. All of these parameters controlled the main hydrological processes, in which soil water and evapotranspiration processes were distinctly important, explaining 54.3% and 23.2% of the runoff variation, respectively.

For NH_4-N concentration simulation, over 90% of observed NH_4-N concentration variation were explained by 14 sensitive parameters which were categorized into hydrological (59.28% of variation), NH_4-N (20.65% of variation) and COD (12.34% of variation) related parameters (Table 3).

3.3.2 Hydrological simulation

The daily runoff hydrographs were reproduced at seven regulated stations and one less-regulated station (viz., the upstream stations unaffected by dams, or downstream stations situated far from dams). The simulations were well fitted with the observations at all the stations from the midstream to downstream (Figure 8 and Table 5). The *bias* at all the regulated stations were very close to 0.0 except the underestimation at Zhoukou (0.24 for calibration and 0.41 for validation), and the overestimation at Mawan (-0.44) and Luohe (-0.52) during the validation period. The obvious biases were caused by the tradeoff among these three evaluation indices. However, the *r* values ranged from 0.61 (Huangqiao for validation) to 0.92 (Yingshang for calibration) with the average value of 0.81 while the *NS* values ranged from 0.23 (Huangqiao for validation) to 0.84 (Yingshang for calibration) with the average value of 0.63. The results at the regulated stations were little worse than that at the less-regulated stations due to the regulations.

Improved SWAT 2000 was also well calibrated to simulate the long-term runoff in the same area (Zhang *et al.*, 2013). Only monthly runoff was simulated because it was

difficult to forecast accurately the regulated daily runoff processes. The f_{runoff} values by SWAT at monthly scale ranged from 0.08 (Luohe for calibration) to 0.61 (Shenqiu for validation) (Table 5). However, our model captured the daily runoff processes very well at most stations, particularly for high flow regimes. Compared with the results of SWAT at the monthly scale, the f_{runoff} was significantly improved for both calibration and validation at most of the downstream stations including Huangqiao, Huaidian, Fuyang, Yingshang and Shenqiu although it became little worse at the upstream stations, viz., from 0.08 to 0.14 at Luohe and from 0.12 to 0.19 at Zhoukou for calibration, from 0.19 to 0.24 for calibration and from 0.30 to 0.38 for validation at Mawan.

Compared the simulation with the observations from 2003 to 2008, the high and low flows were usually overestimated at all the stations without considering the regulations. Except the high flow simulation at Huangqiao and Zhoukou, all the high and low flow events at all the stations were better simulated with the consideration of dams and sluices regulation (Table 6). The best fitting was at Fuyang, especially for the high flow simulation (*bias*=0.10, *r*=0.89 and *NS*=0.78). The improvements of f_{runoff} from unregulation to regulation settings ranged from -0.09 (Fuyang) to -0.20 (Huaidian) for high flow simulation except Huangqiao and Zhoukou (0.01), from -0.06 (Zhoukou) and -0.34 (Mawan) for average flow simulation, and from -3.16 (Huangqiao) and -15.28 (Luohe) for low flow simulation. However, the entire low flows were overestimated at most stations (*bias* < 0.0) except Zhoukou station (0.48), indicating that the simulation performance still needs to be further improved. The probable reasons were that, on the one hand, the low flow forecasting is actually a difficult task especially in the regulated river basins, and the common evaluation criteria are disadvantageous to evaluate low flow simulation (Pushpalatha *et al.*, 2012) and, one the other hand, the dam regulation module of HE^{X}M is still difficult to fully capture the low flow events.

3.3.3 Water quality simulation

NH_4-N concentration is one of the widely used indexes to assess water environmental quality in China (CSEPA, 2002). The observation series at five regulated stations and two less-regulated stations in the middle and downstream were used to calibrate the related parameters of environmental processes. The spatial distribution of nonpoint

source pollutant load was also estimated.

The simulated concentrations showed good agreements with the observations according to the evaluation standard recommend by (Moriasi *et al.*,2007) (Figure 9 and Table 7). The *r* values of all the stations were over 0.60 expect Zhoukou (0.56 for validation), Yingshang (0.49 for validation) and Shenqiu (0.41 for validation) with the average value of 0.66. The *bias* of all the stations were considered as "acceptable" with the range from -0.27 (Fuyang for validation) to 0.29 (Zhoukou for calibration) except Fantaizi (0.45 for calibration and -0.62 for validation). The best simulation was at Luohe. As the accuracy of water quality simulation is directly influenced by hydrological simulation (Gassmann *et al.*, 2014), the unacceptable *bias* at Fantaizi might be attributed to the possible large error in the simulated hydrological processes which was not calibrated due to the lack of observed runoff data. Moreover, the obvious discrepancies between the simulation and observation often appeared in the period from January to May due to the poor simulation performance of low flows.

Compared with the results of Zhang *et al.* (2013), all the values of $f_{NH4\text{-}N}$ decreased obviously with the range from -0.02 (Fantaizi for validation) to -0.61 (Fuyang for validation) except Zhoukou (0.16 for calibration). Although the *bias* of HEXM simulation for some stations was greater than that of SWAT, the overall simulation performance was improved greatly by coupling N cycle model (DNDC) because the values of $f_{NH4\text{-}N}$ were less than that of SWAT. However, the simulation at Zhoukou during the calibration period was an exception due to the runoff simulation error. The simulation was also significantly improved considering the regulation compared to the results without the regulation except Fuyang for calibration. The decreases of $f_{NH4\text{-}N}$ value ranged from 0.06 (Fantaizi for calibration) to 0.49 (Zhoukou for validation). The pollutant concentration usually reduced obviously at the upstream of dams or sluices, due to the degradation and settlement of large water storage. Thus, the simulated concentrations without regulation were greater than the observation or the simulation with regulation, except Huaidian, and the largest difference appeared at Zhoukou.

The spatial pattern of average annual nonpoint source NH_4-N load was shown in Figure 10. The modeled annual yield rates ranged from 0.048 t km^{-2} $year^{-1}$ to 11.00 t km^{-2} $year^{-1}$ with a mean of 0.73 t km^{-2} $year^{-1}$. The yield of each administrative region was summarized properly from the subbasin scale according to the area percentage of

subbasins in each administrative region. Compared with the statistical results of each administrative region given in the official report (HRC., 2011), the *bias* of simulated nonpoint source load in the whole region was 21.31% excluding the great discrepant regions (i.e., Fuyang and Pingdingshan). The high load yield regions were in the middle of Pingdingshan, Xuchang, Zhengzhou, Fuyang and Zhoukou regions. The spatial pattern was significantly correlated with the distribution of paddy fields (r=0.506, p<0.001) and rice yield (r=0.799, p<0.001). The fertilizer loss of paddy fields might be the primary contributor to the nonpoint source NH_4-N load, possibly because the average nitrogen losses coefficient in China was just 30~70% in paddy fields, which was much greater than that in the dry field (20~50%) (Zhu, 2000; Xing and Zhu, 2000). The nitrogen was prone to loss by volatilization to air, dissolution and drainage into rivers with runoff in paddy fields.

Given that the observed average annual point source NH_4-N loads into rivers were about 4.70×10^4 t $year^{-1}$ in Shaying River Catchment, the nonpoint source load contributed 38.57% of the overall NH_4-N load on average from 2003 to 2005, which was little greater than the statistical results (29.37%) given in the official report (HRC., 2011). Compared with the nonpoint source load of each administrative region in 2000, the simulated annual loads tended to increase from 2003 to 2005 except in Kaifeng region. The most increased regions were Fuyang and Pingdingshan regions.

3.3.4 Crop yield simulation

The simulated corn yield and its spatial pattern were shown in Figure 11. The average annual yields were output at subbasin scale and ranged from 0.08 t km^{-2} $year^{-1}$ to 326.95 t km^{-2} $year^{-1}$ with the mean of 76.84 t km^{-2} $year^{-1}$. The yield of each administrative region was also summarized and compared with the data from statistical yearbooks from 2003 to 2005 (Henan Statistical Yearbook, 2003, 2004 and 2005) to test the simulation performance. The high-yield regions were Luohe, Fuyang and Zhoukou regions in the middle and down reaches, whose primary land use were dry land (93.12%, 95.87% and 93.18%, respectively). The yields of Luohe, Nanyang, Kaifeng regions were well simulated. The total yield was underestimated in the whole basin with the *bias* of 19.93%. The boundary mismatch between the administrative region and subbasin might contribute to the discrepancies, as well as the different cropping patterns in such huge region. Higher resolution remote sensing image and

field investigation might further improve the model performance.

4. Conclusions and Discussion

In this study, an integrated water system model (HE^XM) was developed for the effective governance of water issues emerged in the complex basins and the model performance was demonstrated in the Shaying River, China by simulating several key elements of major processes including runoff, water quality concentrations, nonpoint source pollutant load and crop yield.

Theoretically, the processes of hydrology, soil biogeochemistry, environment and ecology are strongly interactive with each other and the hydrological cycle is the critical linkage. Based on certain hypotheses, the integrated water system model describes these processes as a unified system rather than isolated systems for individual processes. HE^XM integrated these multi-scale processes, as well as their interactions based on the hypothesis that the cycles of water and nutrients (N, P and C) are inseparable and acted as the two critical linkages (shown in Fig.1). The processes of hydrological cycle and soil biogeochemistry were depicted more accurately by DTVGM and DNDC, respectively. The proposed model included the major hydrological elements (viz., soil water and evaporation, plant transpiration, runoff and water storage in the dams and sluices), environmental elements (viz., nonpoint source pollutant load of nutrient, water quality variables in water bodies), ecological elements (leaf area index, crop yield and greenhouse gas emission) in the complex basins which could be calibrated if the observations were collected.

In order to solve multi-scale of different water related processes, HE^XM designed the three levels of spatial simulation cells. Flow routing and matter transportation operated at the subbasin scale; water yield, evapotranspiration and soil erosion were considered at the landuse cell; plant physiological and ecological processes and soil biogeochemical processes were simulated at the crop cell. The crop cell varied with the agricultural cultivation structure and timing. The partition of different scale cells was based on the basic spatial information data (e.g. DEM, landuse), the landuse classification and the local agricultural management patterns. All the levels of spatial cell were visible, more reasonable and flexible than hydrologic response units of SWAT without reference to their actual spatial position (Neitsch *et al.*, 2011). HE^XM

considered all the processes at the daily scale, rather than the event or sub-daily scale due to the main purposes for integrated water system management.

HE^{X}M provided a reasonable tool in the integrated water system management by capturing some indicative elements of water related subsystems simultaneously .The case study was shown that hydrological and water environmental modules of HE^{X}M were proved to be closely linked by the statistical test of relative important parameters for runoff and water quality processes. The simulated daily runoffs at all the stations were well fitted with the observations. All the evaluation criteria were acceptable except at one or two stations. HE^{X}M well captured the variation of discontinuous daily NH_4-N concentration, and also properly simulated the spatial patterns of nonpoint source pollutant load and corn yield. The model performances were obviously improved in comparison with the prior SWAT model results. Moreover, HE^{X}M can better simulate runoff and water quality concentration at the daily scale in the highly regulated basins by making a comparison the results with and without dam regulation.

Restricted by the heterogeneity of spatial data in large basins and insufficient observations of every subsystems, not all the results were acceptable and several critical processes were not well calibrated (low flow events, crop yield, soil erosion and nonpoint source pollutant load, greenhouse gas emission, etc.). The model structure could be further developed with the new exploration of water related processes. More complex humanity activities and water-related processes in the dam regulation, agricultural management, urban area and economy system will be incorporated into this model once the interaction mechanisms with natural hydrologic cycle could be depicted accurately. Additionally, there are still several great challenges in combined calibration of multi-component and model uncertainty analysis because of the interactions and tradeoffs among different processes, as well as highly parameterization. It is very difficult to achieve the optima of every processes using the traditional step-by-step calibration. Advanced mathematic analysis technologies should be applied to in the next works, such as multi-objective optimization algorithm, Markov Chain Monte Carlo method.

Appendix A: Hydrological cycle module

The basic water balance equation is

$$P_i + SW_i = SW_{i\text{-}1} + Rs_i + Ea_i + Rss_i + Rg_i + In_i \quad \text{(A1)}$$

where P is precipitation (mm); SW is soil water moisture (mm); Ea is actual evapotranspiration including soil evaporation and plant transpiration (mm); Rs, Rss and Rg is surface runoff, interflow and baseflow (mm), respectively; In is the vegetation interception (mm) and i is the time step (days).

The actual soil evaporation (E_s, mm) and plant transpiration (E_p, mm) is determined by potential evapotranspiration (E_0, mm), leaf area index (LAI, m^2/m^2) and surface soil residues (rsd, t/ha) (Ritchie, 1972), viz.,

$$\begin{cases} E_a = E_t + E_s \le E_0 \\ E_p = f(LAI) \cdot E_0 \\ E_s = f(rsd) \cdot E_0 \end{cases} \quad \text{(A2)}$$

where $f()$ is a linear or nonlinear function. E_0 is calculated by Hargreaves method.

The surface runoff (R_s, mm) yield equation (TVGM; Xia *et al.*,2005) is as following.

$$Rs = g_1 \left(SW_u / W_{sat}\right)^{g_2} \cdot (P - In) \quad \text{(A3)}$$

where SW_u and W_{sat} are surface soil moisture and saturation moisture (mm), respectively; g_1 and g_2 are coefficient of basic runoff and soil moisture, respectively.

The interflow (R_{ss}, mm) and baseflow (R_g, mm) are considered as a linear storage-outflow relationship (Wang *et al.*, 2009).

$$\begin{cases} Rss = k_r \cdot SW_u \\ R_g = k_g \cdot SW_l \end{cases} \quad \text{(A4)}$$

where k_r and k_g are the yield coefficient of interflow and baseflow. SW_l is soil moisture of lower layer (mm).

The infiltration from the upper to lower soil layer is calculated using storage routing methodology (Neitsch *et al.*, 2011), viz.,

$$\begin{cases} W_{\inf} = (SW_u - W_{fc}) \cdot [1 - \exp(-t/T_{\inf})] \\ T_{\inf} = (W_{sat} - W_{fc}) / K_{sat} \end{cases} \quad \text{(A5)}$$

where W_{inf} is water infiltration amount on a given day (mm); W_{fc} is soil field capacity (mm); t and T_{inf} are time step and travel time for infiltration (hrs), respectively; and K_{sat} is saturated hydraulic conductivity (mm/hr).

The overland flow routing is calculated:

$$Q_{overl} = (Q'_{overl} + Q_{stor,i-1}) \cdot [1 - \exp(-1/T_{overl})] \quad \text{(A6)}$$

where Q_{overl} is overland flow discharged into main channel (mm), Q'_{overl} is lateral flow amount generated in the subbasin (mm), $Q_{stor,i-1}$ is lateral flow lagged from the previous day (mm) and T_{overl} is lateral flow travel time (days).

Appendix B: Soil biochemical module

B.1 Soil temperature (Williams *et al.*,1984):

$$T(Z,t) = \overline{T} + (AM/2 \cdot \cos[2\pi \cdot (t-200)/365] + TG - T(0,t)) \cdot \exp(-Z/DD) \quad \text{(B1)}$$

where Z is soil depth (mm); t is time step (days); $\overline{T}$ and TG are average annual temperature and surface temperature, respectively (°C); AM is annual variation amplitude of daily temperature; DD is damping depth of soil temperature (mm).

$$\begin{cases} DD = DP \cdot \exp\{(\ln(500/DP) \cdot [(1-\xi)/(1+\xi)]^2\} \\ DP = 1000 + 2500BD/[BD + 686\exp(-5.63BD)] \\ \xi = SW/[(0.356 - 0.144BD) \cdot Z_M] \\ TG_{IDA} = (1-AB) \cdot (T_{mx} + T_{mn})/2 \cdot (1 - RA/800) + T_{mx} \cdot RA/800 + AB \cdot TG_{IDA-1} \end{cases} \quad \text{(B2)}$$

where DP is maximum damping depth of soil temperature (mm); BD is soil bulk density (t/m^3); ξ is scale parameter; IDA is day of the year; AB is surface albedo; RA is daily solar radiation (ly).

B.2 C and N cycle (Li *et al.*, 1992):

Decomposition: The decomposition of resistant and labile C using the first order kinetic equation, viz.

$$dC/dt = \mu_{CLAY} \cdot \mu_{C:N} \cdot \mu_{t,n} \cdot [S \cdot k_1 + (1-S) \cdot k_2] \quad \text{(B3)}$$

where μ_{CLAY}, $\mu_{C:N}$ and $\mu_{t,n}$ is the reduction factor of clay content, *C: N* ratio and temperature for nitrification, respectively; *S* is labile fraction of organic *C* compounds; k_1 and k_2 is specific decomposition rate of labile faction and resistant fraction, respectively (days^{-1}).

The ammonia amount absorbed by clay and organic matters (FIX_{NH_4}) is estimated using the equation.

$$FIX_{NH_4} = [0.41 - 0.47 \cdot \log(NH_4)] \cdot (CLAY / CLAY_{\max}) \quad \text{(B4)}$$

where NH_4 is NH_4^+ concentration in the soil liquid (g/kg). *CLAY* and $CLAY_{\max}$ are clay content and the maximum clay content, respectively.

$$\begin{cases} \log(K_{NH_4}/K_{H_2O}) = \log(NH_{4m}/NH_{3m}) + pH \\ NH_{3m} = 10^{\{\log(NH_4) - (\log(K_{NH_4}) - \log(K_{H_2O})) + pH\} \cdot (CLAY / CLAY_{\max})} \\ AM = 2 \cdot (NH_3) \cdot (D \cdot t / 3.14)^{0.5} \end{cases} \quad \text{(B5)}$$

where K_{NH_4} and K_{H_2O} are dissociation constant for NH_4^+:NH_3 equilibrium, H^+: OH^- equilibrium, respectively; NH_{4m} and NH_{3m} are NH_4^+ and NH_3 concentration in the liquid phase, respectively (mol/L); *AM* and *D* are accumulated NH_3 loss (mol/cm^2) and diffusion coefficient (cm^2/d^2), respectively.

The nitrification rate (*dNNO*, kg/ha/day) is a function of the available NH_4^+, soil temperature and soil moisture. N_2O emission is a function of soil temperature and soil NH_4^+ concentration, viz.:

$$\begin{cases} dNNO = NH_4(t) \cdot [1 - \exp(-K_{35} \cdot \mu_{t,n} \cdot dt)] \cdot \mu_{m,n} \\ N_2O = (0.0014 \cdot NH_4 / 30.0) \cdot (0.54 + 0.51 \cdot T) / 15.8 \end{cases} \quad \text{(B6)}$$

where $NH_4(t)$ is available NH_4^+ (kg/ha); K_{35} is nitrification rate at 35 ^{o}C (mg/kg/ha); $\mu_{m,n}$ is moisture adjusted factor for nitrification.

Denitrification: The growth rate of denitrifier is proportional to their respective biomass, which is calculated with double Monod kinetics equation.

$$\begin{cases} (dB/dt)_g = \mu_{DN} \cdot B(t) \\ \mu_{DN} = \mu_{t,dn} \cdot (u_{NO_3} \cdot \mu_{PHNO_3} + u_{NO_2} \cdot \mu_{PHNO_2} + u_{N_2O} \cdot \mu_{PHN_2O}) \\ u_{N_xO_y} = u_{N_xO_y,\max} \cdot (C / K_{C,1/2} + C) \cdot (N_xO_y / K_{N_xO_y,1/2} + N_xO_y) \end{cases} \quad \text{(B7)}$$

where B is denitrifier biomass (kg); $(d\mathrm{B}/dt)_g$ is potential growth rate of denitrifier biomass (kg/ha/day); μ_{DN} is relative growth rate of the denitrifiers; $u_{N_xO_y}$ and $u_{N_xO_y,\max}$ are relative and maximum growth rate of NO_2^-, NO_3^- and N_2O. $\mu_{PHN_xO_y}$ and $\mu_{t,dn}$ are reduction factor of soil pH and temperature, respectively.

$$\begin{cases} \mu_{PHNO_3} = 7.14 \cdot (pH - 3.8)/22.8 \\ \mu_{PHNO_2} = 1.0 \\ \mu_{PHN_2O} = 7.22 \cdot (pH - 4.4)/18.8 \\ \mu_{t,dn} = \begin{cases} 2^{(T-22.5)/10} & if\ T < 60°C \\ 0 & if\ T \geq 60°C \end{cases} \end{cases} \quad \text{(B8)}$$

The death rate of denitrifier $(d\mathrm{B}/dt)_d$ (kg/ha/hr) is the proportional to denitrifier biomass, viz.

$$(d\mathrm{B}/dt)_d = M_C \cdot Y_C \cdot B(t) \quad \text{(B9)}$$

where M_C and Y_C are maintenance coefficient of C (1/hr), maximum growth yield of soluble C (kg/ha/hr), respectively.

The consumption rate of soluble C and CO_2 production is calculated as

$$\begin{cases} d\mathrm{C}_{\mathrm{con}}/dt = (\mu_{DN}/Y_C + M_C) \cdot B(t) \\ d\mathrm{CO}_2/dt = dC_{con,t}/dt - (d\mathrm{B}/dt)_d \end{cases} \quad \text{(B10)}$$

The NO_3^-, NO_2^- and N_2O consumption are calculated with Pirt's equation.

$$d\mathrm{N}_xO_y/dt = (u_{N_xO_y}/Y_{N_xO_y} + M_{N_xO_y} \cdot N_xO_y/N) \cdot B(t) \cdot \mu_{PHN_xO_y} \cdot \mu_{t,dn} \quad \text{(B11)}$$

N assimilation is calculated on the basis of the growth rates of denitrifiers and the C: N ratio ($CNR_{D:N}$) in the bacteria, viz.

$$(dN/dt)_{ass} = (dB/dt)_g \cdot (1/CNR_{D:N}) \quad \text{(B12)}$$

The emission rate is a function of adsorption coefficients of the gases in soils and to the air filled porosity of the soil.

$$\begin{cases} P(N_2) = 0.017 + ((0.025 - 0.0013 \cdot AD) \cdot PA \\ P(N_2O) = [30.0 \cdot (0.0006 + 0.0013 \cdot AD) + (0.013 - 0.005 \cdot AD)] \cdot PA \\ P(NO) = 0.5 \cdot [(0.0006 + 0.0013 \cdot AD) + (0.013 - 0.005 \cdot AD) \cdot PA] \end{cases} \quad \text{(B13)}$$

where $P(N_2)$, $P(NO)$ and $P(N_2O)$ are emission rate of N_2, NO, N_2O during a day, respectively; *PA* and *AD* are air-filled fraction of the total porosity and adsorption factor depending on clay content in the soil, respectively.

Nitrate leaching: The NO_3^-N leaching rate is a function of clay content, organic C content and water infiltration in the soil layer.

$$Leach_{NO_3} = W_{\inf} \cdot \mu_{CLAY} \cdot \mu_{soc} \quad \text{(B14)}$$

where $Leach_{NO_3}$ is the NO_3^-N leaching rate; μ_{CLAY} and μ_{soc} are the influence coefficient of clay content and organic C in the soil layer, respectively.

B.3 P cycle

The descriptions of P mineralization, decomposition and sorption are adopted from Neitsch *et al.* (2011) and provided as the supplementary material.

Appendix C : Dams regulation module (Zhang *et al.*, 2013)

The water balance model is used to consider inflow, outflow, precipitation, evapotranspiration and seepage of dam or sluice. The equation is:

$$\Delta V = V_{flowin} - V_{flowout} + V_{pcp} - V_{evap} - V_{seep} \quad \text{(C1)}$$

where ΔV, V_{flowin} and $V_{flowout}$ are water storage variation, water volume of entering and flowing out, respectively (m^3) which are calculated by HCM; V_{pcp}, V_{evap} and V_{seep} are precipitation, evaporation and seepage volume, respectively (m^3), which are functions of water surface area and vary with water storage.

In the design of dam in China, there is a definite correspondence relationship among water level, storage volume and outflow. The water discharge is determined by the water level or water storage volume. Thus, the relationship is described by equation

$$\begin{cases} V_{flowout} = f'(V,H) \\ SA = f''(V,H) \end{cases} \quad \text{(C2)}$$

where *V* and *H* are water storage volume and water level during a day, respectively, $f'()$, $f''()$ are the functions to be determined.

Appendix D: Evaluation indices of model performance

Bias: $$bias = \sum_{i=1}^{N}(O_i - S_i) \Big/ \sum_{i=1}^{N} Q_i \quad \text{(D1)}$$

Relative error: $$re = \sum_{i=1}^{N} \frac{O_i - S_i}{O_i} \times 100\% \quad \text{(D2)}$$

Root mean square error: $$RMSE = \sqrt{\sum_{i=1}^{N}(O_i - S_i)^2 / N} \quad \text{(D3)}$$

Correlation coefficient: $$r = \sum_{i=1}^{N}(O_i - \overline{O}) \cdot (S_i - \overline{S}) \Big/ \sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2 \cdot \sum_{i=1}^{N}(S_i - \overline{S})^2} \quad \text{(D4)}$$

Coefficient of efficiency: $$NS = 1 - \sum_{i=1}^{N}(O_i - S_i)^2 \Big/ \sum_{i=1}^{N}(O_i - \overline{O})^2 \quad \text{(D5)}$$

where, O_i and S_i are the i^{th} observed and simulated value, respectively; $\overline{O}$ and $\overline{S}$ are the average observed and simulated value, respectively. N is the length of series.

Acknowledgements

This study was supported by the Natural Science Foundation of China (No. 41271005), the Key Project for the Strategic Science Plan in IGSNRR, CAS (No. 2012ZD003), the Endeavour Research Fellowship and China Visiting Scholar Project from China Scholarship Council. Thanks to Dr Yongqiang Zhang, Mr. James R Frankenberger during our internal review procedure, the editor (Dr. Christian Stamm) and two anonymous reviewers for their valuable comments which helped improving the quality and presentation of the manuscript.

References

Arthington, A.H., 2012. *Environmental Flows: Saving Rivers in the Third Millennium*. University of California Press, Berkeley, CA. 406pp.

Arnold, J. G., Srinivasan,R., Muttiah, R. S., and Williams, J. R.: Large-area hydrologic modeling and assessment: Part I. Model development. *J. Am. Water Resour. Assoc.* 34(1):73-89,1998.

Brakensiek, D.L. Estimating the effective capillary pressure in the Green and Ampt infiltration equation. *Water Resour. Res.*, 13(3): 680-682, 1977.

Bicknell, B. R., Imhoff, J. C., Kittle, J. L., Donigian, A. S., and Johanson, R. C.: Hydrologic Simulation Program –FORTRAN (HSPF): User's Manual for Release 10. Report No. EPA/600/R-93/174. Athens, Ga.: U.S. EPA Environmental Research Lab, 1993.

Bingner,R.L., Theurer, F.D., and Yuan,Y.: AnnAGNPS Technical Processes: Documentation Version 2. Unpublished Report, USDA-ARS National Sedimentation Laboratory, Oxford, Miss, 2001.

Blöschl, G., Sivapalan,M. Scale issues in hydrological modelling: a review. *Hydrol. Process.*, 9(3–4): 251-290, 1995.

Borah, D. K., and Bera, M.: Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Trans. ASAE* 46(6): 1553-1566, 2003.

Bouraoui,F., and Dillaha,T.A.:ANSWERS-2000: Runoff and sediment transport model. *J. Environ. Eng.*, 122(6): 493-502, 1996.

Brown, L. C. and Barnwel,T. O.: The enhanced stream water quality models QUAL2E and QUAL2E-UNCAS: documentation and user manual. Env. Res. Laboratory. US EPA, 1987.

Burt,T.P. and Pinay,G.: Linking hydrology and biogeochemistry in complex landscapes. *Prog. Phys. Geog.*, 29(3): 297-316, 2005.

Ceola, S., Montanari, A., Koutsoyiannis, D. Toward a theoretical framework for integrated modeling of hydrological change. *WIREs Water*, 1:427–438,2014.

China's national standard (CNS), (2007). Current land use condition classification (GB/T21010-2007).

China State Environmental Protection Administration (CSEPA), (2002). *Environmental quality standards for surface water* -GB 3838-2002. Beijing: China Environmental Science Press. (In Chinese).

Deng, J., Zhu,B., Zhou,Z., *et al.* : Modeling nitrogen loadings from agricultural soils in southwest China with modified DNDC. *J. Geophys. Res.: Biogeosci.*

(2005–2012), 116(G2) , 2011.

Di Toro, D.M., Fitzpatrick, J.J., and Thomann, R.V.: Water quality analysis simulation program (WASP) and model verification program (MVP)-Documentation. Hydroscience, Inc., Westwood, NY, for U.S. EPA, Duluth, MN, Contract No. 68-01-3872,1983.

Duan, Q., Sorooshian, S., and Gupta, V. K.: Optimal use of the SCE-UA global optimization method for calibrating watershed models. *J.Hydrol.*, 158(3): 265-284, 1994.

Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review. *Hydrol. Sci. J.*,55:58-78,2010.

Gassman, P.W., Reyes, M.R., Green,C.H., and Arnold, A.G.: The soil and water assessment tool: historical development, applications, and future research directions. *T. ASABE*, 1211-1250, 2007.

Gassmann, M., Brito, D., Olsson, O., 2014. Estimation of phosphorus export from a Mediterranean agricultural catchment with scarce data. *Hydrol. Sci. J.* 59 (1), 221–233.

Gleick, P.H. Water in crisis: paths to sustainable water use. *Ecol. Appl.*, 8(3): 571-579, 1998.

Goldberg,D.E.: Genetic algorithms in search, optimization, and machine learning. Reading Menlo Park: Addison-wesley, 1989.

Hamrick, J. M.: A three-dimensional environmental fluid dynamics computer code: theoretical and computational aspects. The College of William and Mary, Virginia Institute of Marine Science. Special Report 317, 1992.

Hargreaves, G. H., and Samani, Z. A.: Estimating potential evapotranspiration. *J. Irrigat. Drain. Div.*, 108(3), 225-230,1982.

Henan Statistical Yearbook in 2003,2004 and 2005. China Statistics Press, Beijing.

Horst, W.J., Kamh, M., Jibrin, J.M. and Chude, V.O.: Agronomic measures for increasing P availability to crops, *Plant. Soil.* 237: 211-223, 2001.

HRC (The Huai River Commission of the Ministry of Water Resources, P.R.C), *Huai River Basin and Shangdong Peninsula Integrated Water Resources Plan*, 2011.

Johnes P J.: Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. *J. Hydrol.*, 183(3): 323-349, 1996.

Jones, C.A. and Kiniry J.R., eds. CERES-Maize. College Station: Texas A&M Univ. Press, 1986.

Jordan, Y. C., Ghulam, A., and Hartling, S.: Traits of surface water pollution under climate and land use changes: A remote sensing and hydrological modeling approach. *Earth-Sci. Rev.*, 128, 181-2014-3-9,2014.

Kirchner J.W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res*, 42(3) , 2006.

Kirchner, J.W.: A double paradox in catchment hydrology and geochemistry. *Hydrol. Process.* 17(4): 871-874,2003.

Kindler, J.: Integrated water resources management: the meanders. *Water Int.*, 25:312-319, 2000.

King, K. W., Arnold, J. G., Bingner, R. L. Comparison of Green-Ampt and curve number methods on Goodwin Creek watershed using SWAT. *Trans. ASAE*, 42(4), 919-925,1999.

Kennedy, J., and Eberhart,R.: Particle swarm optimization[C]//Proceedings of IEEE international conference on neural networks, 4(2): 1942-1948,1995.

Letcher, R.A., Croke, B.F.W., Jakeman, A.J. Integrated assessment modelling for water resource allocation and management: A generalised conceptual framework. *Environ. Modell. Softw.*, 22(5): 733-742, 2007.

Li,C., Frolking,S., Frolking,T.A.: A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity. *J. Geophys. Res.*(1984–2012), 97(D9): 9759-9776, 1992.

Lohse,K.A., Brooks,P.D., McIntosh,J.C., Meixner, T., and Huxman, T. E. Interactions between biogeochemistry and hydrologic systems. *Ann. Rev. Environ. Resour.*, 34: 65-96, 2009.

Mantovan, P., and Todini, E.: Hydrological forecasting uncertainty assessment:

Incoherence of the GLUE methodology, *J. Hydrol.*, 330, 368–381, 2006.

Mantovan, P., Todini, E. and Martina, M. L.V.: Reply to comment by Keith Beven, Paul Smith, and Jim Freer on "Hydrological forecasting uncer tainty assessment: Incoherence of the GLUE methodology", *J. Hydrol.*, 338, 319–324,2007.

Milly, P. C. D., Wetherald, R. T., Dunne, K. A. and Delworth,T. L.: Increasing risk of great floods in a changing climate. *Nature* 415, 514-517,2002.

McDonnell, J.J., Sivapalan, M., Vach éK., *et al*. : Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology. *Water Resour. Res*., 43(7) , 2007.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Binger, R. D. Harmel, and T. Veith.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, *T. ASABE*, 50(3), 885-900,2007.

Neitsch, S., Arnold, J., Kiniry, J., Williams, J.R., 2011. *SWAT2009 Theoretical Documentation*. Texas Water Ressources Institute, Temple, Texas.

Onstad, C. A. and Foster, G. R.: Erosion modeling on a watershed. *T.ASAE* 18(2):288-292,1975.

Pimentel, D., Berger, B., Filiberto, D., Newton, M., Wolfe, B., Karabinakis, E., *et al.*: Water resources: agricultural and environmental issues. *BioSci.*, 54(10), 909-918, 2004.

Pushpalatha, R., Perrin, C., Le Moine, N., and Andr éassian, V.: A review of efficiency criteria suitable for evaluating low-flow simulations. *J.Hydrol.*, 420-421, 171-182, 2012.

Rallison, R.E. and Miller.,N.: Past, present and future SCS runoff procedure. 353-364,1981. In V.P. Singh (ed.). Rainfall runoff relationship. Water Resources Publication, Littleton, CO.

Revenga,C., Brunner,J., Henninger,N., Kassem,K., and Payne,R.: Pilot analysis of global ecosystems: freshwater systems.Washington DC:World Resources Institute, 2000.On-line at: http://www.wri.org/wr2000/freshwater_page.htm.

Ritchie, J.T.: A model for predicting evaporation from a row crop with incomplete cover. *Water Resour. Res*. 8:1205-1213,1972.

Ritter.A, and Muñoz-Carpena,R.: Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *J. Hydrol.*, 480: 33-45, 2013.

Schiermeier,Q.: Increased flood risk linked to global warming. *Nature*, 470, 316,2011.

Sharpley, A.N. and Williams, J.R.: EPIC-erosion/productivity impact calculator: 1. Model documentation. Technical Bulletin-United States Department of Agriculture, 1990.

Singh,V.P. and Woolhiser, D.A.: Mathematical modeling of watershed hydrology. *J. Hydrol. Eng.*, 7(4): 270-292, 2002.

Sivapalan, M. and Kalma, J. D.: Scale problems in hydrology: contributions of the Robertson Workshop. *Hydrol. Process.*, 9(3-4), 243-250,1995.

Sivapalan, M.: Process complexity at hillslope scale, process simplicity at the watershed scale: Is there a connection? *Hydrol. Process.*, 17: 1037–1041, 2003.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J.Hydrol.*, 324(1), 10-23,2006.

Vinogradov, Y. B., Semenova, O. M., and Vinogradova, T. A.: An approach to the scaling problem in hydrological modelling: the deterministic modelling hydrological system. *Hydrol. Process.*, 25(7), 1055-1073, 2011.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., *et al.*: Global threats to human water security and river biodiversity. *Nature*, 467(7315), 555-561,2010.

Wang, G., Xia, J., and Chen, J.: Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: A case study of the Chaobai River basin in northern China. *Water Resour. Res.*, 45(7) ,2009.

Wang, J.Q., Ma, W.Q., Jiang, R.F. and Zhang, F.S.: Analysis about amount and ratio of basal fertilizer and topdressing fertilizer on rice, wheat, maize in China. *Chin. J.Soil Sci..*, 39(2):329-333,2008.(In Chinese)

Watson, N.: Integrated river basin management: A case for collaboration. *Intl. J. River Basin Manag.*, 4(2): 243–257,2004.

Wigmosta,M.S., Vail,L.W., and Lettenmaier,D.P.: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, 30(6): 1665-1679, 1994.

Wilhite, D. A.: Drought and water crises: science, technology and management issues. CRC Press, 2005.

Williams, J.R., Jones, C.A., and Dyke, P.T. Modeling approach to determining the relationship between erosion and soil productivity. *Trans. ASAE*, 27(1): 129-144,1984.

Williams, J. R., Jones, C. A., Kiniry, J. R., and Spanel, D. A. The EPIC crop growth model. . *Trans. ASAE*, 32(2):497-511,1989.

Xia, J., Wang, G.S., Tan, G., Ye, A.Z., and Huang, G. H.: Development of distributed time-variant gain model for nonlinear hydrological systems. *Sci. China: Earth Sci.*, 48(6), 713-723,2005.

Xia, J.: Identification of a constrained nonlinear hydrological system described by Volterra Functional Series, *Water Resour. Res.*, 27(9): 2415−2420, 1991.

Xing, G. X., and Zhu, Z. L.: An assessment of N loss from agricultural fields to the environment in China. *Nutr. Cycl. Agroecosys.*, 57(1): 67-73, 2000.

Zhai, X.Y., Zhang, Y.Y., Wang X.L., Xia J. and Liang, T.: Non-point source pollution modeling using Soil and Water Assessment Tool and its parameter sensitivity analysis in Xin'anjiang Catchment, China. *Hydrol. Process*. 28, 1627-1640,2014.

Zhang, Y.Y., Xia, J., Shao, Q.X., and Zhai, X.Y.: Water quantity and quality simulation by improved SWAT in highly regulated Huai River Basin of China. *Stoch. Env Res. Risk A.*, 27(1), 11-27,2013.

Zhou, Y., Khu, S. T., Xi, B., Su, J., Hao, F., Wu, J., and Huo, S.: Status and challenges of water pollution problems in China: learning from the European experience. *Environ. Earth Sci.*, 1-12, 2014.

Zhu, Z.L.: Loss of fertilizer N from plants-soil system and the strategies and techniques for its reduction. *Soil Environ. Sci.*, 9(1):1-6, 2000. (in Chinese)

1 Table 1. The widely used models and their main focuses

Element / Model	Hydrology			Environment			Ecology	
	soil water	surface water	ground water	nonpoint source	instream water quality	lake water quality	crop growth	soil biochemistry
HSPF		√		√	√	√		
SHE	√	√	√		√			
ANSWERS	√	√		√	√	√		
AnnAGNPS	√	√		√	√	√		
WASP		√			√	√		
QUAL2K					√	√		
EFDC		√			√	√		
DNDC		√					√	√
EPIC							√	√
SWAT	√	√	√	√	√	√	√	√

2

1 Table 2. The data sets and their categories used in HE^{X}M

category	Data	Objectives of HE^{X}M	Controlled processes of HE^{X}M
GIS	DEM	Elevation, slopes and lengths of each subbasin and channel	Hydrology and environment
	Land use map	Land use types and their corresponding areas in each subbasin	Hydrology, environment and ecology
	Soil map	Soil physical properties of each subbasin such as bulk density, texture, saturated conductivity	
Weather	Daily precipitation	Daily precipitation of each subbasin	Hydrology
	Daily maximum and minimum temperature	Daily maximum and minimum temperature of each subbasin	
Hydrology	Daily runoff observations	Hydrological parameter calibration	Hydrology
Environment	The outlets and the discharge data	Point source pollutant load	Environment
	The concentration observation	Environmental parameter calibration	
Ecology	Crop yield, Leaf area index	Ecological parameter calibration	Ecology
Economy	The basic economic statistical indictors	Populations, breeding stock of large animals and livestock, water withdrawal in each subbasin	Hydrology and environment
Water projects	The reservoir's design data attribute parameters	Regulation rules of reservoirs or sluices	Hydrology
Agricultural management	Fertilization time and amount, the time of seeding and harvest, crop types	Agricultural management rules of each subbasin	Environment and ecology

Table 3 The agricultural management scheme in the Shaying River Catchment

Crop	Management	Time	Ratio of annual fertilization	
		(month-day)	TN	TP
Early rice	Basal fertilization	4-1	0.60	0.86
	Planting	4-15	-	-
	Topdressing Fertilization	5-1	0.40	0.14
	Harvest & Kill	7-31	-	-
Late rice	Basal fertilization	8-1	0.50	0.86
	Planting	8-15	-	-
	Topdressing Fertilization	9-1	0.50	0.14
	Harvest & Kill	10-31	-	-
Winter wheat	Basal fertilization	10-1	0.64	0.02
	Planting	10-15	-	-
	Topdressing Fertilization	1-1	0.36	0.98
	Harvest & Kill	6-1	-	-
Corn	Basal fertilization	6-1	0.41	0.88
	Planting	6-15	-	-
	Topdressing Fertilization	7-15	0.59	0.12
	Harvest & Kill	9-30	-	-

Table 4 Sensitive parameters, corresponding ranges, and relative importance for runoff and NH_4-N simulation

Name	Min	Max	Definition	Relative Importance for runoff (%)	Relative Importance for NH_4-N (%)
WMc	0.20	0.45	Field capacity of soil	32.73	11.10
WM	0.45	0.75	Saturated moisture capacity of soil	11.68	11.83
g_1	0	3	Basic surface runoff coefficient	7.30	10.34
g_2	0	3	Influence coefficient of soil moisture	10.54	12.11
KET_p	0	3	Adjustment factor of evapotranspiration	23.21	10.71
k_r	0	1	Interflow yield coefficient	9.55	3.20
T_g	1	100	Delay time for aquifer recharge	1.74	-
k_g	0	1	Baseflow yield coefficient	2.91	-
f_c	0	120	Steady state infiltration rate	0.33	-
rk1	0.02	3.4	COD deoxygenation rate at 20 ^oC	-	6.62
rk3	-0.36	0.36	COD settling rate at 20 ^oC	-	3.60
bc1	0.1	1	Bio-oxidation rate of NH_4-N at 20 ^oC	-	1.97
res_set (NH_4-N)	0	100	Settling rate of NH_4-N in the reservoirs	-	14.17
res_rk1	0.02	3.4	COD deoxygenation rate in the reservoirs at 20^oC	-	2.12
res_bc1	0.1	1.0	Bio-oxidation rate of NH_4-N in the reservoirs at 20 ^oC	-	4.51
Total relative importance				100.00	92.27

Table 5 Runoff simulation results of regulated and less-regulated stations (given in brackets) and the comparisons with the existing study. SWAT does not have daily results because it is calibrated at monthly scale.

Stations	Periods	Daily flow				Monthly flow: HE^{X}M(SWAT)			
		bias	r	NS	f	bias	r	NS	f
Regulated stations									
Huangqiao	Calibration	0.00	0.86	0.72	0.14	0.00 (0.05)	0.88(0.69)	0.75(0.40)	0.12(0.32)
	Validation	0.05	0.61	0.23	0.40	0.05(0.21)	0.81(0.83)	0.52(-0.25)	0.24(0.54)
Mawan	Calibration	0.00	0.68	0.46	0.29	0.00(0.05)	0.74(0.82)	0.54(0.66)	0.24(0.19)
	Validation	-0.44	0.63	0.38	0.48	-0.44(0.46)	0.79(0.95)	0.52(0.62)	0.38(0.30)
Luohe	Calibration	0.00	0.84	0.70	0.15	0.00(-0.04)	0.87(0.94)	0.71(0.87)	0.14(0.08)
	Validation	-0.52	0.75	0.51	0.42	-0.52(-0.56)	0.87(0.81)	0.67(0.54)	0.33(0.40)
Zhoukou	Calibration	0.24	0.87	0.73	0.21	0.24(0.10)	0.90(0.94)	0.76(0.88)	0.19(0.12)
	Validation	0.41	0.79	0.55	0.36	0.41(0.34)	0.91(0.89)	0.70(0.68)	0.26(0.26)
Huaidian	Calibration	0.03	0.88	0.77	0.13	0.03(-0.10)	0.91(0.85)	0.81(0.72)	0.10(0.18)
	Validation	0.12	0.76	0.54	0.27	0.12(-0.01)	0.87(0.72)	0.70(0.46)	0.18(0.28)
Fuyang	Calibration	0.00	0.90	0.81	0.10	0.00(0.03)	0.95(0.92)	0.89(0.84)	0.05(0.09)
	Validation	0.14	0.88	0.76	0.17	0.14(-0.41)	0.94(0.85)	0.86(0.63)	0.11(0.31)
Yingshang	Calibration	-0.13	0.92	0.84	0.12	-0.13(-0.34)	0.92(0.82)	0.84(0.61)	0.12(0.30)
	Validation	0.16	0.87	0.74	0.18	0.16(-0.27)	0.93(0.85)	0.82(0.69)	0.13(0.24)
Less–regulated stations									
Shenqiu	Calibration	0.00	0.91	0.82	0.09	0.00(-0.09)	0.94(0.81)	0.88(0.54)	0.06(0.25)
	Validation	-0.13	0.83	0.67	0.21	-0.14(-0.72)	0.98(0.78)	0.94(0.12)	0.08(0.61)

Table 6. The comparison of runoff simulation results at regulated stations when the dam regulation is considered or not. Range means the difference of objective function value between considering regulation and without considering regulation. If the range value is less than 0.0, the simulation of considering regulation is better than that of without regulation. Otherwise, the simulation of without regulation is better.

Stations	Regulated capacity (%)	Flow event	Regulation considered				Regulation not considered				Range
			bias	r	NS	f	bias	r	NS	f	
Huangqiao	2.01	High	0.19	0.79	0.53	0.29	-0.11	0.80	0.47	0.28	0.01
		Low	-2.83	0.01	-9.89	4.90	-4.80	0.03	-17.42	8.06	-3.16
		Average	0.02	0.79	0.59	0.21	-0.36	0.81	0.51	0.35	-0.14
Mawan	0.29	High	0.09	0.67	0.45	0.32	-0.46	0.69	0.37	0.47	-0.15
		Low	-	-	-	-	-	-	-	-	-
		Average	-0.14	0.66	0.44	0.35	-1.01	0.68	0.26	0.69	-0.34
Luohe	0.26	High	-0.01	0.83	0.68	0.17	-0.42	0.82	0.54	0.35	-0.18
		Low	-1.82	0.01	-81.02	28.28	-3.96	0.02	-124.70	43.56	-15.28
		Average	-0.15	0.82	0.66	0.22	-0.68	0.82	0.51	0.45	-0.23
Zhoukou	1.31	High	0.28	0.85	0.67	0.26	-0.24	0.85	0.64	0.25	0.01
		Low	0.48	0.04	-7.91	3.45	-1.65	0.16	-20.29	7.93	-4.48
		Average	0.30	0.85	0.70	0.25	-0.41	0.86	0.63	0.30	-0.06
Huaidian	1.37	High	0.12	0.85	0.71	0.19	-0.47	0.85	0.46	0.39	-0.20
		Low	-0.35	0.06	-9.49	3.93	-2.67	0.04	-37.82	14.15	-10.22
		Average	0.06	0.86	0.73	0.16	-0.74	0.86	0.42	0.49	-0.33
Fuyang	2.21	High	0.10	0.89	0.78	0.15	-0.29	0.89	0.69	0.24	-0.09
		Low	-0.40	0.04	-6.09	2.82	-2.28	0.06	-21.54	8.59	-5.77
		Average	0.05	0.90	0.80	0.12	-0.50	0.90	0.68	0.31	-0.19
Yingshang	1.76	High	0.11	0.88	0.77	0.15	-0.35	0.88	0.68	0.26	-0.11
		Low	-0.39	0.02	-8.39	3.59	-2.49	0.00	-28.62	11.04	-7.45
		Average	0.05	0.89	0.79	0.12	-0.60	0.89	0.66	0.35	-0.23

Table 7 The comparison of NH_4-N simulation results between HE^XM and improved SWAT, and between considering dams regulation and no regulation

Stations	Periods	Unregulated			Regulated: HE^XM(SWAT)		
Regulated stations		bias	r	f	bias	r	f
Luohe	Calibration	-0.67	0.60	0.54	-0.02(-0.13)	0.93(0.25)	0.05(0.44)
	Validation	-	-	-	-	-	-
Zhoukou	Calibration	-0.56	0.38	0.59	0.29(0.01)	0.61(0.66)	0.34(0.18)
	Validation	-1.35	0.66	0.85	0.27(0.19)	0.56(0.04)	0.36(0.57)
Huaidian	Calibration	0.49	0.80	0.35	0.22(0.01)	0.73(0.42)	0.25(0.30)
	Validation	0.22	0.51	0.36	0.02(0.02)	0.67(0.29)	0.18(0.37)
Fuyang	Calibration	0.26	0.80	0.23	0.28(0.00)	0.78(-0.20)	0.25(0.60)
	Validation	-0.38	0.56	0.41	-0.27(-1.13)	0.76(0.41)	0.26(0.86)
Yingshang	Calibration	0.25	0.58	0.34	0.24(-0.13)	0.79(0.31)	0.23(0.41)
	Validation	-0.76	0.62	0.57	-0.24(0.49)	0.49(0.46)	0.38(0.51)
Less-regulated stations							
Shenqiu	Calibration	0.13	0.62	0.26	0.13(-)	0.62(-)	0.26(-)
	Validation	0.16	0.41	0.37	0.16(0.27)	0.41(0.33)	0.37(0.47)
Fantaizi	Calibration	0.38	0.51	0.44	0.45(-0.01)	0.69(0.18)	0.38(0.42)
	Validation	-1.02	0.73	0.64	-0.62(0.54)	0.61(0.49)	0.51(0.53)

List of Figure Captions

Figure 1. The structure of HE^{X}M and the interactions among the major modules (1: hydrological part; 2: water environmental part; 3: ecological part; 4: dams regulation part; 5: parameter analysis tool)

Figure 2. The flowchart of hydrological cycle module in HE^{X}M and the interactions with other modules

Figure 3. The flowchart of soil biochemical module (a) and crop growth module (b) in ecological part of HE^{X}M and the interactions with other modules

Figure 4. The flowchart of soil erosion (a), matter migration (b) and water quality (c) module in water environmental part of HE^{X}M and the interactions with other modules

Figure 5. The multiple scale of different processes and the solutions of HE^{X}M

Figure 6. The location of study area

Figure 7. The digital delineation of subbasin, population, animal stock and fertilization for HE^{X}M

Figure 8. The daily runoff simulation at all the stations

Figure 9. The simulated NH_4-N concentration variation at all the situations

Figure 10. The spatial pattern of nonpoint source NH_4-N load and paddy area at the subbasin and regional scale in Shaying River Catchment

Figure 11. The spatial pattern of corn yield at the subbasin and regional scale in Shaying River Catchment

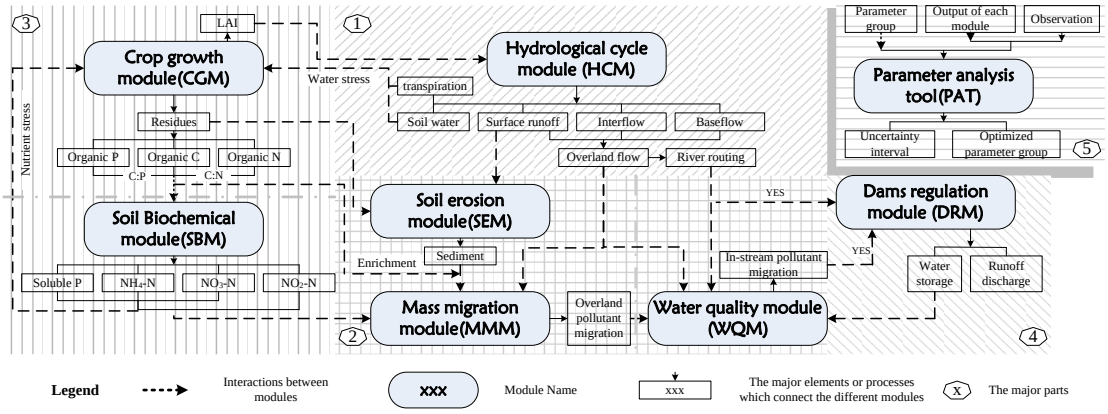


Figure 1. The structure of HE^{X}M and the interactions among the major modules (1: hydrological part; 2: water environmental part; 3: ecological part; 4: dams regulation part; 5: parameter analysis tool)

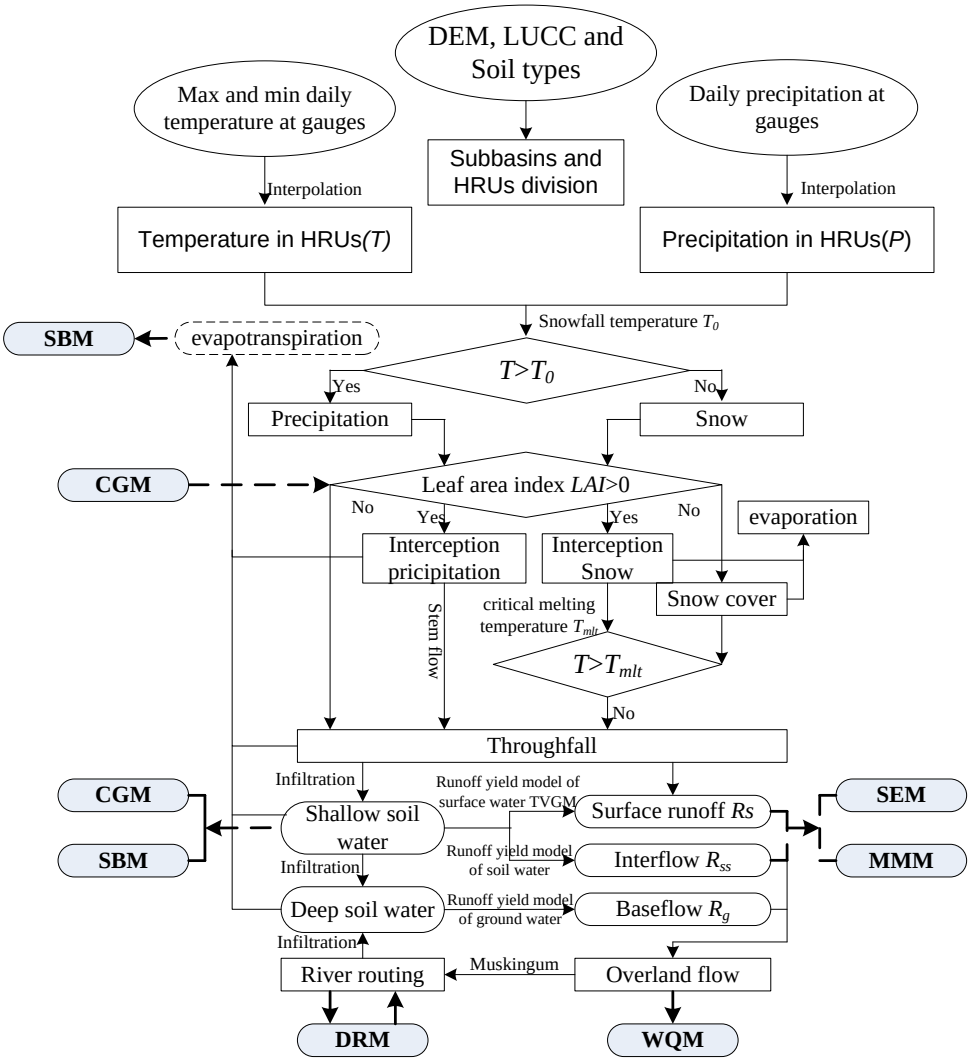


Figure 2. The flowchart of hydrological cycle module in HE^{X}M and the interactions with other modules

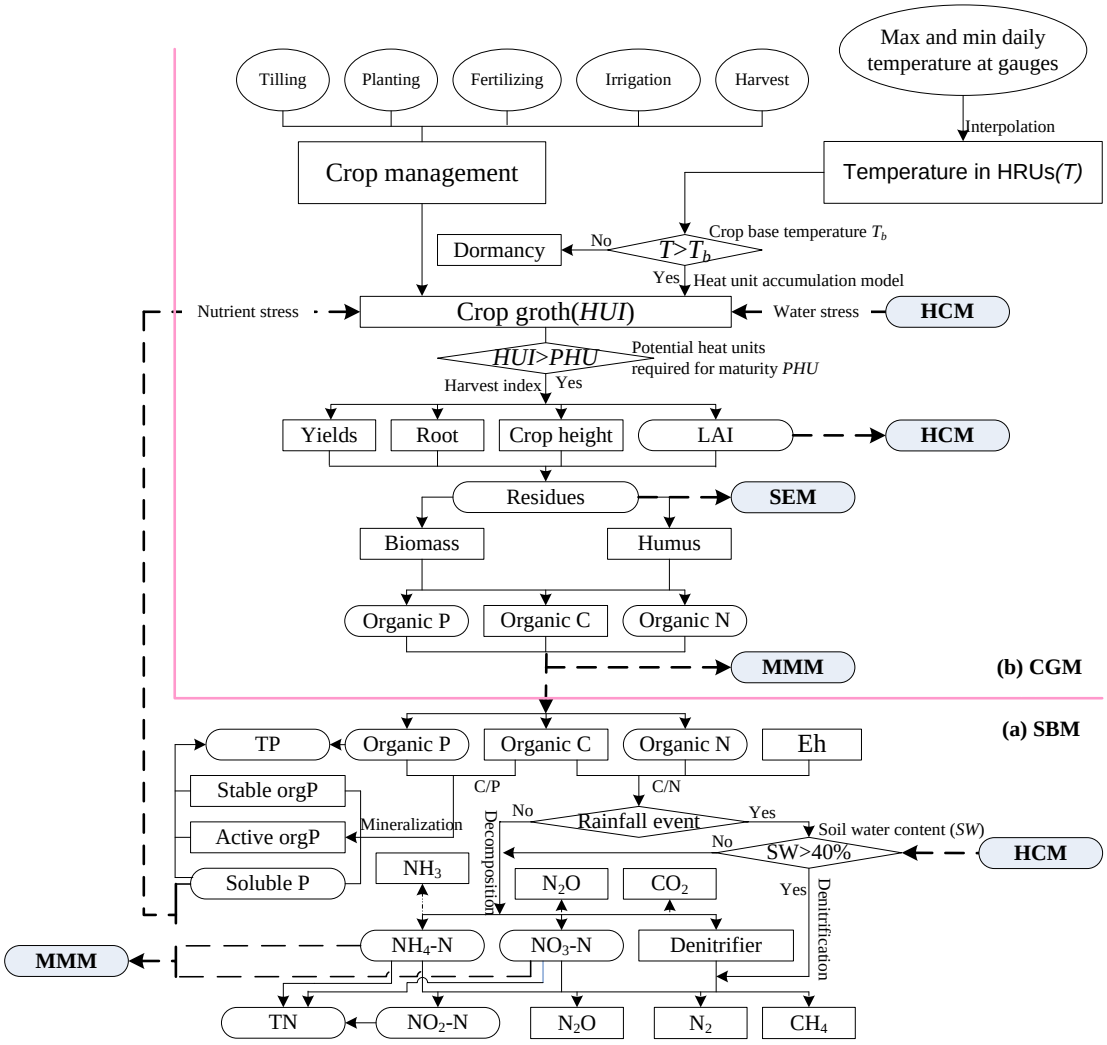


Figure 3. The flowchart of soil biochemical module (a) and crop growth module (b) in ecological part of HE^{X}M and the interactions with other modules

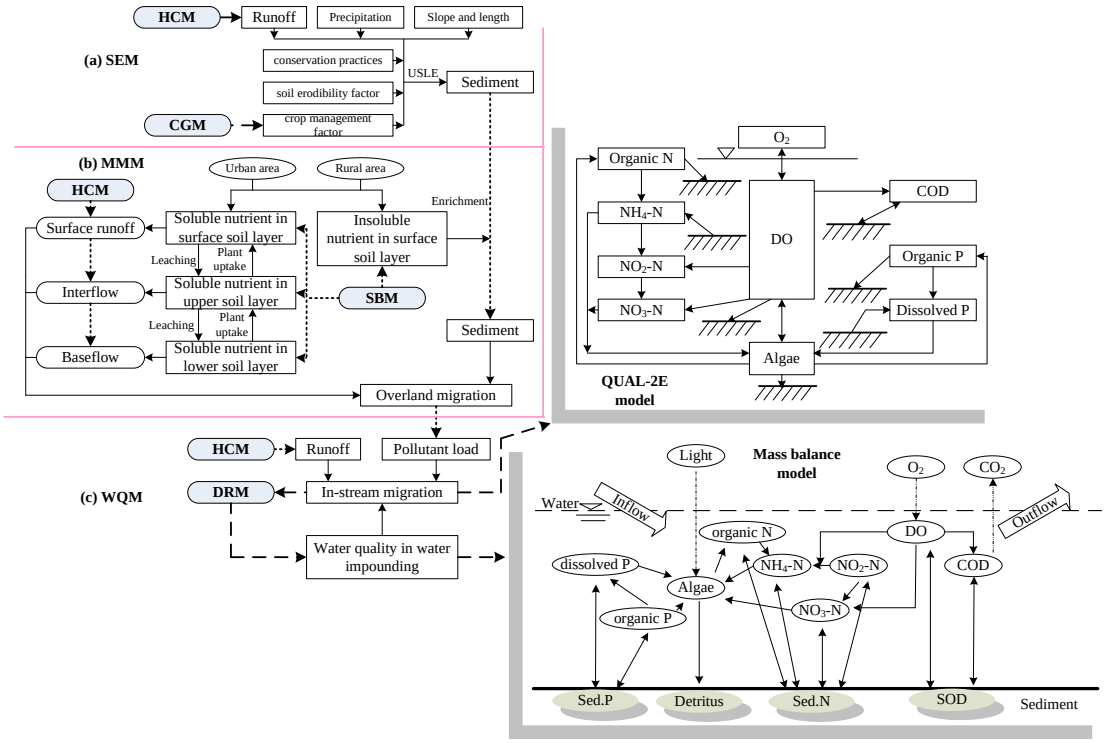


Figure 4. The flowchart of soil erosion (a), matter migration (b) and water quality (c) module in water environmental part of HE^{X}M and the interactions with other modules

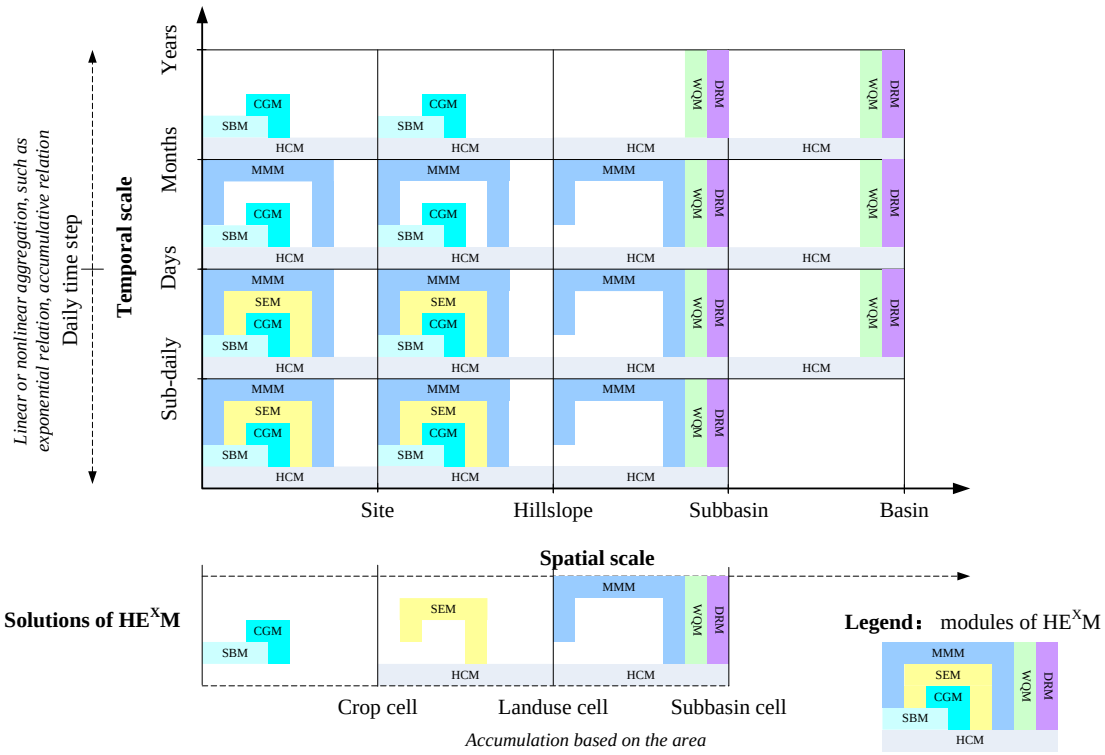


Figure 5. The multiple scales of different processes and the solutions of HE^{X}M

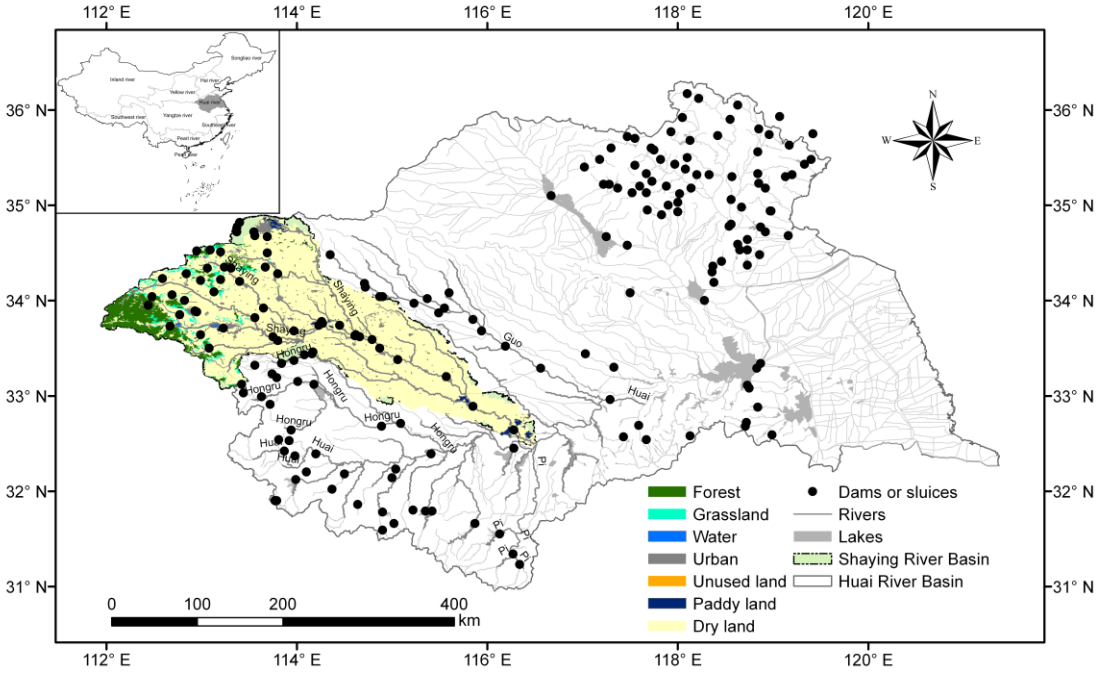


Figure 6. The location of study area

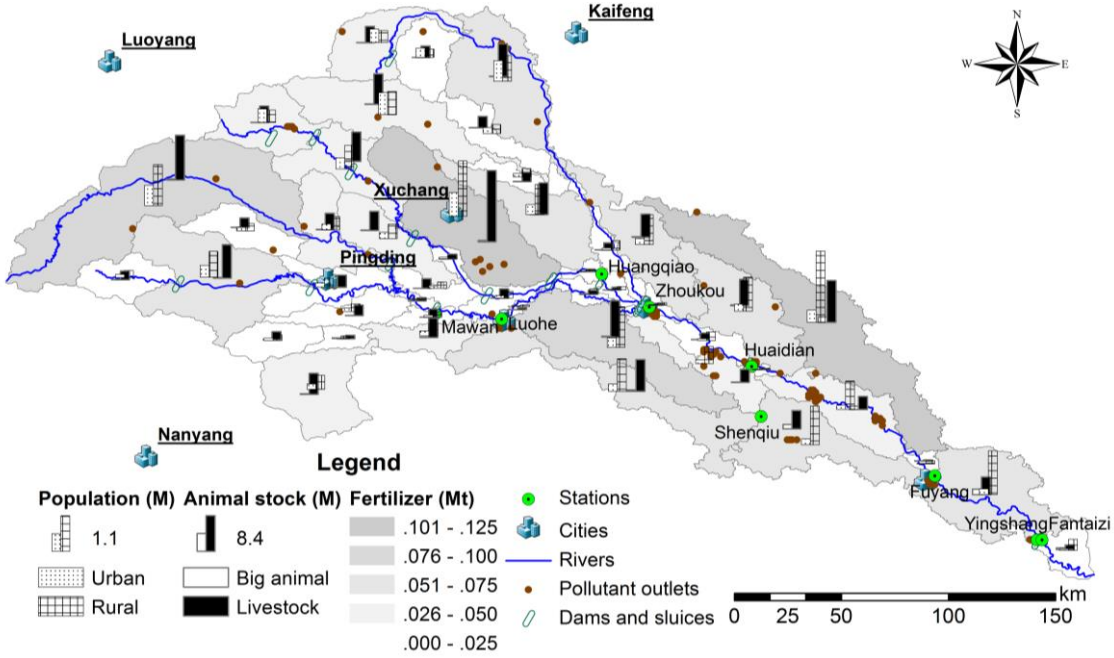
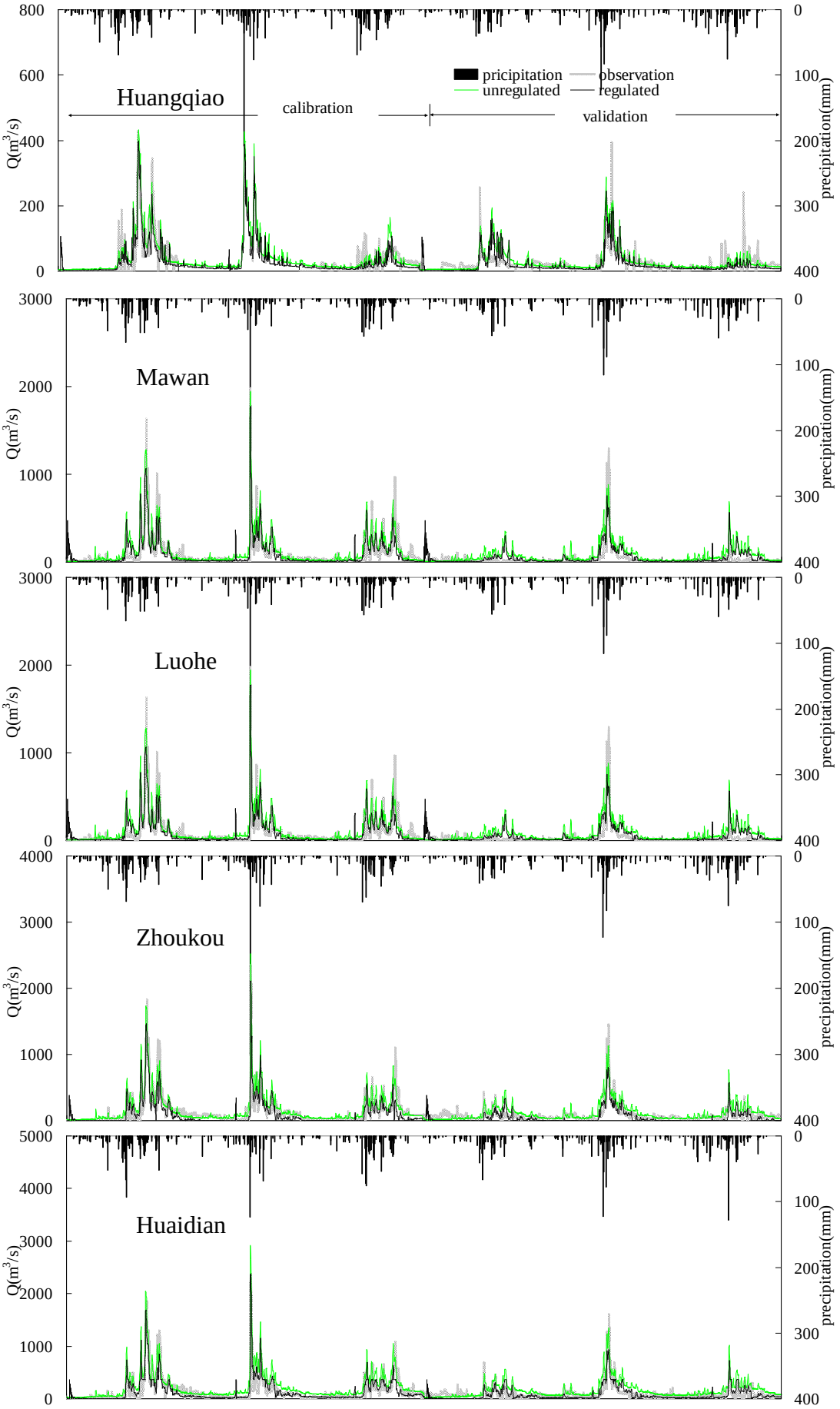


Figure 7. The digital delineation of subbasin, population, animal stock and fertilization for HE^{X}M


Huangqiao
calibration
validation
pricipitation
observation
unregulated
regulated
Q(m3/s)
precipitation(mm)
Mawan
Luohe
Zhoukou
Huaidian

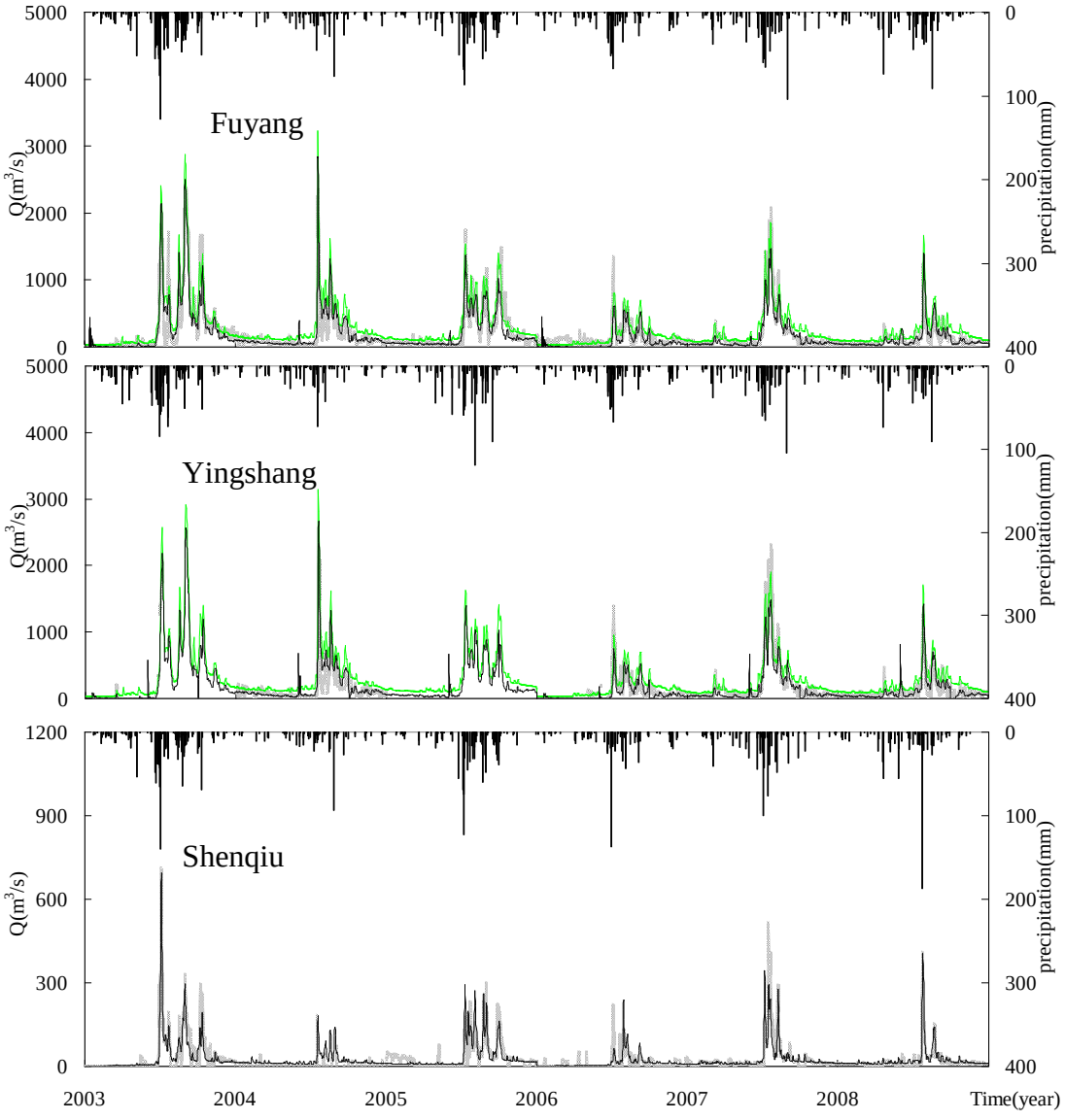
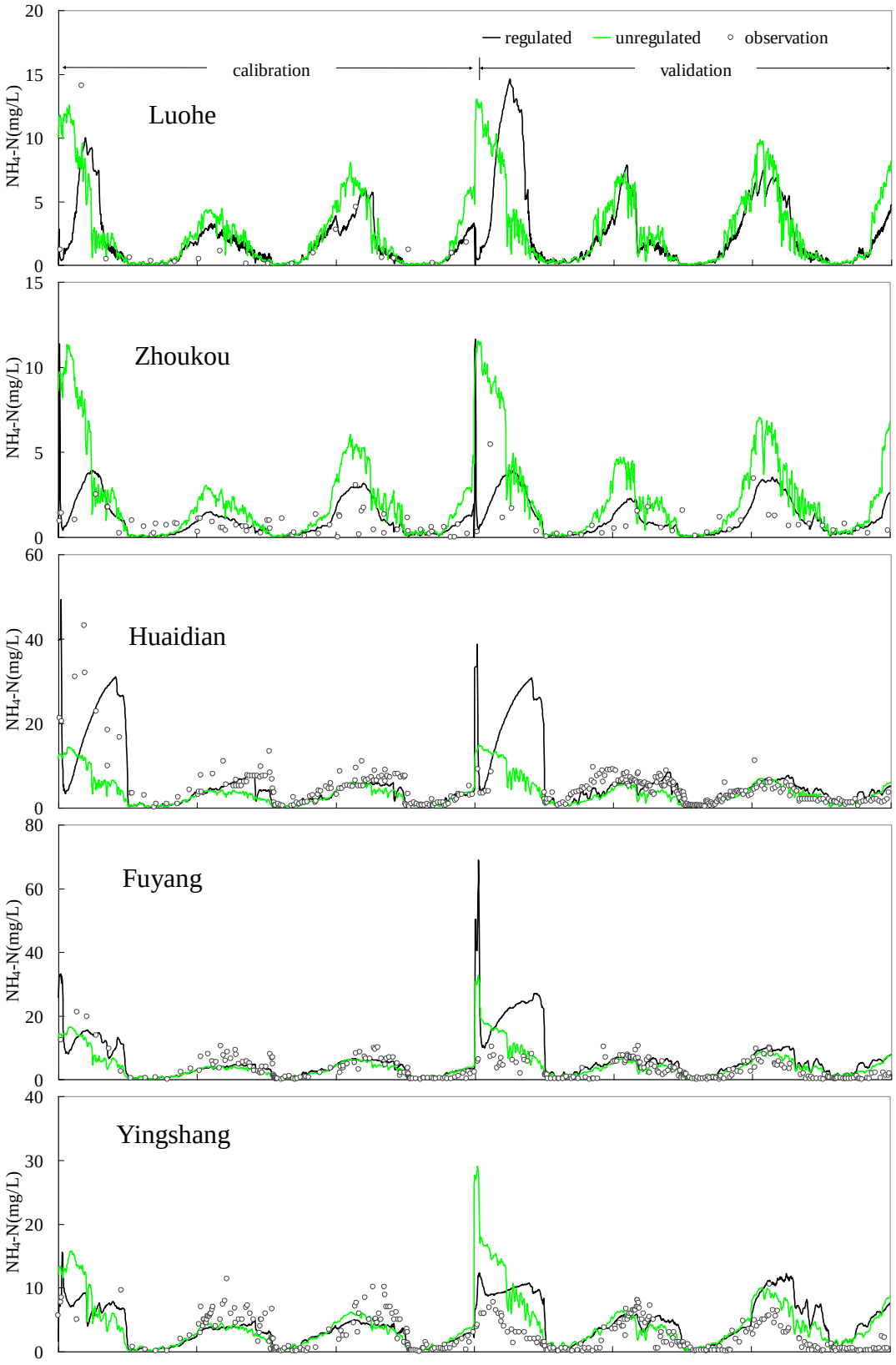


Figure 8. The daily runoff simulation at all the stations


regulated
unregulated
observation
calibration
validation
Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
NH4-N(mg/L)

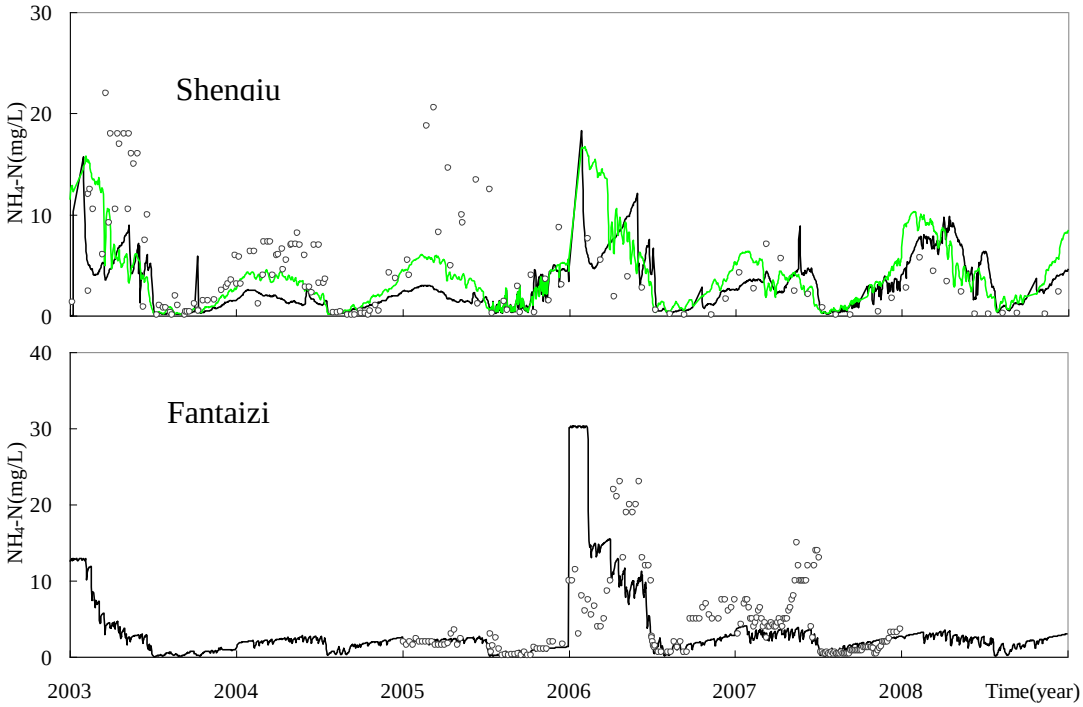
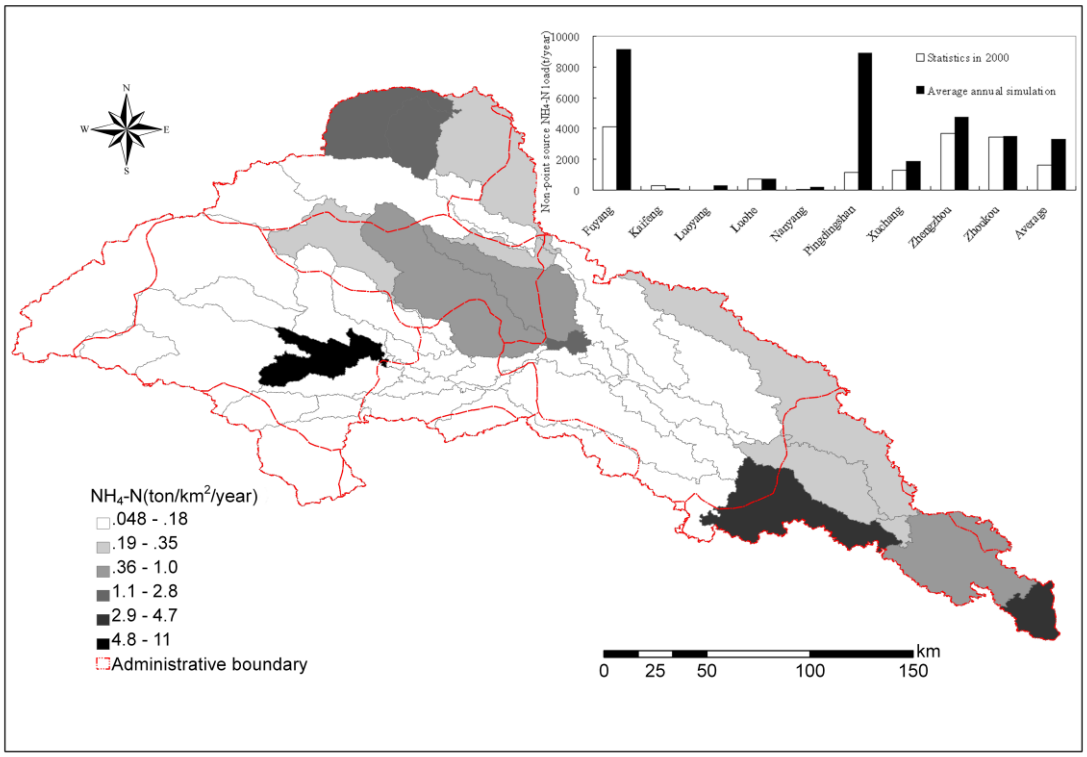
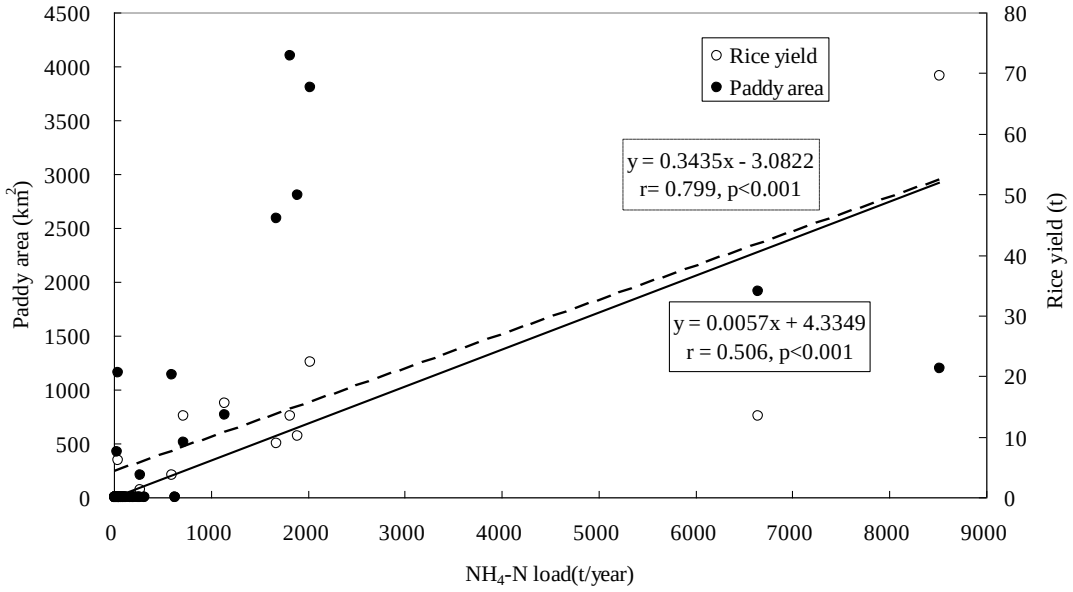


Figure 9. The simulated NH_4-N concentration variation at all the situations



1



2

3 Figure 10. The spatial pattern of nonpoint source NH_4-N load and paddy area at the

4 subbasin and regional scale in Shaying River Catchment

5

6

7

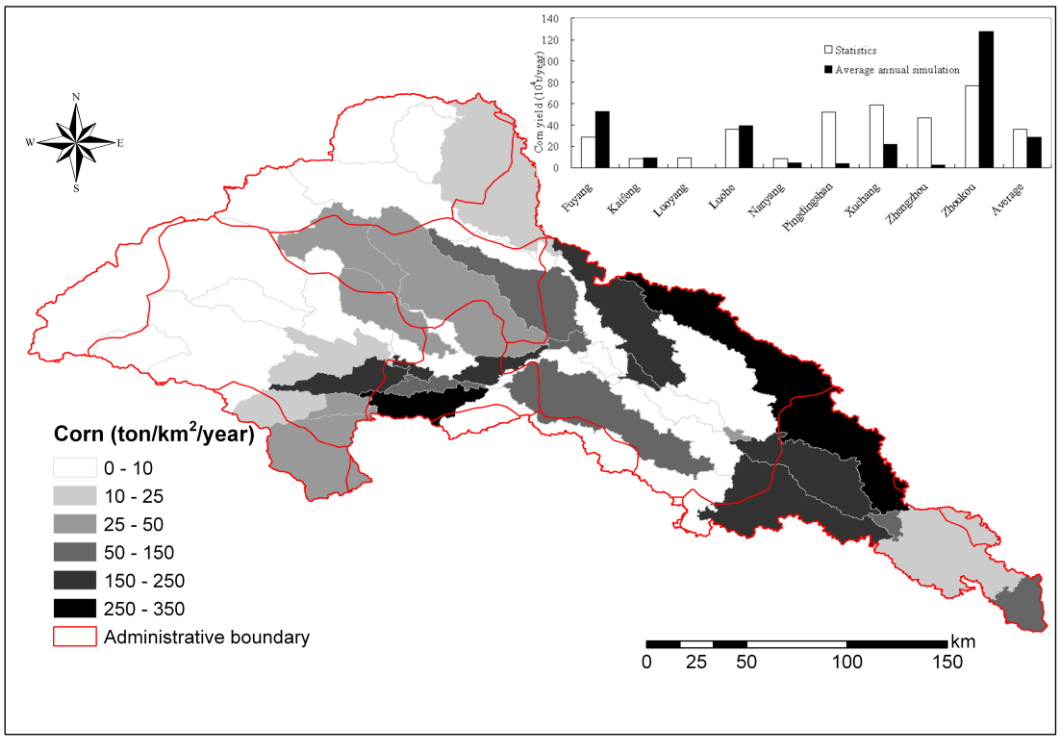


Figure 11. The spatial pattern of corn yield at the subbasin and regional scale in Shaying River Catchment

Supplementary material

1. Soil P cycle simulation (Neitsch, *et al.*, 2002)

Mineralization: The mineralized P is added to solution P pool. The amount of active and stable organic P is calculated as

$$\begin{cases} orgP_{act} = orgP_{hum} \cdot orgN_{act} / (orgN_{act} + orgN_{sta}) \\ orgP_{sta} = orgP_{hum} \cdot orgN_{sta} / (orgN_{act} + orgN_{sta}) \end{cases} \quad \text{(S1)}$$

where $orgP_{act}$ and $orgP_{sta}$ are the amount of P in active organic pool and stable organic pool, respectively (kg/ha); $orgP_{hum}$ is the humic organic P in the layer (kg/ha); $orgN_{act}$ and $orgN_{sta}$ are the amount of N in active organic pool and stable organic pool, respectively (kg/ha).

The mineralized rate of humus active organic P pool (*RHP*) is calculated

$$RHP = 1.4 \cdot \beta_{\min} \cdot (\gamma_{tmp} \cdot \gamma_{SW})^{1/2} \quad \text{(S2)}$$

where $\beta_{\min}$ is the rate coefficient for mineralization of humus active organic nutrients; γ_{tmp} and γ_{SW} are temperature factor and soil water factor.

The mineralized of the residue fresh organic P pool (*RRP*) is calculated as

$$\begin{cases} RRP = 0.8 \cdot \delta_{ntr} \\ \delta_{ntr} = \beta_{rsd} \cdot \gamma_{ntr} \cdot (\gamma_{tmp} \cdot \gamma_{SW})^{1/2} \end{cases} \quad \text{(S3)}$$

where δ_{ntr} and β_{rsd} are the residue decay rate and the mineralization coefficient of residue fresh organic nutrients. γ_{ntr} is the nutrient cycling residue composition factor.

Decomposition: The decomposition rate of the residue fresh organic P pool (*DRP*) is

$$DRP = 0.2 \cdot \delta_{ntr} \quad \text{(S4)}$$

Sorption: The P movement between soluble and active mineral pools ($P_{sol|act}$, kg/ha) and between active and stable mineral pools ($P_{act|sta}$, kg/ha) are

$$P_{sol|act} = \begin{cases} P_{sol} - \min P_{act} \cdot pai/(1-pai) & if \quad P_{sol} > \min P_{act} \cdot pai/(1-pai) \\ 0.1 \cdot [P_{sol} - \min P_{act} \cdot pai/(1-pai)] & if \quad P_{sol} < \min P_{act} \cdot pai/(1-pai) \end{cases} \quad \text{(S5)}$$

$$P_{act|sta} = \begin{cases} 0.0006 \cdot (4 \cdot \min P_{act} - \min P_{sta}) & if \quad \min P_{sta} < 4 \cdot \min P_{act} \\ 0.00006 \cdot \beta_{eqP} \cdot (4 \cdot \min P_{act} - \min P_{sta}) & if \quad \min P_{sta} > 4 \cdot \min P_{act} \end{cases} \quad \text{(S6)}$$

where P_{sol}, $minP_{act}$ and $minP_{sta}$ are soluble, mineral active and stable P, respectively (kg/ha); *pai* is P availability index.

2. Crop growth module

2.1 Crop yield (Williams *et al.*,1989)

The crop growth process depends on the accumulation of daily heat (Sharpley and Williams, 1990). The accumulated heat (*HU*, °C) during a day and heat unit index (*HUI*) is calculated as:

$$\begin{cases} HU_K = (T_{mx,K} + T_{mn,K})/2 - T_{b,j} \\ HUI_i = \sum_{K=1}^{i} HU_K \Big/ PHU_j \end{cases} \quad \text{(S7)}$$

where T_{mx} and T_{mn} are maximum and minimum daily temperature (°C), respectively; T_b is the base temperature of a certain crop (°C). *PHU* is potential heat unit required for crop maturity (°C). The range of *HUI* is from 0.0 at the seeding time to 1.0 at the physiological maturity.

The potential increased biomass for a day is estimated as follow:

$$\begin{aligned} \Delta B_{p,i} &= 0.001 \cdot BE_i \cdot PAR_i \cdot [1 + \Delta HRLT_i]^3 \\ &= 0.0005 \cdot BE_i \cdot RA_i \cdot [1 - \exp(-0.65 \cdot LAI)] \cdot [1 + \Delta HRLT_i]^3 \end{aligned} \quad \text{(S8)}$$

where ΔB_p is daily potential increased biomass (t/ha); *BE* is crop parameter for converting energy to biomass (kg·ha·m^2/MJ); *HRLT* and $\Delta HRLT$ are length of a day (hr) and its variation (hr/d); *PAR* is intercepted photosynthetic active radiation (MJ/m^2). *RA* is solar radiation (MJ/m^2) and *LAI* is leaf area index (m^2/m^2), which is a function of heat units, crop stress, and crop development stages.

From emergence to the start of leaf decline, *LAI* is estimated with the equation:

$$\begin{aligned} LAI_i &= LAI_{i-1} + \Delta LAI \\ &= LAI_{i-1} + (\Delta HUF)(LAI_{mx})(1 - \exp(5 \cdot (LAI_{i-1} - LAI_{mx}))) \cdot \sqrt{REG_i} \end{aligned} \quad \text{(S9)}$$

From the start of leaf decline to the end of the growing season,

$$LAI_i = LAI_0 \cdot \left(1 - HUI_i / 1 - HUI_0\right)^{ad_j} \quad \text{(S10)}$$

where *HUF* is heat unit factor. *REG* is minimum crop stress factor. *Ad* is a parameter controlled *LAI* decline rate for crop *j* and HUI_0 is *HUI* value when *LAI* begins to decline.

But the biomass growth is constrained by water, temperature, nutrient, and aeration.

$$\Delta B = \Delta B_p \cdot REG = \Delta B_p \cdot \min(WS, TS, SN, SP, AS) \quad \text{(S11)}$$

where *REG* is the crop growth regulating factor.

The water stress: $WS_i = \sum_{l=1}^{M} u_{i,l} \Big/ E_{P,i}$ (S12)

The temperature stress: $TS_i = \sin[\pi \cdot (T_{g,i} - T_{b,j}) / 2(T_{o,j} - T_{b,j})] \quad 0 \le TS_i \le 1$ (S13)

The nitrogen stress:
$$\begin{cases} SN_{S,i} = 2[1 - \sum_{K=1}^{i} UN_K \Big/ (c_{NB,i} \cdot B_i)] \\ SN_i = 1 - SN_{S,i} / [SN_{S,i} + \exp(3.39 - 10.93 SN_{S,i})] \end{cases} \quad \text{(S14)}$$

The phosphorus stress:
$$\begin{cases} SP_{S,i} = 2[1 - \sum_{K=1}^{i} UP_K \Big/ (c_{NP,i} \cdot B_i)] \\ SP_i = 1 - SP_{S,i} / [SP_{S,i} + \exp(3.39 - 10.93 SP_{S,i})] \end{cases} \quad \text{(S15)}$$

The aeration stress:
$$\begin{cases} SAT = SW1 / PO1 - CAF_j \\ AS_{S,i} = 1 - SAT / [SAT + \exp(-1.291 - 56.1 \cdot SAT)] \quad SAT > 0.0 \end{cases} \quad \text{(S16)}$$

where T_g and T_0 are average daily soil surface temperature and the optimal temperature (^{o}C) for crop *j*, respectively; *SAT* is saturation factor; *SW1* and *PO1* are water content and porosity of the top 1m of soil (mm) , respectively; *CAF* is critical aeration factor for crop *j*; *AS* is aeration stress factor.

The crop yield is estimated using the harvest index, viz.:

$$YLD_j = HI_j \cdot B_{AG} \quad \text{(S17)}$$

where *YLD* is total amount yield harvested from the field (t/ha), and *HI* is harvest index; *BAG* is the above-ground biomass (t/ha). For non-stressed conditions, harvest index increases nonlinearly from zero at seedling to *HI* at maturity. Affected by water

stress, the harvest index is calculated as following

$$HIA_i = HIA_{i-1} - HI_j \cdot WSYF_j \cdot FHU_i \cdot (0.9 - WS_i) / [1 + WSYF_j \cdot FHU_i \cdot (0.9 - WS_i)] \quad \text{(S12)}$$

where HI_j is normal harvest index of crop j; HIA is adjusted harvest index; $WSYF_j$ is sensitivity parameter of harvest index to draught for crop j; FHU is a function of crop growth stage. The crop growth stage function is calculated as

$$FHU_i = \begin{cases} \sin[\pi \cdot (HUI_i - 0.3)/0.6] & 0.3 \le HUI_i \le 0.90 \\ 0. & HUI_i < 0.3, HUI_i > 0.9 \end{cases} \quad \text{(S18)}$$

2.2 Water use

The potential water use from surface soil to any root depth is calculated as：

$$U_{p,i} = E_{p,i} \cdot [1 - \exp(-\Lambda \cdot Z/RZ)]/[1 - \exp(-\Lambda)] \quad \text{(S19)}$$

The potential water use ($U_{p,l}$, mm/day) in layer l is calculated by taking the difference between $U_{p,i}$ values at the layer boundaries, viz.,

$$U_{p,l} = E_{p,i} \cdot [\exp(-\Lambda \cdot Z_{l-1}/RZ) - \exp(-\Lambda \cdot Z_l/RZ))]/[1 - \exp(-\Lambda)] \quad \text{(S20)}$$

where UP is the total water used to depth Z m on day i (mm); RZ is the root zone depth (m); Λ is a water use distribution parameter.

Restricted by soil water content, the potential water use (U_l, mm/day) in layer l is calculated with the following equations when soil water content is less than 25% of plant available soil water (Jones and Kiniry, 1986).

$$U_l = \begin{cases} U_{p,l} \cdot \exp[20 \cdot (SW_{l,i} - WP_l)/(FC_l - WP_l) - 1] & if\ SW_{l,i} < (FC_l - WP)_l/4 + WP_l \\ U_{p,l} & if\ SW_l \ge (FC_l - WP_l)/4 + WP_l \end{cases} \quad \text{(S21)}$$

2.3 Nutrient uptake

The daily crop nutrient uptake (N and P) is the difference between crop nutrient demand and ideal nutrient content for day i.

$$\begin{cases} UND_i = c_{NB,i} \cdot B_i - \sum_{K=1}^{i} UN_K \\ UPD_i = c_{PB,i} \cdot B_i - \sum_{K=1}^{i} UP_K \end{cases} \quad \text{(S22)}$$

where *UND* and *UNP* are *N* and *P* uptake amount (kg/ha); *UN* and *UP* are the actual uptake of *N* and *P* (kg/ha); c_{NB} and c_{NP} are the optimal *N* and *P* concentration of the crop (kg/t); *B* is the accumulated biomass for day *i* (t/ha).

The soluble N (NO_3-N and NH_4-N) mass flow to the roots is used to distribute potential *N* uptake among soil layers.

$$\begin{cases} UN_{l,i} = u_{l,i} \cdot (WN_l / SW_l)_i \\ UNS_i = \sum_{K=1}^{M} UN_{l,i} \end{cases} \quad \text{(S23)}$$

where *WN* is NO_3^-N or NH_4-N amount in soil (kg/ha). The total N available for uptake by mass flow *UNS* is estimated by summing *UN* of all layers.

The total *P* available for uptake is calculated using the equation

$$\begin{cases} UPS_i = 1.50 \cdot UPD_i \cdot \sum_{l=1}^{M} LF_{u,l} \cdot (RW_l / RWT_i) \\ LF_{u,l} = 0.1 + 0.9 \cdot c_{LP,l} / [c_{LP,l} + 117 \cdot \exp(-0.283 \cdot c_{LP,l})] \end{cases} \quad \text{(S24)}$$

where *UPS* is the amount of *P* supplied by soil (kg/ha); *RW* and *RWT* are the root weight in layer *l* and in total (kg/ha); LF_u is the labile *P* factor for uptake (g/t).

A portion of uptake *N* will be fixed by legumes, viz.,

$$\begin{cases} WFX_i = FXR_i \cdot UND_i & WFX \le 6.0 \\ FXR = \min(1.0, FXW, FXN) \cdot FXG \end{cases} \quad \text{(S25)}$$

where *FXG* is the growth stage factor; *FXW* and *FXN* are the factors of soil water and *NO3-N*, respectively. All of these factors are calculated using the follow equations.

$$FXG_i = \begin{cases} 0.0 & HUI_i \le 0.15, HUI_i \ge 0.75 \\ 6.67 HUI_i - 1.0 & 0.15 < HUI_i \le 0.3 \\ 1.0 & 0.3 < HUI_i \le 0.55 \\ 3.75 - 5.0 HUI_i & 0.55 < HUI_i < 0.75 \end{cases} \quad \text{(S26)}$$

$$FXW_i = (SW_{0.3,i} - WP_{0.3}) / 0.85 \cdot (FC_{0.3} - WP_{0.3}) \quad SW_{0.3} < 0.85(FC_{0.3} - WP_{0.3}) + WP_{0.3} \quad \text{(S27)}$$

$$FXN_i = \begin{cases} 0.0 & WNO_3 > 300 kg \cdot ha^{-1} \cdot m^{-1} \\ 1.5 - 0.005 \cdot WNO_3 / RD & 100 < WNO_3 \le 300 \\ 1.0 & WNO_3 \le 100 \end{cases} \quad \text{(S28)}$$

where $SW_{0.3}$, $WP_{0.3}$ and $FC_{0.3}$ are the water contents in the top 0.3 m soil, at wilting point and field capacity (mm), respectively.

3. Soil erosion module (Onstad and Foster, 1975)

The soil erosion by precipitation is estimated using the improved USLE equation (Onstad and Foster, 1975), viz.,

$$Y = \begin{cases} (0.646EI + 0.45Q \cdot q_p^{0.333}) \cdot K \cdot CE \cdot PE \cdot LS & Q > 0. \\ 0 & Q \le 0. \end{cases} \quad \text{(S29)}$$

where Y is the sediment yield (t/ha); Q is runoff volume (mm); q_p is peak runoff rate (mm/hr); K is soil erodibility factor determined by the soil type; PE is erosion control practice factor.

LS is the factor of slope length and steepness:

$$\begin{cases} LS = (\lambda/22.1)^{\xi}(65.41S^2 + 4.56S + 0.065) \\ \xi = 0.6 \cdot [1 - \exp(-35.835S)] \end{cases} \quad \text{(S30)}$$

CE is the crop management factor:

$$CE = (0.8 - CE_{mn,j})\exp(-0.00115CV) + CE_{mn,j} \quad \text{(S31)}$$

EI is the rainfall energy factor:

$$EI = R \cdot [12.1 + 8.9 \cdot (\log r_p - 0.434) \cdot r_{0.5}]/1000 \quad \text{(S32)}$$

where S is land surface slope (m/m) and λ is slope length (m); ξ is a parameter dependent upon slop; $CE_{mn,j}$ is the minimum crop management factor of crop j; CV is soil cover (above ground biomass and residue) (kg/ha). R is daily rainfall amount (mm) and r_p, $r_{0.5}$ is the peak rainfall rate and maximum 0.5 h rainfall intensity (mm/hr). The value of r_p is obtained according to the exponential rainfall distribution.

4. Matter migration module

4.1 Nutrient loss in urban and rural area

The nutrient loss in urban area takes place along the surface runoff and is estimated

using the export coefficient model (Johnes,1996) .

$$V_{N\_urban} = 100 \cdot c_{N\_urban} \cdot Area_{urban} \quad \text{(S33)}$$

where V_{N_urban}, c_{N_urban} and $Area_{urban}$ are the amount of nutrient loss in urban area (kg); the export coefficient (kg/ha/year) and urban area (km^2), respectively.

The living and livestock farming in the rural area is also one of important source of nutrient. The total loss is estimated using the following equations.

$$\begin{cases} V_{N\_living} = c_{N\_living} \cdot Pop_{rural} \\ V_{N\_livestock} = c_{N\_livestock} \cdot Pop_{stock} \end{cases} \quad \text{(S34)}$$

where V_{N_living} and $V_{N\_livestock}$ are the amount of nutrient loss from living and livestock farming in the rural area, respectively (kg/year). c_{N_living} and $c_{N\_livestock}$ are the export coefficient of living (kg/day/person) and livestock (kg/day/animal), respectively; Pop_{rural} and Pop_{stock} are the population and the animal stock, respectively.

4.2 Soluble nutrient migration

The loss of soluble nutrient is considered to happen in both upper and lower layer of soil. The loss weight of NO_3^-N, NH_4-N and soluble P are calculated using the equation, respectively

$$\begin{cases} V_{N\_up} = W_{N\_up} \cdot [1 - \exp(-\dfrac{R_s + R_{ss}}{UL})] \\ V_{N\_low} = W_{N\_low} \cdot [1 - \exp(-\dfrac{R_g}{UL})] \end{cases} \quad \text{(S35)}$$

where W_{N_up} and W_{N_low} are the soluble nutrient weight in the upper and lower soil layer, respectively (kg/ha); *UL* is maximum soil water content (mm); V_{N_up} and V_{N_low} is soluble nutrient loss in the upper and lower soil layer, respectively (kg/ha).

4.3 Insoluble nutrient migration(Neitsch *et al.*, 2011)

The amount of organic *N* or *P* migrated with the sediment is estimated using the equation

$$Y_{ON} = 0.001 \cdot Y \cdot c_{ON} \cdot ER \quad \text{(S36)}$$

where Y_{ON} is loss of organic N or P (kg/ha); c_{ON} is organic N or P concentration in the soil layer (g/m^3); ER is enrich ratio.

5. Water quality module

The basic equation of in-stream water quality module (Brown and Barnwell 1987) is

$$dC/dt = -(K_d + K_{set}) \cdot C + \sum S_{out} \quad \text{(S37)}$$

where C is the pollutant concentration (mg/L); K_d and K_{set} are degradation and settling coefficient of pollutant (day^{-1}), respectively; and $\sum S_{out}$ is the external source items (mg/L/day).

The equation of water quality module of water impounding is as follow.

$$\begin{cases} dh/dt = [Q_{in} - Q_{out}]/A + P - E \\ dC_L/dt = [C_{in}Q_{in} - C_L Q_{out}]/Ah - K_{set}C_L - K_d C_L + K_{scu}C_s \cdot d/h \\ dC_s/dt = h/d \cdot K_{set}C_L - K_{scu}C_s - K_{bur}C_s \end{cases} \quad \text{(S38)}$$

where h and d are water and sediment depth, respectively (m); Q_{in} and Q_{out} are inflow and outflow, respectively (m^3/s); C_{in} and C_{out} are pollutant concentration into and out of the water body (mg/L); P and E are precipitation and evapotranspiration (m/s); C_L and C_s are constituent concentration in the water body and the sediment (mg/L); K_{scu} and K_{bur} are resuspension and decay coefficient of pollutant in the sediment (day^{-1}), respectively; A is water surface area (km^2).