Dear authors,

The three reviewers were in principle happy with the revisions in your manuscript as most of their comments were adequately addressed. Two reviewers pointed out some minor points which are given in detail below and which i would encourage you to fully address before the manuscript can be accepted for publication:

Response: Thank you very much for your comments. We revised the manuscript carefully by following all the reviewers' comments. All the changes were highlighted in blue color.

Reviewer #2:

The abstract has a good balance, with an introduction, aims, methods, results and discussion/conclusion. The results section in the abstract is perhaps too long. It is confusing as to why a present tense is used in the abstract. I would prefer a past tense to describe results.

Response: Thanks very much for your comments and suggestions. We tried our best to shorten the results in the abstract and some redundant words were removed (See P2 L3-12).

We also changed tenses in the statements for the results. Indeed, both present and past tenses are used to describe results in scientific papers. For consistency, we also used the past tense to present the results in the revision (See P2 L3-11; P4 L25-P5 L18; P14 L23-P22 L28; P23 L11-12; P23 L25-29; P24 L6-P25 L7).

'Earth' should be spelt with a capital.

Response: Thanks very much for your careful reviewing. It was changed accordingly (see P2 L28).

The manuscript mixes the use of a serial comma with non-use. A consistent usage or non-usage must be implemented. For example, at line 1 page 3.

Response: Thanks very much for your careful reviewing and sorry for the carelessness. The redundant commas were deleted in the revision (P1 L23; P4 L7, L20; P5 L6; P8 L1, L9, L25, L31; P9 L4, L7, L29; P10 L27; P11 L5, L12; P13 L19, L25; P14 L1, L2; P15 L6).

Line 9 page 3 : 'Richard's'

Response: Thanks very much for your comments. "Richards" is the surname of a researcher, not "Richard".

Inconsistent citation formats at:
page 2 line 32
page 3 line 26
page 3 line 32

page 4 line 19
page 4 line 23

Response: Thanks very much for your careful reviewing. They were revised accordingly (P2 L32; P3 L16, L26, L33; P4 L25; P8 L16).

The use of a present tense here is also unusual and I don't feel comfortable with it as a reader.

Response: Thanks very much for your comments. We changed the present tense to a past tense when we described the objectives and results in the revision (See P2 L3-11; P4 L25-P5 L18; P14 L23-P22 L28; P23 L11-12; P23 L25-29; P24 L6-P25 L7).

Serial commas at:
Page 7 line 29
Page 8 line 6
Page 8 line 22
Page 9 line 4

Response: Thanks very much for your careful reviewing. The extra commas were deleted in the revision (P1 L23; P4 L7, L20; P5 L6; P8 L1, L9, L25, L31; P9 L4, L7, L29; P10 L27; P11 L5, L12; P13 L19, L25; P14 L1, L2; P15 L6).

Inconsistent citation format and serial comma at page 8 line 13

Response: Thanks very much for your careful reviewing. The inconsistent citation format and serial comma were revised accordingly (P2 L32; P3 L16, L26, L33; P4 L25; P8 L16).

n-type dash to indicate a range: page 8 line 15

Response: Thanks very much for your careful reviewing. N-type dash was revised in the manuscript (See P8 L18; P14 L28; P15 L1, L4-5, L7; P18 L18, L31-32; P19 L1, L25; P20 L31-32; and all the references from P32 L1 to P38 L27).

'Equifinality' is mentioned on page 11 line 15 but there is no attempt to explain what the term means or even provide a citation

Response: Thanks very much for your comments. A reference about "equifinality" was added in the revision (See P11 L16).

Forests, water bodies etc. are referred to as land use, but are these not more representative of land cover?

Response: Thanks very much for your comments. Yes, water bodies are land cover. In order to shorten the description, we did not tend to distinguish land use and land cover and changed "land-use" to "land-use/cover" by following popular scientific papers (See P12 L13, P13 L8,L9,L14,L15,L18, L22, L23, L26, L27; P14 L18, L19; P15 L26; P16 L5; P20 L22; P21 L25).

Page 7 line 2-3: how is surface runoff different from overland flow?

Response: Thanks very much for your comments. The overland flow is the total runoff yield including the surface runoff, interflow and baseflow. It was clarified in P7 L6 and Figures 2 and 4.

Figure 4B: for the overland water quality module, are only two broad land use types considered, namely urban and rural areas? I would surely make sense to consider a greater range of land use types?

Response: Thanks very much for your comments. We were very sorry that the other land-use/cover types were missed. This figure was redrawn (See Fig.4).

The WQM does not mention algal growth. Does the model simulate this and the effect on nutrients?

Response: Thanks very much for your comments. We selected the QUAL-2E model and mass balance model to describe the water quality processes in the in-stream and water impounding, respectively. Both of these model considered the algal growth. We were very sorry this figure was confusing and we redraw this figure (See Fig.4).

Not enough information is provided on the WQM in terms of simulating water quality in reservoirs: stratification? Water quality sink?

Response: Thanks very much for your comments. The mass balance model was adopted to describe the water quality processes in the water impounding (reservoir, lake). We were very sorry this figure was confusing and we redraw this figure. The detailed descriptions of water quality processes were showed at the top right-hand corner.

Water quality sink process was considered in the water quality modules of water impounding (P10 L16 and equation S48 in the supplementary material). However, we did not considered the stratification in this study. The main reasons were that the high resolution bathymetric data of individual dams or lakes were needed for stratification but frequently not available, and the extended model focused on the processes of water and nutrients at the large basin scale. This issue was mentioned in the model limitations in the discussion section (See P24 L9-11).

Page 10 line 11 'spatial consistent'. Please reformulate this sentence into readable English.

Response: Thanks very much for your careful reviewing. It was revised to "The model is solved at the sub-basin scale rather than at the fine grid scale in order to maintain spatial consistency with the hydrological cycle module" (P10 L10-12).

Page 15. In terms of ranges and using hyphens, use the n-type dash.

Response: Thanks very much for your careful reviewing. N-type dash was revised in the manuscript (See P8 L18; P14 L28; P15 L1, L4-5, L7; P18 L18, L31-32; P19 L1, L25; P20 L31-32; and all the references from P32 L1 to P38 L27).

Page 15 line 3 and 4: incorrect use of brackets. Reformulate the sentence.

Response: Thanks very much for your careful reviewing. The sentence was revised to "The average annual stocks include 8.30 million of big animals (cattle, pigs and sheep) and 178.42 million of poultries" (See P15 L2-3).

Page 15 line 30 to 32: use a thousands separator (,) as you have in other parts of the document.

Response: Thanks very much for your careful reviewing. There were revised accordingly in the revision (See P14 L30; P16 L6-7).

Page 11 line 24: You mention the 'NS' to evaluate model performance. Please state what this is. I assume it is the Nash Sutcliffe efficiency? Please provide the citation.

Response: Thanks very much for your comments. Yes, "NS" was the "Nash-Sutcliffe efficiency". It was revised and a reference was also provided in the revision (See P11 L24; P31 L9).

Page 16 line 25: Using the log (NS) is less sensitive to extreme values.

Response: Thanks very much for your comments. We agreed with you that the log (NS) was less sensitive to extreme values. However, it will cover the problem on model fitting for extreme values. We did not tend to extend the discussion in this direction because the paper is already very long.

I think the manuscript would benefit from some discussion on the strategy of more complex models (as taken here) as opposed to simpler models that attempt to simulate the most important processes explaining the majority of water quality variation (requisite simplicity), and the corresponding benefits/disadvantages of each broad type of model in regards to the required observed data, ease of use, and equifinality. Perhaps the model described here needs to be places in a more global context: could it be used in other countries? Would it be of benefit in developing countries where data and resources are a limitation?

Response: Thanks very much for your summarization and suggestions. The sub-section of "comparison with other models" (4.1) was revised following the suggestions of you and Reviewer 3. More comparisons and the limitations of the proposed model were discussed in P22 L6, L8, L17-26; P25 L7-9, L16-18.

Reviewer #3:

Thank you for responding to most of my own and the other reviewers concerns. I have just a few minor points for revision.

1. Scientific purpose - The authors state that the scientific purpose is the integration of multiple bio-geo-chemical processes. It would be nice if they could turn this into a question: "Does including these extra processes in an integrated manner improve model results compared to models that are focuses only on one component?" I think they show this to some extent in their comparisons with SWAT, but could slightly reword their introduction and discussion to better demonstrate this. The discussion now is a slightly boring comparison with SWAT. Remove the long list of number comparisons and discuss whether or not including more processes in the new model allows for an improvement of the overall model results compared to SWAT. Incidentally, if you really wanted to demonstrate the added value of integrated components, you could allow interaction between the model modules in calibration, e.g. when calibrating water quality you may find issues to improve water quantity (such as the need for more surface runoff to increase particulate phosphorous erosion). This is why sometimes manual calibration can actually be superior to automatic.

Response: Thanks very much for your good suggestion. A sentence quoting this question was added in P4 L12 to lead the review on SWAT model.

Following your suggestion and reviewer #2's suggestion, we revised the introduction section, the discussion section and the sub-section of "comparison with other models". More comparisons were discussed in P22 L6, L8 and L17-26. The long list of number comparisons were revised to "Compared with the results of these models, our model generally performed better on the runoff or water quality simulations. In particular, our model performed even better than SWAT at the regulated stations as more detailed dam regulation rules and soil biochemical processes were considered. For example, the average values of f_{runoff} at the monthly scale decreased from 0.32 (SWAT in Zhang *et al.*, 2012) to 0.15 (our model) at the regulated stations. The average values of $f_{NH4\text{-}N}$ decreased from 0.47 (SWAT in Zhang *et al.*, 2012) to 0.27 (our model). Moreover, both the Xinanjiang model and DTVGM are limited to simulate the flow series at the unregulated or less-regulated stations because they do not consider the dam regulation in their current model frameworks (Shi *et al.*, 2013; Ma *et al.*, 2014)." (See P22 L17-26.)

2. Input Data: I'm not sure the authors understood my concerns about input data: My question to them is "Does the resolution and quality of the input data support the complexity of the process descriptions in the model?" How accurate are the agricultural N and P inputs or the point sources? Does the government have good control over this in China or do farmers and industries flaunt environmental regulations? This varies considerably from country to country and therefore affects the quality of the model inputs. This was touched upon in the crop model results but should be more clearly mentioned in the discussion.

Response: Thanks very much for your comments. The detailed spatial and temporal resolution and source of data were presented in Table S2 in the supplementary material. The resolutions of GIS and weather input data were quite satisfactory for our case study. However, most data for water quality, ecology and agricultural management were at monthly or annual temporal scale. The data for economy, agricultural management and diffuse source load were collected from individual administrative regions. Both the temporal and spatial scales were bigger than the required daily scale and the spatial calculation units (sub-basin, land-use/cover and crop). The data were uniformly distributed to solve this problem as usual. Thus, the scale discrepancies would directly affect the model performance during the model calibration. To reduce such problems, we assess the model performance at the original observed scales. Nevertheless, all the collected data source could drive our model and satisfactory results were obtained. If higher-resolution data are available, better simulation performances would be achieved, particularly for water quality, diffuse source and crop yield. The detailed description of model input data were given in P 15 L16-24.

Furthermore, data quality is a general issue and the uncertainty of data accuracy is still a hot topic in the water related sciences. It affects the performance of all mathematical models, not only our model. This issue was also discussed in P25 L7-9.

3. Some minor points:

- Please check the spelling and grammar one last time - there are some small errors remaining

Response: Thanks very much for your comments. We carefully checked the manuscript again to minimize the grammar problems.

- The HYPE model IS developed for integrated water quality and hydrology (N, P and to some extent C) and also includes regulation of dams. See Strömqvist et al. 2011.

Response: Thanks very much for your comments. We agreed the HYPE is developed for integrated water quality and hydrology. The improved version S-HYPE (Strömqvist *et al.*, 2011) also considered the regulation of dams. However, we thought this model still belongs to the hydrology-based model. It was cited in P 3 L27 and P37 L12-14.

Integrated water system simulation by considering hydrological and biogeochemical processes: model development, with parameter sensitivity and autocalibration

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

[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 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@igsnrr.ac.cn) and

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

Abstract

Integrated water system modeling is a feasible approach to understanding severe water crises in the world and promoting the implementation of integrated river basin management. In this study, a classic hydrological model (the time variant gain model: TVGM) was extended to an integrated water system model by coupling multiple water-related processes in hydrology, biogeochemistry, water quality, ecology and considering the interference of human activities. A parameter analysis tool, which included sensitivity analysis, autocalibration and model performance evaluation, was developed to improve modelling efficiency. To demonstrate the model performances, the Shaying River Catchment, which is the largest, highly regulated and heavily polluted tributary of the Huai River Basin in China, was selected as the case study area. The model performances were evaluated on the key water-related components including runoff, water quality, diffuse pollution load (or nonpoint source) and crop yield. Results showed that our proposed model simulated most components reasonably well. The simulated daily runoff at most regulated and less-regulated stations matched well with the observations. The average correlation coefficient and Nash–Sutcliffe efficiency were 0.85 and 0.70, respectively. Both the simulated low and high flows at most stations were improved when the dam regulation was considered. The daily ammonium-nitrogen (NH_4-N) concentration was also well captured with the average correlation coefficient of 0.67. Furthermore, the diffuse source load of NH_4-N and the corn yield were reasonably simulated at the administrative region scale. This integrated water system model is expected to improve the simulation performances with extension to more model functionalities, and to provide a scientific basis for the implementation in integrated river basin managements.

1. Introduction

Severe water crises are global issues that have emerged as a consequence of the rapid development of social economy, and include flooding, water shortages, water pollution and ecological degradation. These crises have hindered the equitable development of regions by compromising the sustainability of vital water resources and ecosystems. It is impossible to address these crises within a single scientific discipline (e.g., hydrology, hydraulics, water quality or aquatic ecology) because of the complicated interactions among physical, chemical and ecological components of an aquatic ecosystem (Kindler, 2000; Paola *et al.*, 2006). The paradigm of integrated river basin management may be a sensible solution at basin scale by focusing on the coordinated management of water resources in term of social economy, water quality and ecosystems. Integrated water system models have been popular since last decade due to the rapid development of water-related sciences, computer science, Earth observation technologies and the availability of open data.

Hydrological cycle has been known as a critical linkage among other water-related processes (e.g., physical, biogeochemical and ecological processes) and energy fluxes at the basin scale (Burt and Pinay, 2005). For examples, physiological and ecological processes of vegetation affect evapotranspiration, soil moisture distribution and nutrient movement. In the meantime, soil moisture and nutrient constrain the vegetation growth. Overland flow is a carrier of pollutants to water bodies. Therefore, all the processes should be considered simultaneously to capture the interactions and feedbacks between individual cycles. Multidisciplinary research provides an effective way to enable breakthroughs in the integrated water system modeling by integrating the theories in water-related sciences (e.g., accumulated temperature law for phenological development, Darcy's law for groundwater flow, Saint-Venant equation for flow routing, balance equation for mass and momentum, Richards' equation for unsaturated zone, Horton theory for infiltration, Penman-Monteith equation for evapotranspiration). Abundant open data sources further support the implementation of integrated water system model, e.g., high-resolution spatial information data, chemical and isotopic data from field experiments (Singh and Woolhiser, 2002; Kirchner, 2006).

Several models have been developed since the 1980s (Di Toro *et al.*, 1983; Brown and Barnwell, 1987; Johnsson *et al.*, 1987; Hamrick, 1992; Li *et al.*, 1992; Abrahamsen and Hansen, 2000; Tattari *et al.*, 2001; Singh and Woolhiser, 2002). Owing to the complexity of the integrated water system and the scale conflicts between different processes, most existing models focus on only one or two major water-related processes, and can be categorized into three major classes. (1). Hydrological models emphasize the rainfall-runoff relationship and link with some dominating water quality and biogeochemical processes. These models generally show satisfactory performances in simulating the hydrological processes. Some widely accepted models are TOPMODEL (Beven and Kirkby, 1979), SHE (Abbott *et al.*, 1986), HSPF (Bicknell *et al.*, 1993), VIC (Liang *et al.*, 1994), ANSWERS (Bouraoui and Dillaha, 1996), HBV-N (Arheimer and Brandt, 1998 and 2000), HYPE (Lindström *et al.*, 2010) and its improved version S-HYPE (Strömqvist *et al.*, 2012). (2). Water quality models focus on the migration and transformation processes of pollutants in water bodies. These models can simulate the water quality variables at high spatial and temporal resolutions in river networks by adopting multi-dimensional dynamic equations. However, they have difficulties to simulate the overland processes of water and pollutants. Typical models include WASP (Di Toro *et al.*, 1983), QUAL2E (Brown and Barnwell, 1987) and EFDC (Hamrick, 1992). (3). Biogeochemical models have advantages in simulating the physiological and ecological processes of vegetation and the vertical movements of nutrients and water in soil layers at the field or experimental catchment scales. However, these models lack accurate hydrological features (Deng *et al.*, 2011) and are hard to simulate the movements of water, nutrients and their losses along flow pathways in the basin. Some biogeochemical models are SOILN (Johnsson *et al.*, 1987), EPIC (Sharpley and Williams, 1990), DNDC (Li *et al.*, 1992), Daisy (Abrahamsen and Hansen, 2000) and ICECREAM (Tattari *et al.*, 2001). Overall, most models usually achieve good performances on their oriented processes and only approximate the results for other processes outside of the model's focus in the integrated river basin management. An important scientific question is "does including these extra processes in an integrated manner improve model results compared to models that are focuses only on one component?"

SWAT is an integrated water system model that can simulate most water-related processes over a long period at large scales (Arnold *et al.*, 1998). However, not all water-related processes can be well captured in practice because of the inaccurate descriptions of some processes, such as daily extreme flow events (Borah and Bera, 2004), soil nitrogen and carbon (Gassman *et al.*, 2007) and regulation rules of dams or sluices in regulated basins (Zhang *et al.*, 2012). Particularly, the simulation methods of surface runoff yield in SWAT have been questioned, e.g., the general applicability of the curve number (Rallison and Miller, 1981) and the scale limitations of the Green-Ampt infiltration model (King *et al.*, 1999). Furthermore, SWAT has difficulties in accurately capturing the complicated dynamic processes of soil nitrogen and carbon by comparing with other biochemical models (Gassman *et al.*, 2007). Several modified versions have been developed, such as SWIM (Krysanova *et al.*, 1998) and SWAT-N (Pohlert *et al.*, 2006).

In this study, we tended to develop an integrated water system model based on a hydrological model. The time variant gain model (TVGM) proposed by Xia (1991) is a lumped hydrological model based on the rainfall and runoff observations from many basins with different scales all over the world. In TVGM, the rainfall-runoff relationship is considered to be nonlinear because the surface runoff coefficient varies over time and is significantly affected by antecedent soil moisture. TVGM has strong mathematical basis because this nonlinear relationship is transformed into a complex Volterra nonlinear formulation. Wang *et al.* (2002) extended TVGM to the distributed time variant gain model (DTVGM) by taking the advantages of better computing facilities and available data sources. Currently, DTVGM is performed well in many basins with different scales and climate zones to investigate the effect of human activities and climate change on runoff (Xia *et al.*, 2005; Wang *et al.*, 2009).

In the model development, we would like to produce reasonable simulations simultaneously in both hydrological and water quality processes and to include more water-related processes such as soil biogeochemistry and crop growth for better understandings of the complicated water related processes and their interactions in the real basins. Our proposed model was built by extending DTVGM through coupling the detailed interactions and linkages among hydrological, water quality, soil biogeochemical and ecological processes, as well as considering the prevalent regulations of water projects (dams and sluices) at the basin scale. In order for readers to use the proposed model easily, a parameter analysis module, which included popular objective functions, autocalibration approaches and summary statistics, was also developed. To demonstrate the model performances, we simulated several key water-related components including flow regimes, diffuse source (or nonpoint source) pools of nutrients, water quality variables in water bodies and crop yield in a highly regulated and heavily polluted catchment (Shaying River Catchment) in China.

2. Methods and material

2.1 Model framework

Our proposed model includes eight major modules, namely hydrological cycle module (HCM), soil biochemical module (SBM), crop growth module (CGM), soil erosion module (SEM), overland water quality module (OQM), water quality module of water bodies (WQM) and dam regulation module (DRM). The parameter analysis tool (PAT) is also designed for model calibration. The model structure is shown in Figure 1. More detailed descriptions of each module and its interactions with other modules are given in sub-sections 2.1.1 to 2.1.5. The main equations of each module are deferred to the appendix and supplementary materials for readers who are interested in the mathematical details.

Our model is 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. It takes full advantages of the existing models, i.e., the powerful interconnections of the hydrological models with other processes at the spatial scale, the elaborative descriptions of the ecological models on nutrient vertical movement in soil layers, and the elaborative descriptions of the water quality models on nutrient movements along river networks. First, several key components, simulated by the hydrological cycle module (HCM) (e.g., evapotranspiration, soil moisture and flow), are treated as critical linkages in all the modules (Section 2.1.1). Second, the soil biochemical processes determine the nutrient loads absorbed in the crop growth process (CGM) and migrated into water bodies as the diffuse pollution source (OQM and WQM). The accurate descriptions of soil biochemical processes are helpful in improving the simulation of diffuse source processes in responding to agricultural management (Section 2.1.2). Third, the hydrological cycle module (HCM) provides a function for describing the connections between spatial calculation units to simulate the overland and in-stream movements of water and nutrients at the basin scale (Sections 2.1.1 and 2.1.3).

2.1.1 Hydrological cycle module (HCM)

Surface runoff calculation is the core of hydrological simulation. TVGM is adopted to calculate the surface runoff yields for different land-use/cover areas, such as forest, grassland, water body, urban area, unused land, paddy land and dryland agriculture. The potential evapotranspiration is calculated using Hargreaves method (Hargreaves and Samani, 1982) because only the available daily maximum and minimum temperatures are used. The actual plant transpiration is expressed as a function of potential evapotranspiration and leaf area index, whereas soil evaporation is expressed as a function of potential evapotranspiration and surface soil residues (Neitsch *et al.*, 2011). The yields of interflow and baseflow have linear relationships with the soil moisture in the upper and lower layers, respectively (Wang *et al.*, 2009). The infiltration from the upper to lower soil layers is calculated using storage routing method (Neitsch *et al.*, 2011). The Muskingum method or kinetic wave equation is used for river flow routing.

Figure 2 shows that the shallow soil moisture from the hydrological cycle module is a major factor that connects the crop growth module (to control crop growth) and the soil biochemical module (to control the vertical migration and reaction of nutrients in the soil layers). Plant transpiration is also linked to the soil biochemical module (to drive the vertical migration of nutrients in the soil layers). The surface runoff is linked to the soil erosion module, while the overland flow (surface runoff, interflow and baseflow) is connected to the overland water quality module (to drive the movements of nutrients and sediment along flow pathways) and the water quality module of water bodies (rivers and lakes) for runoff routing. Moreover, the hydrological cycle module provides the inflows for individual dams or sluices in the dam regulation module.

2.1.2 Modules for ecological processes

The ecological processes are described in the soil biochemical module and the crop growth module. The crop growth and soil biochemical processes directly affect the soil moisture, evapotranspiration, nutrient transformation and loss from the soil layers. Therefore, our model incorporates the water cycle, nutrient cycle, crop growth and their key linkages.

2.1.2.1 Soil biochemical module (SBM)

The soil biochemical module simulates the key processes of Carbon (C), Nitrogen (N) and Phosphorus (P) dynamics in the soil layers, including decomposition, mineralization, immobilization, nitrification, denitrification, leaching and plant uptake. Different forms of N and P outputted from the soil biochemical module are connected to the crop growth module as the nutrient constraints of crop growth and to the overland water quality module as the main diffuse sources to water bodies (Figure 3a).

Soil C and N cycle. The sub-models of daily step decomposition and denitrification in DNDC (Li *et al.*, 1992) are adopted to simulate the soil biogeochemical processes of C and N at the field scale. The decomposition and other oxidation processes are the dominant microbial processes in the aerobic condition. The three conceptual organic C pools are the decomposable residue C pool, microbial biomass C pool and stable C pool. The decomposition of each C pool is treated as the first-order decay process with the individual decomposition rates constrained by the soil temperature and moisture, clay content and C: N ratio. The major simulated processes of decomposition under aerobic condition are mineralization, immobilization, ammonia (NH_3) volatilization and nitrification. The mineralization and immobilization of mineral N (NH_4^+ and NO_3^-) are determined by the flow rates of soil organic carbon (SOC) pools. NH_3 volatilization is controlled by the NH_4^+ concentration, clay content, pH, soil moisture and temperature. NH_4^+ is oxidized to NO_3^- during nitrification and nitrous oxide (N_2O) is emitted into the air during the nitrification. Denitrification occurs under the anaerobic condition, which is controlled by soil moisture, temperature, pH and dissolved SOC content. The detailed descriptions are given in Appendix B and Li *et al.* (1992).

Soil P cycle. The major processes of soil P cycle are simulated according to the study of Horst *et al.* (2001). Six P pools are considered including three organic pools (stable and active pools for plant uptake, fresh pool associated with plant residue) and three mineral pools (dissolved mineral, stable and active pools). The involved processes are the P release, mineralization and decomposition from fertilizer, manure, residue, microbial biomass, humic substances and the sorption by plant uptake (Horst *et al.*, 2001; Neitsch *et al.*, 2011).

Soil profile is divided into three layers, namely, surface (0–10 cm) and user defined upper and lower layers, all of which are consistent with the soil layers of hydrological cycle module to smoothly exchange the values through the linkages (e.g., soil moisture) among different modules.

2.1.2.2 Crop growth module (CGM)

The crop growth module is developed based on EPIC crop growth model (Hamrick, 1992). It simulates total dry matter, leaf area index, root depth and density distribution, harvest index, nutrient uptake and so on (Williams *et al.*, 1989; Sharpley and Williams, 1990). The crop respiration and photosynthesis drive the vertical movements of water and nutrients. The output of leaf area index is a main factor connecting the hydrological cycle module (to control the transpiration) and the crop residue left in the fields is a main source of organic nutrients (C, N and P) connecting to the soil biochemical module for soil biochemical processes, to the overland water quality module and to the soil erosion module as one of the five constraint factors (Figure 3b).

2.1.3 Modules for water quality processes

The water quality processes focus on the migration and transformation of water quality variables (e.g., sediment, different forms of nutrients, biochemical oxygen demand: BOD and chemical oxygen demand: COD) along the flow pathways in the land surface and river network. The main modules are the soil erosion module for the sediment yield, the overland water quality module for the migration of overland diffuse source to water bodies and the water quality module for the migration and transformation of point and diffuse pollution sources in water bodies.

2.1.3.1 Soil erosion module (SEM)

The soil erosion by precipitation is estimated using the improved USLE equation (Onstad and Foster 1975) based on runoff yields outputted from the hydrological cycle module and crop management factor outputted from the crop growth module. The soil erosion module simulates the sediment load for the overland water quality module to provide the carrier for the migration of insoluble organic matters along overland transport paths and water bodies (Figure 4a).

2.1.3.2 Overland water quality module (OQM)

This module simulates the overland losses and migration loads of diffuse source pollutants (e.g., sediment, insoluble and dissolved nutrients, BOD and COD) (Figure 4b). The main diffuse sources include the nutrient loss from the soil layers and urban areas, the farm manure from livestock in rural areas. The nutrient loss from the soil layers, as the primary diffuse source in most catchments, is determined by the overland flow and sediment yield (Williams *et al.*, 1989) and the other sources are estimated using the export coefficient method (Johnes, 1996). The overland migration processes contain the dissolved pollutant migration with overland flow and the insoluble pollutant migration with sediment. All the processes occur along the overland transport paths.

2.1.3.3 Water quality module of water bodies (WQM)

This module simulates the transformation and migration of water quality variables in different types of water bodies (in-stream and water impounding) (Figure 4c). The simulated variables include water temperature, dissolved oxygen (DO), sediment, different forms of nutrients (N and P), BOD and COD. Point pollution sources are also considered. Point sources are directly added to the surface water in the model according to their geographic positions. Common point sources are urban water treatment plants and industrial plants.

Two modules are designed for the different types of water bodies, i.e., the in-stream water quality module and the water quality module for water impounding (reservoir or lake). The enhanced stream water quality model (QUAL-2E) (Brown and Barnwell 1987), is adopted to simulate the longitudinal movement and transformation of water quality variables in the in-streams. The model is solved at the sub-basin scale rather than at the fine grid scale in order to maintain spatial consistency with the hydrological cycle module. The water quality outputs provide the water quality boundary of dams or sluices in the dam regulation module. The water quality module for water impounding assumes that water body is at the steady state and focuses on the vertical interaction of water quality processes. The main processes include water quality degradation and settlement, sediment resuspension and decay.

2.1.4 Dam regulation module (DRM)

Dams and sluices highly alter flow regimes and associated water quality processes in most river networks. Thus, the dam and sluice regulation should be considered in the water system models. The dam regulation module provides the regulated boundaries (e.g., water storage and outflow) to the hydrological cycle module for flow routing and to the water quality module of water bodies for pollutant migration.

Given that different types of dams and sluices are likely to show completely different regulation behaviors, we try to reproduce their common functionalities for either the flood control or water supply in this module. Three methods are proposed to calculate the water storage and outflow of dams or sluices, namely, the measured outflow, controlled outflow with target water storage and the relationship between outflow and water storage volume. The first method requires users to provide the measured outflow series during the simulation period. The second method simplifies the regulation rules of dams or sluices for long-term analysis based on the assumption that water is stored according to the usable water level during non-flooding season and the flood control level during flooding season, and the surplus water is discharged. This 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 based on the relationships among water level, water surface area, storage volume and outflow according to the designed dam data or long-term observed data (Zhang *et al.*, 2013) (Appendix C).

2.1.5 Parameter analysis tool (PAT)

In our model, 66 lumped and 94 distributed parameters involve the hydrological, ecological and water quality processes. The distributed parameters are divided into 37 overland parameters, 17 stream parameters and 40 parameters of water projects (only for the sub-basin with reservoir or sluice) 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 spatial calculation units share many common parameter values if their properties are the same.

Owing to a large number of parameters, it is hard to find optimal parameter values by manual tuning. Limited number of observed processes causes equifinality in model calibration (Beven, 2006). Therefore, the parameter sensitivity analysis and calibration are important steps to alleviate equifinality in the applications of highly parameterized models, particularly for integrated water system models (Mantovan and Todini, 2006; Mantovan *et al.*, 2007; McDonnell *et al.*, 2007). The PAT is designed for parameter sensitivity analysis, autocalibration and model performance evaluation (Figure 5).

To evaluate model performance, five traditionally used criteria are included in the PAT, i.e., bias (*bias*), relative error (*re*), root mean square error (*RMSE*), correlation coefficient (*r*) and Nash-Sutcliffe efficiency (*NS* defined by Nash and Sutcliffe, 1970). The detail definitions of these criteria are given in Appendix D. Furthermore, flow duration curve and cumulative distribution function are also provided for capturing multiple signatures of calibrated processes. More criteria can also be proposed by the users. The objective function(s) to calibrate the model can be formed by a single or multiple criteria or their function (such as weighted average).

The parameter analysis algorithms in the PAT include the parameter sensitivity method (Latin hypercube one factor at a time: LH-OAT) (van Griensven *et al.*, 2006), the single objective auto-optimization methods such as particle swarm optimization

(PSO) (Kennedy, 2010), genetic algorithm (GA) (Goldberg, 1989) and shuffled complex evolution (SCE-UA) (Duan *et al.*, 1994), as well as the multi-objective auto-optimization methods such as weighted sum method and nondominated sorting genetic algorithm II (NSGA-II) (Deb *et al.*, 2002). The method can be selected on the basis of the specific requirements of users.

In order to obtain the optimal parameter values, the following treatments are adopted in the PAT. First, the prior ranges of all the parameter values or their prior distributions (i.e., uniform or normal) are preset by referring the literatures or similar basins. The constraints on parameters are also considered in both parameter sensitive analysis and autocalibration. In the hydrological cycle module, the constraints on soil moisture parameters are "W_m (minimum moisture) < W_w (moisture at permanent wilting point) < W_{fc} (field capacity) < W_{sat} (saturated moisture capacity)". The basic surface runoff coefficient (g_1) for different land-use/covers are set in ascending order (water body, paddy land, urban area, forest, dryland agriculture, unused land and grassland). The interflow yield coefficient (Kss) is greater than the baseflow coefficient (K_{bs}). In the water quality module of water bodies, the settling rates of water quality variables (K_{set}) in the water impounding are greater than the resuspension rates (K_{scu}) and the settling rates (R_{set}) in channels. Second, the sensitive parameters are determined to reduce the parameter dimensions by sensitivity analysis. Third, the selected sensitive parameters are calibrated by auto-optimization method, while the insensitive parameters remain as their default values which are given by referring the literatures, or other models (e.g., SWAT, EPIC and DNDC) in the same/ similar basins.

The PAT connects with other modules through the parameter values which are used to simulate the processes of other modules and evaluate the objective functions in sensitivity analysis and autocalibration. Depending on the algorithm used, the parameter values are (randomly) sampled from the multi-dimensional parameter spaces to drive our model and the objective function value of each parameter set is then obtained. For the parameter sensitivity analysis, the sensitivity index of each parameter set is evaluated by comparing the variation of the objective function value along with the change of parameter value. For the parameter autocalibration, the good parameter sets are kept or updated by the auto-optimization method until the convergence or the maximum number of iterations is achieved.

2.2 Model operation

2.2.1 Multi-scale solution

The spatial heterogeneities of basin attributes and the different time scales used in individual processes cause inconsistent spatial and temporal scales in model integration (Sivapalan and Kalma, 1995; Singh and Woolhiser, 2002). For the spatial scale, three levels of spatial calculation units are designed, namely, sub-basin, land-use/cover and crop from largest to smallest. These units are defined as the minimum polygons with similar hydrological properties, land-use/covers and agriculture crop cultivation patterns, respectively. The sub-basins are defined on the basis of digital elevation model (DEM), the positions of gauges and water projects, and are used in the hydrological cycle module (e.g., flow routing in both land and in-stream), overland water quality module, water quality module of water bodies and dam regulation module. Seven specific land-use/cover units of each sub-basin are partitioned by the land-use/cover classification (i.e., forest, grassland, water, urban, unused land, paddy land and dryland agriculture) and are used in the hydrological cycle module (e.g., water yield, infiltration, interception and evapotranspiration) and the soil erosion module. Moreover, several specific land-use/cover units (paddy land, dryland agriculture, forest and grassland), where agricultural activities usually occur, are divided further into the crop units for the detailed analysis of the impact of agricultural management on water and nutrient cycles. In the current version of our model, these four land-use/cover units are divided into 10 specific categories of crop units as fallow for all these land-use/cover units, grass for grassland unit, fruit tree and non-economic tree for forest unit, early rice and late rice for paddy unit, spring wheat, winter wheat, corn and mixed dry crop for dryland agriculture unit. The crop unit of a specific land-use/cover pattern varies depending on crop cultivation structure and timing. The related modules are the soil biochemical module and the crop growth module. All of the outputs of the crop unit are summarized at the land-use/cover scale or sub-basin scale based on the area percentages in different crop units.

For the temporal scale, it is practical to use a daily time-step as this is consistent with the underlying rainfall-runoff module and the data availability. The sub-daily scale may improve the performance in some modules (e.g., SEM and WQM). However, most observations (e.g., climate data sets, soil nutrient availability and water quality concentrations) are at the daily scale, leading to potential uncertainties or instabilities to disaggregate the observations into a sub-daily scale. Linear or nonlinear aggregation functions are used to transform different time scales to daily scale (Vinogradov *et al.*, 2011), such as exponential functions for flow infiltration and overland flow routing processes in the hydrological cycle module, for soil erosion processes in the soil erosion module (equations A5, A6 and S32 in the Appendices), and accumulation functions for the crop growth process in the crop growth module (equation S7 in the supplementary material).

2.2.2 Basic datasets and spatial delineation

The indispensable datasets for model setup are GIS data, daily meteorological data series, social and economic data series and dam attribute data. Several monitoring data series are needed for model calibration, such as runoff and water quality series in river sections, soil moisture and crop yield at the field scale. Table 1 shows all of the detailed datasets and their usages.

The hydrological toolset of Arc GIS platform is used to delineate all the spatial calculation units based on DEM, land-use/cover data. The sub-basin attributes (e.g., location, evaluation, area, land surface slope and slope length, land-use/cover areas) and flow routing relationship between sub-basins are obtained during this procedure.

2.3 Study area and model testing

In this study, our model was applied to a highly regulated and heavily polluted catchment (the Shaying River Catchment) in China. The simulated water-related components contained daily runoff and water quality concentrations at river sections, spatial patterns of diffuse source pollution load and crop yield at sub-basin scale.

2.3.1 Study area

The Shaying River Catchment (112°45′–113°15′E, 34°20′–34°34′N), which is the largest sub-basin of the Huai River Basin in China, is selected as the study area (Figure 6a). The drainage area is 36,651 km^2 with a mainstream of 620 km. The average annual population (2003–2008) (Figure 6b) is 32.42 million with rural population of 23.70 million. The average annual stocks include 8.30 million of big animals (cattle, pigs and sheep) and 178.42 million of poultries (Figure 6c). The average annual use of chemical fertilizer is 1.55 million ton (N: 38%–51%, P: 16%–25% and others: 23%–47%) (Figure 6d). The catchement is located in the typical warm temperate and semi-humid continental climate zone. The annual average temperature and rainfall are 14–16°C and 769.5 mm, respectively. The Shaying River is the most seriously polluted tributary with a pollutant load contribution of over 40% in the whole Huai River and is usually known as the water environment barometer of the Huai River mainstream. To reduce flood or drought disasters, 24 reservoirs and 13 sluices, whose regulation capacities are over 50% of the total annual runoff, have been constructed and fragmented the river into several impounding pools.

2.3.2 Model setup

All data sets for model setup and calibration were collected from the government bureaus, official books and scientific references. The detailed descriptions were presented in Tables S2 and S3 of the supplementary material. The resolutions of GIS and weather input data were quite satisfactory for the model application. However, most data on water quality, ecology and agricultural management were at monthly or annual temporal scale. The data for economy, agricultural management and diffuse source load were collected from individual administrative regions. Both the temporal and spatial scales were larger than the required daily scale or spatial calculation units (sub-basin, land-use/cover and crop). In these cases, the data values were uniformly distributed to the required temporal and/or spatial scales, such as the input of point sources, social and economic data.

The Shaying River Catchment was divided into 46 sub-basins. According to the land-use/cover classification standard of China (CNS,2007), the main land-use/cover types were dryland agriculture (84.04%), forest (7.66%), urban (3.27%), grassland (2.68%), water (1.43%), paddy land (0.91%) and unused land (0.01%).The soil input parameters (the contents of sand, clay and organic matters) were calculated based on the percentage of soil types in each sub-basin. The main crops were early rice and late rice in the paddy land, winter wheat and corn in the dryland agriculture. The main agricultural management schemes (fertilize, plant, harvest and kill) were summarized by field investigation in the studies of Wang *et al*. (2008) and Zhai *et al.* (2014) (Table S3). Crop rotations and management schemes were considered in the model by setting the start time, the duration of management and the fertilizer amounts. Two fertilizations (base and additional fertilization) were considered in the model during the complete growth cycle of a certain crop. The areas of sub-basin, land-use/cover and crop units ranged from 46.48 km^2 to 3,771.15 km^2, from 0.04 km^2 to 2,762.5 km^2, and from 3.73 km^2 to 2,762.5 km^2, respectively.

The daily precipitation series from 2003 to 2008 at 65 stations were interpolated to each sub-basin using the inverse distance weighting method, while the daily temperature series at six stations were interpolated using the nearest-neighbor interpolation method. The social and economic data (e.g., population and livestock in the rural area, chemical fertilizer amounts) were calculated for each sub-basin based on the area percentage.

Moreover, 5 reservoirs, 12 sluices and over 200 wastewater discharge outlets were considered according to their geographical positions. The farm manure from rural living and livestock farming were considered as diffuse source owing to their scattered characteristics and the deficient sewage treatment facilities in the rural areas.

2.3.3 Model evaluation

The observation series of daily runoff and NH_4-N concentration were used to calibrate the model parameters. There were five regulated stations (Luohe, Zhoukou, Huaidian, Fuyang and Yingshang) and one less-regulated station (Shenqiu) which is the downstream station situated far from water projects. Moreover, given that the observed yields of diffuse pollutant loads and crops were hard to collect for the whole catchment, only the statistical results from official reports or statistical yearbooks (Wang, 2011; Henan Statistical Yearbooks, 2003, 2004 and 2005) were collected to validate the model performances.

We selected LH-OAT for parameter sensitivity analysis and SCE-UA for parameter calibration in the PAT. To reduce the dimensions of the calibration problem, we restricted SCE-UA to calibrate only the sensitive parameters defined by LH-OAT, whereas the rest parameters remained constants. The selected evaluation indices of model performance were *bias*, *r* and *NS*. However, *NS* was sensitive to extreme value, outlier and number of the data points, and was not commonly used in environmental sciences (Ritter and Muñoz-Carpena, 2013). Thus NS was not used to evaluate the NH_4-N concentration simulation.

The model calibration was conducted by the following steps. Hydrological parameters were calibrated first against the observed runoff series at each station from upstream to downstream, and then water quality parameters against the observed NH_4-N concentration series. The calibration and validation periods were from 2003 to 2005 and from 2006 to 2008, respectively. The weighted sum method was usually used to comprehensively handle multi-objectives (Efstratiadis and Koutsoyiannis, 2010). In this study, single objective functions were formed by equally weighting the evaluation indices as (f_{runoff} and $f_{NH4\text{-}N}$) because the case study was only a demonstration of the model performance.

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

Moreover, the effect of dam regulation was considered because of the high regulation in most rivers. The dam and sluice regulation usually altered the intra-annual distribution of flow events, such as flattening high flow and increasing low flow. The simulation performances of high and low flows were separately evaluated and the effectiveness of the DRM was tested by comparing the simulation with and without the consideration of dam regulation. The high and low flows were determined by the cumulative distribution function (CDF). A threshold of 50% was used for easy presentation, i.e., the flow was treated as high flow (or low flow) if its percentile was greater than (or smaller than) the threshold.

3. Results

3.1 Parameter sensitivity analysis

Nine sensitive parameters were detected for runoff simulation by LH-OAT (Table 2), including soil related parameters W_{fc} (field capacity), W_{sat} (saturated moisture capacity), K_r (interflow yield coefficient) and K_{sat} (steady state infiltration rate); TVGM parameters g_1 (basic surface runoff coefficient) and g_2 (influence coefficient of soil moisture); baseflow parameters K_g (baseflow yield coefficient) and T_g (delay time for aquifer recharge); and evapotranspiration parameter K_{ET} (adjusted factor of actual evapotranspiration). All of these parameters controlled the main hydrological processes in which soil water and evapotranspiration processes were distinctly important and explained 54.3% and 23.2% of the runoff variation, respectively.

For NH_4-N concentration simulation, over 90% of observed NH_4-N concentration variations 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. The main explanation was that hydrological processes provided the hydrological boundaries that affected the diffuse source load into rivers and the degradation and settlement processes of NH_4-N in water bodies (van Griensven *et al.*, 2002). NH_4-N concentration was further influenced by the settlement and biological oxidation. Moreover, it was a competitive relationship between COD and NH_4-N to consume DO of water bodies in a certain limited level (Brown and Barnwell, 1987).

3.2 Hydrological simulation

The runoff simulations fitted the observations well at all the stations (Figure 7 and Table 3). The *biases* were very close to 0.0 at all the regulated stations except Zhoukou with an underestimation (*bias*: 0.24 for calibration and 0.41 for validation) and Luohe with an overestimation (bias: –0.52 for validation). The obvious biases were caused by the average objective function of all three evaluation rather than the *bias* only. The *r* values ranged from 0.75 (Luohe for validation) to 0.92 (Yingshang for calibration) with the average value of 0.85, whereas the *NS* values ranged from 0.51 (Luohe for validation) to 0.84 (Yingshang for calibration) with the average value of 0.70. The results of the regulated stations were a little worse than those of the less-regulated station (Shenqiu) owing to the regulation.

By comparing the simulations with the observations from 2003 to 2008, we saw that the high and low flows were always overestimated if the model did not consider the regulations (Figure 8). Except the high flows at Zhoukou, both high and low flows at all the stations were simulated well when the dam and sluice regulation was considered (Table 4). The best fitting was at Fuyang, particularly for the high flow simulation (*bias*=0.10, *r*=0.89 and *NS*=0.78). From unregulation to regulation settings, the improvements measured by f_{runoff} ranged from –0.08 (Zhoukou) to –0.29 (Huaidian) for high flow simulations, from –0.05 (Zhoukou) to –0.31 (Huaidian) for average flow simulations, and from –1.97 (Fuyang) to –3.91 (Yingshang) for low flow simulations except Zhoukou (1.28). The improvements in the low flow simulations were very obvious. However, their performances still needed to be improved further, particularly for the underestimation at Zhoukou and Huaidian. The possible reasons were as follows. On one hand, the applied evaluation indices (*r* and *NS*) were known to emphasize the high flow simulation rather than the low flow simulation (Pushpalatha *et al.*, 2012) and the objective of autocalibration was to obtain the optimal solution for the average of three evaluation indices rather than the *bias* only. The slight sacrifice of *bias* improved the overall simulation performance evaluated by all three indices. One the other hand, the dam regulation module still could not fully capture the low flows.

Furthermore, the model performances on monthly flows were even better, particularly for *r* and *NS*. The *r* values ranged from 0.87 (Luohe for both calibration and validation) to 0.95 (Fuyang for calibration) with the average value of 0.92, whereas the *NS* values ranged from 0.67 (Luohe for validation) to 0.94 (Shenqiu for validation) with the average value of 0.80. Compared with the existing results at the same stations by SWAT (Zhang *et al.*, 2013), the flow simulations at the downstream stations were improved although they became a little worse at the upstream stations (Luohe and Zhoukou for calibration). In particular, the total water volume and agreements with the observations (i.e., *bias* and *NS*) were well captured.

3.3 Water quality simulation

The simulated concentrations of NH_4-N matched well with the observations according to the evaluation standard recommend by Moriasi *et al.* (2007) (Figure 9 and Table 5). The *r* values were over 0.60 for all the stations except Zhoukou (0.56 for validation), Yingshang (0.49 for validation) and Shenqiu (0.41 for validation) and the average value was 0.67. The *biases* were considered as “acceptable” with a range from –0.27 (Fuyang for validation) to 0.29 (Zhoukou for calibration). The best simulation was at Luohe Station. The obvious discrepancies between the simulations and observations often appeared in the period from January to May because of the poor simulation performances on the low flows. Although the *biases* changed markedly from calibration to validation at Fuyang and Yingshang stations, the model performances were still acceptable. The possible explanation was that the *biases* for corresponding runoff simulations at these two stations also changed.

Compared with the results without the consideration of regulation, the simulation results were obviously improved when the regulation was considered except those at Fuyang Station in the calibration period. The decreases in $f_{NH4\text{-}N}$ value ranged from 0.10 (Huaidian for calibration) to 0.49 (Zhoukou for validation) although there was a slight increase at Fuyang for the calibration (0.02). Therefore, it was concluded that the consideration of dam and sluice regulation played an important role in the water quality simulation. In the upper stream of Shaying River, the flow was small and the NH_4-N concentration decreased obviously because of the degradation and settlement of large water storage. In the downstream of Shaying River, the NH_4-N concentration increased because of the pollutant accumulation and the decreasing flow from dams and sluices owing to the regulation (Zhang *et al*., 2010). Therefore, the simulated concentrations without regulation were usually overestimated or higher than the simulation with regulation at the upstream stations (Luohe and Zhoukou). However, the concentrations were underestimated at the downstream stations (Huaidian, Fuyang and Yingshang). The largest differences between the simulations with and without the consideration of regulation appeared at Zhoukou.

The spatial pattern of average annual load of diffuse source NH_4-N was shown in Figure 10a. The estimated annual yield rates ranged from 0.048 t km^{-2} $year^{-1}$ to 11.00 t km^{-2} $year^{-1}$ with the average value of 0.73 t km^{-2} $year^{-1}$. The yield in each administrative region was summarized from the results of each sub-basin according to the area percentage of sub-basin in each administrative region. Compared with the statistical load of each administrative region based on the soil erosion, land-use/cover and fertilizer amount in the official report (Wang, 2011), the *bias* of simulated diffuse source load in the whole region was 21.31% when the two regions with the biggest *biases* (Fuyang and Pingdingshan) were excluded as outliers. The high load regions were in the middle of Pingdingshan, Xuchang, Zhengzhou, Fuyang and Zhoukou regions. The spatial pattern was significantly correlated with the distribution of paddy area (r=0.506, p<0.001) and rice yield (r=0.799, p<0.001) (Figures 10 b and c). The fertilizer losses in the paddy areas might be the primary contributor to the diffuse source NH_4-N load because the average nitrogen loss coefficient in China was just 30%–70% in the paddy areas, which was higher than that in the dryland agriculture (20%–50%) (Zhu, 2000; Xing and Zhu, 2000).

Summarized from the collected data for model input, the observed average load of point source NH_4-N into rivers was approximately 4.70×10^4 t $year^{-1}$ in the Shaying River Catchment. The diffuse source contributed 38.57% of the overall NH_4-N load on average from 2003 to 2005, and this value was slightly higher than the statistical results (29.37%) given in the official report (Wang, 2011). Moreover, the diffuse source contributions at the stations ranged from 31.72% (Huaidian) to 47.13% (Shenqiu). Compared with the diffuse source loads in the individual administrative regions in 2000, the simulated loads tended to increase from 2003 to 2005 except in Kaifeng region. The yields in Fuyang and Pingdingshan regions increased at highest rates. The primary pollution source in the Shaying River Catchment was still the point source, but the diffuse source was also an important concern. In term of spatial variation, the contribution of diffuse source to the pollutant load was high in the upstream and low in the middle and downstream because the point source emission was usually concentrated in the middle and downstream. Therefore, compared with the results in Zhang *et al.* (2013), the overall simulation performance of NH_4-N concentration was also improved remarkably by considering the detailed nutrient processes in the soil layers.

3.4 Crop yield simulation

The simulated corn yield and its spatial pattern were shown in Figure 11. The average annual yields were summarized at sub-basin scale and ranged from 0.08 to 326.95 t km^{-2} $year^{-1}$ with the average value of 76.84 t km^{-2} $year^{-1}$. The yield of each administrative region was further summarized and compared with the data from statistical yearbooks from 2003 to 2005 (Henan Statistical Yearbook, 2003, 2004 and 2005). The high-yield regions were Luohe, Fuyang and Zhoukou in the middle and downstream where the primary land-use/cover was the dryland agriculture (93.12%, 95.87% and 93.18%, respectively). The crop yields in Luohe, Nanyang and Kaifeng regions were well simulated. The total yield was underestimated in the whole basin with a *bias* of 19.93%. The discrepancies might be caused by the boundary mismatch between the administrative region and sub-basin, spatial heterogeneities of human agricultural activities and inaccurate cropping pattern used in such huge regions. A high-resolution remote sensing image and field investigation might be helpful to improve the model performance.

4. Discussion

4.1 Comparison with other models

It is a natural tendency that models grow in complexity in order to capture more interactions of complex water-related processes in the real basins because of more and more available observations and improved accuracies (Beven, 2006). Our proposed model was developed in this direction and tended to benefit integrated river basin management although the model applicability needs to be further evaluated in different regions. In comparison with most existing models, our proposed model considered all the water-related processes as an integrated system rather than isolated systems for individual processes.

Our model provided competitive simulation results in the Huai River Basin (Figures 7–9; Tables 3–5). Several typical models were also applied in this basin, such as SWAT for the monthly runoff and water quality simulation at the regulated stations (Zhang *et al.*, 2012), SWAT and Xinganjiang models for the daily runoff simulation at the unregulated upstream stations (Shi *et al.*, 2013) and DTVGM for daily runoff simulation (Ma *et al.*, 2014). Compared with the results of these models, our model generally performed better on the runoff or water quality simulations. In particular, our model performed even better than SWAT at the regulated stations as more detailed dam regulation rules and soil biochemical processes were considered. For example, the average values of f_{runoff} at the monthly scale decreased from 0.32 (SWAT in Zhang *et al.*, 2012) to 0.15 (our model) at the regulated stations. The average values of $f_{NH4\text{-}N}$ decreased from 0.47 (SWAT in Zhang *et al.*, 2012) to 0.27 (our model). Moreover, both the Xinanjiang model and DTVGM are limited to simulate the flow series at the unregulated or less-regulated stations because they do not consider the dam regulation in their current model frameworks (Shi *et al.*, 2013; Ma *et al.*, 2014).

4.2 Equifinality

Until now, our understandings of water-related processes are still ambiguous and it is hard to describe all these processes in the real-word systems from strong physical foundations (Beven and Freer, 2001; Beven, 2006; Hrachowitz *et al.*, 2014).

Empirical equations are usually adopted to approximate the physical processes with numerous unknown parameters, especially in the large scale models. A single output variable of models is associated with multiple processes and many parameters. For examples, SWAT contains over 200 parameters (Arnold *et al.*, 1998) and DNDC has nearly 100 parameters (Li *et al.*, 1992). Pohlert *et al.* (2006) reported that six hydrological and 12 N-cycle sensitive parameters were detected in SWAT-N for the simulation of water flow and N leaching. In the case study, nine and 14 sensitive parameters of our model were detected for runoff and NH_4-N simulation, respectively (Table 2). Therefore, due to the large numbers of model parameters and limited observations, most existing models are subject to equifinality, which is more serious if more water-related processes are considered, or more sub-basins are delineated for the distributed models.

Several strategies would be helpful to alleviate the equifinality, such as field experiments on the physical parameters (Kirchner, 2006), the utilization of more observed processes, multiple evaluation measures for a single predicted component (Her and Chaubey, 2015), parameter regularization and process constraints (Tonkin and Doherty, 2005; Pokhrel *et al.*, 2008; Euser *et al.*, 2013). Moreover, some attempts are made to move away from traditional curve fitting towards more process consistency and efficient model selection techniques (Hrachowitz *et al.*, 2014; Fovet *et al.*, 2015).

For our model, all the independent calibration and validation data sets were specified in Table 1 and most widely-used measures of model performances were also provided in the PAT. In the case study, we also employed several observation sources (e.g., runoff and water quality observations at different stations, the diffuse pollution load and crop yield data) and used three measures to evaluate model performance for the individual components (e.g., *bias*, *r* and *NS*). To make full use of the existing data in practice, parameter sensitivity analysis would be an effective way to reduce dimensionality in model calibration, and then focus only on the critical processes and parameters that are sensitive to model outputs (van Griensven *et al.*, 2006). Model autocalibration would be efficient to obtain the optimal simulations from numerous samples in multi-dimensional parameter spaces.

4.3 Model limitations

It should be noted that our extended model still has several limitations:

(1). The mathematical descriptions of groundwater, crop growth processes and agriculture management practices were still inaccurate. The current version focused on the detailed descriptions of hydrological and nutrient cycle in the soil layers and water bodies, and the consideration of dam regulation. Satisfactory performances on water quantity and quality simulation were achieved in our case study. However, the simulations for groundwater, diffuse pollution, crop yield in the agriculture regions could be improved further. The stratification of water impounding in the water quality module should be considered if the high resolution bathymetric data of dams or lakes are available.

(2). High parameterization is an inevitable issue because of its all-inclusive framework. Our model considered the main water-related processes in the hydrological, ecology and water quality subsystems but numerous processes were still controlled by unmeasurable parameters because of their empirical and/or scale dependent nature (Her and Chaubey, 2015). Although the parameter sensitivity analysis and calibration are widely used to handle the high parameterization issue, the equifinality and parameter uncertainty are still inevitable because of the insufficient observations and the complex interactions among different subsystems.

5. Conclusions

In this study, TVGM hydrological model was extended primarily to an integrated water system model to address the complex water issues emerging in the basins. The model performance was demonstrated in the Shaying River Catchment, China. The model provided a reasonable tool for the effective water governance by simultaneously simulating several indicative components of water-related processes including the hydrological components (e.g., runoff, soil moisture, evaporation and plant transpiration, water storage in the dams and sluices), water quality components (e.g., diffuse pollution source load, water quality concentrations in water bodies) and ecological components (e.g., crop yield) which could be calibrated if observations were available. The case study showed that the simulated runoffs at most stations fitted the observations well in the highly regulated Shaying River Catchment. All the evaluation criteria were acceptable for both the daily and monthly simulations at most stations. This model well simulated the discontinuous daily NH_4-N concentration and properly captured the spatial patterns of diffuse pollution load and corn yield.

Owing to the heterogeneity of spatial data in large basins and insufficient observations of individual subsystems, not all the results were acceptable and several processes were still not well calibrated (such as low flow events, diffuse pollution source load and crop yield). More available data and improved data quality will reduce the model uncertainty and equifinality problem, especially the higher-resolution data for surface conditions, water quality, agricultural management and socio-economic data. The model would be improved by further considering more accurate human activities in the agricultural management, calibrating multiple components by multi-objective optimization and model uncertainty analysis because of the interactions and tradeoffs among different processes. The over-parameterization and the reasonable prior parameter conditions should also be treated carefully in applications. Advanced analysis technologies would benefit the future model development, such as model selection techniques, parameter regularization. Moreover, an easy-used operational software package can broaden the model's applications in different regions. More case studies are needed to further demonstrate its applicability.

Appendix A: Hydrological cycle module

The basic water balance equation is

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

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

E_s and E_p are determined by the potential evapotranspiration (E_0, mm), leaf area index (LAI, m^2/m^2) and surface soil residues (rsd, t/ha) (Ritchie, 1972) as

$$
\begin{cases}
E_a = E_t + E_s \le E_0 \\
E_p = \begin{cases} LAI \cdot E_0/3 & 0 \le LAI \le 3.0 \\ E_0 & LAI > 3.0 \end{cases} \\
E_s = E_0 \cdot exp(-5.0 \times 10^{-5} \cdot rsd)
\end{cases}
\tag{A2}
$$

where E_0 is calculated by Hargreaves method (Hargreaves and Samani, 1982).

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

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

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

The interflow (Rss, mm) and baseflow (Rbs, mm) have linear relationships with the soil moistures in the upper and lower layers, respectively (Wang *et al.*, 2009) as

$$
\begin{cases}
Rss = k_{ss} \cdot SW_u \\
Rbs = k_{bs} \cdot SW_l
\end{cases}
\tag{A4}
$$

where k_{ss} and k_{bs} are the yield coefficients of interflow and baseflow, respectively; SW_l is the soil moisture in the lower layer (mm).

The infiltration from the upper to lower soil layers is calculated using storage routing method (Neitsch *et al.*, 2011) as

$$
\begin{cases}
W_{\inf} = (SW_u - W_{fc}) \cdot [1 - \exp(-24/T_{\inf})] \\
T_{\inf} = (W_{sat} - W_{fc}) / K_{sat}
\end{cases}
\tag{A5}
$$

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

The calculation of overland flow routing is adopted from Neitsch *et al.* (2011) as

$$
\begin{cases}
Q_{overl} = (Q'_{overl} + Q_{stor,i-1}) \cdot \left[1 - \exp(-T_{retain}/T_{route})\right] \\
T_{route} = T_{overl} + T_{rch} = \dfrac{L_{overl}^{0.6} \cdot n_{overl}^{0.6}}{18 \cdot slp_{overl}^{0.3}} + \dfrac{0.62 \cdot L_{rch} \cdot n_{rch}^{0.75}}{A^{0.125} \cdot slp_{rch}^{0.375}}
\end{cases}
\tag{A6}
$$

where Q_{overl} is the overland flow discharged into main channel (mm); Q'_{overl} is the lateral flow amount generated in the sub-basin (mm), $Q_{stor,i-1}$ is the lateral flow in the previous day (mm); T_{retain} is the retain time of flow (days); T_{route} is the flow routing times in sub-basin (days); T_{overl} and T_{rch} are the routing times of overland flow and river flow, respectively (days); L_{overl} and L_{rch} are the lengths of sub-basin slope and river, respectively (km); slp_{overl} and slp_{rch} are the slopes of sub-basin and river, respectively (m/m); n_{overl} and n_{rch} are the Manning's roughness coefficients for sub-basin and river, respectively (m/m); and A is the sub-basin area (km^2).

Appendix B: Soil biochemical module

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

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

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

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

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

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

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

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

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

The NH_4 amount (FIX_{NH4}, kg/ha) absorbed by clay and organic matters is estimated by

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

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

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

where K_{NH4} and K_{H2O} are the dissociation constants for NH_4^+:NH_3 equilibrium, H^+: OH^- equilibrium, respectively; NH_{4m} and NH_{3m} are the NH_4^+ and NH_3 concentrations (mol/L) in the liquid phase, respectively; AM and D are the accumulated NH_3 loss (mol/cm^2) and diffusion coefficients (cm^2/d^2), respectively.

The nitrification rate ($dNNO$, kg/ha/day) is a function of the available NH_4^+, soil temperature and moisture; N_2O emission is a function of soil temperature and soil NH_4^+ concentration, and are given as

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

where K_{35} is the nitrification rate at 35 °C (mg/kg/ha); $\mu_{sw,n}$ is the soil moisture adjusted factor for nitrification.

Denitrification: The growth rate of denitrifier ($(dB/dt)g$, kg/ha/day) is proportional to their respective biomass and is calculated by double Monod kinetics equation as

$$\begin{cases} (d\mathrm{B}/dt)_g = \mu_{DN} \cdot B(t) \\ \mu_{DN} = \mu_{t,dn} \cdot (u_{NO_3} \cdot \mu_{PH,NO_3} + u_{NO_2} \cdot \mu_{PH,NO_2} + u_{N_2O} \cdot \mu_{PH,N_2O}) \\ u_{N_xO_y} = u_{N_xO_y,\max} \cdot (C / K_{C,1/2} + C) \cdot (N_xO_y / K_{N_xO_y,1/2} + N_xO_y) \end{cases} \qquad \text{(B7)}$$

where B is the denitrifier biomass (kg); μ_{DN} is the relative growth rate of the denitrifiers; u_{NxOy} and $u_{NxOy,max}$ are the relative and maximum growth rates of NO_2^-, NO_3^- and N_2O denitrifiers, respectively. $K_{C,1/2}$ and $K_{NxOy,1/2}$ are the half velocity constants of C and N_xO_y, respectively; $\mu_{PH,NxOy}$ and $\mu_{t,dn}$ are the reduction factors of soil pH and temperature, respectively. The mathematical expressions are given as

$$\begin{cases} \mu_{PH,NO_3} = 7.14 \cdot (pH - 3.8)/22.8 \\ \mu_{PH,NO_2} = 1.0 \\ \mu_{PH,N_2O} = 7.22 \cdot (pH - 4.4)/18.8 \\ \mu_{t,dn} = \begin{cases} 2^{(T-22.5)/10} & if\ T < 60°C \\ 0 & if\ T \geq 60°C \end{cases} \end{cases} \quad \text{(B8)}$$

The death rate of denitrifier ($(dB/dt)_d$, kg/ha/hr) is proportional to denitrifier biomass and is given as

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

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

The consumption rates of dissolved C and CO_2 production are calculated as

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

where $\mu_{sw,d}$ is the soil moisture adjusted factor for denitrification.

The NO_3^-, NO_2^-, NO and N_2O consumption are calculated as

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

where M_{NxOy} and Y_{NxOy} are the maintenance coefficient (1/hr), maximum growth yield on NO_3^-, NO_2^-,NO or N_2O (kg/ha/hr), respectively.

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 rates are the functions of adsorption coefficients of the gases in soils and to the air filled porosity of the soil and are given as.

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

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

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

$$Leach_{NO_3} = W_{\inf} \cdot \mu_{CLAY} \cdot \mu_{soc} \tag{B14}$$

where $Leach_{NO3}$ is the NO_3^- leaching rate; μ_{CLAY} and μ_{soc} are the influence coefficients of clay content and soil organic C, respectively.

B.3 P cycle

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

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

The water balance model of dam or sluice is considered the inflow, outflow, precipitation, evapotranspiration, seepage and water withdraw. The equation is:

$$\Delta V = V_{flowin} - V_{flowout} + V_{pcp} - V_{evap} - V_{seep} - V_{withd} \tag{C1}$$

where ΔV, V_{flowin} and $V_{flowout}$ are the water storage variation, water volumes of entering and flowing out, respectively (m^3), and are calculated by HCM; V_{pcp}, V_{evap} and V_{seep} are the volumes of precipitation, evaporation and seepage, respectively (m^3), and are the functions of surface water area and water storage. V_{withd} is the water withdraw volume (m^3) by human and is given as a model input.

According to the design data of dam and sluice in China, there is a particular relationship among water level, storage and outflow. The outflow is determined by the water level or water storage volume. The relationships are described by equations.

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

where *V* and *H* are the water storage volume (m^3) and water level (m) during a day, respectively; *f'()* and *f''()* are the functions which could be determined by statistical analysis methods (e.g., correlation analysis, linear or non-linear regression analysis, polynomial regression analysis and least squares fitting).

Appendix D: Evaluation indices of model performance

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

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

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

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

Nash–Sutcliffe efficiency: $$NS = 1 - \sum_{i=1}^{N}(O_i - S_i)^2 \Big/ \sum_{i=1}^{N}(O_i - \overline{O})^2 \tag{D5}$$

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

Acknowledgements

This study was supported by the Natural Science Foundation of China (No. 41271005), the China Youth Innovation Promotion Association CAS (No. 2014041), the Key Project for the Strategic Science Plan in IGSNRR, CAS (No. 2012ZD003), the Endeavour Research Fellowship, China Visiting Scholar Project from China Scholarship Council, and the CSIRO Computational and Simulation Sciences Research Platform. The authors would like to thank Dr. Yongqiang Zhang, Mr. James R Frankenberger for their participation in our internal review procedure, Dr. Markus Hrachowitz and Dr. Christian Stamm for improving the quality and presentation of the manuscript, and the anonymous reviewers for their valuable comments and suggestions.

References

Abbott, M.B, Bathurst, J.C., Cunge, J.A., O'Connell, P.E. and Rasmussen, J.: An Introduction to the European System: Systeme Hydrologique Europeen (SHE). *J. Hydrol.* 87: 61–77, 1986.

Abrahamsen, P., and Hansen, S. Daisy: an open soil-crop-atmosphere system model. *Environ. Modell. Softw.*, 15(3): 313–330, 2000.

Arheimer, B. and Brandt, M.: Modelling nitrogen transport and retention in the catchments of southern Sweden. *Ambio* 27(6):471–480, 1998.

Arheimer, B. and Brandt, M.: Watershed modelling of non-point nitrogen pollution from arable land to the Swedish coast in 1985 and 1994. *Ecol. Engin.* 14:389–404, 2000.

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.

Beven, K.J. and Kirkby, M.J.: A physically based variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* , 24(1):43–69, 1979.

Beven, K.J. A manifesto for the equifinality thesis. *J. Hydrol.*, 320(1–2):18–36, 2006.

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.

Borah, D. K., and Bera, M.: Watershed-scale hydrologic and nonpoint-source pollution models: Review of application. *Trans. ASAE* 47(3): 789–803, 2004.

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.

China's national standard (CNS): *Current land use condition classification* (GB/T21010–2007), General administration of quality supervision, inspection and quarantine of China and Standardization administration of China, Beijing, China, 2007.

Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA–II. *IEEE T. Evolut. Comput*. 6(2), 182–197, 2002.

Deng, J., Zhu, B., Zhou, Z. X., Zheng, X. H., Li, C. S., Wang, T., and Tang, J. L.: 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.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures. *Hydrol. Earth Syst. Sci.*, 17 (5), 2013.

Fovet, O., Ruiz, L., Hrachowitz, M., Faucheux, M., and Gascuel-Odoux, C.: Hydrological hysteresis and its value for assessing process consistency in catchment conceptual models, *Hydrol. Earth Syst. Sci.*, 19, 105–123, 2015.

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*, 50: 1211–1250, 2007.

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

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

Her, Y., and Chaubey, I.: Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty. *Hydrol. Process.*, 29:4220–4237, 2015.

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.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., and Gascuel‐Odoux, C.: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity." *Water Resour. Res.* 50, 9: 7445–7469, 2014.

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.

Johnsson, H., Bergstrom, L., Jansson, P. E., and Paustian, K.: Simulated nitrogen dynamics and losses in a layered agricultural soil. *Agr. Ecosyst. Environ.*, 18(4), 333–356,1987.

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. W03S04 doi: 10.1029/2005WR004362.

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

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

Kennedy, J.: Particle swarm optimization, Encyclopedia of Machine Learning. Springer USA, 760–766, 2010.

Krysanova, V., Mueller-Wohlfeil, D.I. and Becker, A.: Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecol. Model.*, 106, 261–289, 1998.

Li, C., Frolking, S. and 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.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges: A Simple hydrologically based model of land surface water and energy fluxes for GSMs, *J. Geophys. Res.*, 99(D7), 14,415–14,428, 1994.

Lindström, G., Pers, C.P., Rosberg, R., Strömqvist, J. and Arheimer, B.:. Development and test of the HYPE (Hydrological Predictions for the Environment) model - A water quality model for different spatial scales. *Hydrol. Res.* 41.3–4:295–319,2010.

Ma, F., Ye, A., Gong, W., Mao, Y., Miao, C., and Di, Z. An estimate of human and natural contributions to flood changes of the Huai River. *Global Planet Change*, 119, 39–50, 2014.

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 uncertainty assessment: Incoherence of the GLUE methodology”, *J. Hydrol.*, 338, 319–324, 2007.

McDonnell, J. J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz, C., Hooper, R., Kirchner, J., Roderick, M.L., Selker, J., and Weiler, M.: Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resour. Res.*, 43, W07301, doi:10.1029/2006WR005467,2007.

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

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part I—A discussion of principles. *J. Hydrol*. 27(3), 282–290, 1970.

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

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

Paola, C., Foufoula-Georgiou, E., Dietrich, W.E., Hondzo, M., Mohrig, D., Parker, G., Power, M.E., Rodriguez-Iturbe, I., Voller, V., Wilcock, P.: Toward a unified science of the Earth's surface: opportunities for synthesis among hydrology, geomorphology, geochemistry, and ecology. *Water Resour. Res*., 42, 2006. W03S10. DOI: 10.1029/2005WR004336.

Pushpalatha, R., Perrin, C., Le Moine, N., and Andréassian, 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.

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.

Pohlert, T., L. Breuer, J.A. Huisman, and H.-G. Frede.: Integration of a detailed biogeochemical model into SWAT for improved nitrogen predictions-model development, sensitivity and uncertainty analysis. *Ecol. Model.* 203:215–228, 2006.

Pokhrel, P., Gupta, H. V. and Wagener, T.: A spatial regularization approach to parameter estimation for a distributed watershed model, *Water Resour. Res*., 44, 2008. W12419, doi:10.1029/2007WR006615.

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

Shi, P., Chen, C., Srinivasan, R., Zhang, X., Cai, T., Fang, X., Qu, S., Chen, X., and. Li, Q.: Evaluating the SWAT model for hydrological modeling in the Xixian watershed and a comparison with the XAJ model. *Water Resour. Manag.*, 25(10), 2595–2612, 2011.

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.

Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C., and Lindström, G.: Water and nutrient predictions in ungauged basins: set-up and evaluation of a model at the national scale. *Hydrolog. Sci. J.*, 57(2), 229–247,2002.

Tattari, S., Bärlund, I., Rekolainen, S., Posch, M., Siimes, K., Tuhkanen, H. R., and Yli-Halla, M.: Modeling sediment yield and phosphorus transport in Finnish clayey soils. *T. ASABE*, 44(2), 297–307, 2001.

Tonkin, M. J., and Doherty, J.: A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resour. Res.*, 41(10), 2005. W10412, doi:10.1029/2005WR003995.

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.

Wang, G. S., Xia J., Tan G., and Lu A.F.: A research on distributed time variant gain model: A case study on Chao River basin (in Chinese), *Prog. Geogr.,* 21(6), 573–582, 2002.

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, W00A11, doi:10.1029/2007WR006768,2009.

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)

Wang, X.: Summary of Huaihe River Basin and Shandong Peninsula Integrated Water Resources Plan, *China water resources*, 23,112–114, 2011.

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.

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)

Table 1. The data sets and their categories used in the model

Category	Data	Objectives	Controlled processes
GIS	DEM	Elevation, area, longitude and latitude, slopes and lengths of each sub-basin and channel	Hydrology and water quality
	Land-use/cover map	Land-use/cover types and their corresponding areas in each sub-basin	Hydrology, water quality and ecology
	Soil map	Soil physical properties of each sub-basin such as bulk density, saturated conductivity	
Weather	Daily precipitation	Daily precipitation of each sub-basin	Hydrology
	Daily maximum and minimum temperature	Daily maximum and minimum temperature of each sub-basin	
Hydrology	Observed runoff or other hydrological components, etc.	Hydrological parameter calibration	Hydrology
Water quality	Urban wastewater discharge outlets and discharge load	Model input of point source pollutant load	Water quality
	Water quality observations (concentration or load), etc.	Water quality parameter calibration	
Ecology	Crop yield, leaf area index, etc.	Ecological parameter calibration	Ecology
Economy	Basic economic statistical indictors	Populations, breeding stock of large animals and livestock, water withdrawal in each sub-basin	Hydrology and water quality
Water projects	Design data attribute parameters	Regulation rules of dams or sluices	Hydrology
Agricultural management	Fertilization and irrigation types, timing and amount, time of seeding and harvest, and crop types	Agricultural management rules of each sub-basin	Water quality and ecology

Table 2 Sensitive parameters, their value ranges and relative importance for runoff and NH_4-N simulations

Variables	Range	Definition	Relative importance for runoff (%)	Relative importance for NH_4-N (%)
W_{fc}	0.20 to 0.45	Field capacity of soil	32.73	11.10
W_{sat}	0.45 to 0.75	Saturated moisture capacity of soil	11.68	11.83
g_1	0 to 3	Basic surface runoff coefficient	7.30	10.34
g_2	0 to 3	Influence coefficient of soil moisture	10.54	12.11
K_{ET}	0 to 3	Adjustment factor of evapotranspiration	23.21	10.71
K_{ss}	0 to 1	Interflow yield coefficient	9.55	3.20
T_g	1 to 100	Delay time for aquifer recharge	1.74	-
K_{bs}	0 to 1	Baseflow yield coefficient	2.91	-
K_{sat}	0 to 120	Steady state infiltration rate	0.33	-
R_d(BOD)	0.02 to 3.4	BOD deoxygenation rate at 20 °C	-	6.62
R_{set}(BOD)	-0.36 to 0.36	BOD settling rate at 20 °C	-	3.60
R_d(NH_4)	0.1 to 1	Bio-oxidation rate of NH_4-N at 20 °C	-	1.97
K_{set}(NH_4)	0 to 100	Settling rate of NH_4-N in the reservoirs	-	14.17
K_d(BOD)	0.02 to 3.4	BOD deoxygenation rate in the reservoirs at 20°C	-	2.12
K_d(NH_4)	0.1 to 1.0	Bio-oxidation rate of NH_4-N in the reservoirs at 20 °C	-	4.51
Total relative importance			100.00	92.27

Table 3 Runoff simulation results for regulated and less-regulated stations

Stations	Periods	Daily flow				Monthly flow			
		bias	r	NS	f	bias	r	NS	f
Regulated stations									
Luohe	Calibration	0.00	0.84	0.70	0.15	0.00	0.87	0.71	0.14
	Validation	-0.52	0.75	0.51	0.42	-0.52	0.87	0.67	0.33
Zhoukou	Calibration	0.24	0.87	0.73	0.21	0.24	0.90	0.76	0.19
	Validation	0.41	0.79	0.55	0.36	0.41	0.91	0.70	0.26
Huaidian	Calibration	0.03	0.88	0.77	0.13	0.03	0.91	0.81	0.10
	Validation	0.12	0.76	0.54	0.27	0.12	0.87	0.70	0.18
Fuyang	Calibration	0.00	0.90	0.81	0.10	0.00	0.95	0.89	0.05
	Validation	0.14	0.88	0.76	0.17	0.14	0.94	0.86	0.11
Yingshang	Calibration	-0.13	0.92	0.84	0.12	-0.13	0.92	0.84	0.12
	Validation	0.16	0.87	0.74	0.18	0.16	0.93	0.82	0.13
Less–regulated stations									
Shenqiu	Calibration	0.00	0.91	0.82	0.09	0.00	0.94	0.88	0.06
	Validation	-0.13	0.83	0.67	0.21	-0.13	0.98	0.94	0.08

Table 4. The runoff simulation results at regulated stations with and without the dam regulation considered. Range means the difference of objective function value between regulations considered and not considered. If the range value is less than 0.0, then the simulation with regulation is better than that without regulation. Otherwise, the simulation without regulation is better.

Stations	Regulated capacity (%)	Flow event	Regulation considered				Regulation not considered				Range
			bias	r	NS	f	bias	r	NS	f	
Luohe	0.26	High	-0.16	0.97	0.92	0.09	-0.62	0.97	0.80	0.29	-0.20
		Low	-0.02	0.98	0.69	0.12	-1.46	0.99	-5.53	2.67	-2.55
		Average	-0.15	0.97	0.93	0.08	-0.68	0.96	0.82	0.30	-0.22
Zhoukou	1.31	High	0.21	0.98	0.93	0.10	-0.38	0.98	0.87	0.18	-0.08
		Low	1.00	0.00	-2.57	1.86	-0.64	0.99	-0.08	0.58	1.28
		Average	0.30	0.99	0.93	0.13	-0.41	0.98	0.89	0.18	-0.05
Huaidian	1.37	High	0.02	0.98	0.95	0.03	-0.64	0.98	0.68	0.32	-0.29
		Low	0.36	0.97	0.43	0.32	-1.51	0.98	-5.88	2.80	-2.48
		Average	0.06	0.98	0.96	0.04	-0.74	0.98	0.72	0.35	-0.31
Fuyang	2.21	High	0.04	0.98	0.96	0.03	-0.39	0.99	0.86	0.18	-0.15
		Low	0.17	0.99	0.87	0.10	-1.43	0.99	-3.78	2.07	-1.97
		Average	0.05	0.99	0.97	0.03	-0.50	0.99	0.88	0.21	-0.18
Yingshang	1.76	High	0.03	0.98	0.95	0.03	-0.44	0.99	0.86	0.20	-0.17
		Low	0.18	0.99	0.82	0.12	-1.77	0.95	-9.26	4.03	-3.91
		Average	0.05	0.99	0.96	0.03	-0.60	0.98	0.86	0.25	-0.22

Table 5. The comparison of NH_4-N simulation results between with and without dam regulation considered.

Stations	Periods	Regulated			Unregulated			Range	Ratio of diffuse source load (%)
		bias	r	f	bias	r	f		
Regulated stations									
Luohe	Calibration	-0.02	0.93	0.05	-0.67	0.60	0.54	-0.49	46.10
	Validation	-	-	-	-	-	-		
Zhoukou	Calibration	0.29	0.61	0.34	-0.56	0.38	0.59	-0.25	44.54
	Validation	0.27	0.56	0.36	-1.35	0.66	0.85	-0.49	
Huaidian	Calibration	0.22	0.73	0.25	0.49	0.80	0.35	-0.10	31.72
	Validation	0.02	0.67	0.18	0.22	0.51	0.36	-0.18	
Fuyang	Calibration	0.28	0.78	0.25	0.26	0.80	0.23	0.02	33.12
	Validation	-0.27	0.76	0.26	-0.38	0.56	0.41	-0.15	
Yingshang	Calibration	0.24	0.79	0.23	0.25	0.58	0.34	-0.11	33.26
	Validation	-0.24	0.49	0.38	-0.76	0.62	0.57	-0.19	
Less-regulated stations									
Shenqiu	Calibration	0.13	0.62	0.26	-	-	-	-	47.13
	Validation	0.16	0.41	0.37	-	-	-	-	

List of Figure Captions

Figure 1. The model structure and the interactions among the major modules (1: hydrological part; 2: water quality part; 3: ecological part; 4: dam regulation part; 5: PAT).

Figure 2. The flowchart of HCM and the interactions with other modules**.**

Figure 3. The flowchart of SBM (a) and CGM (b) in the ecological part and the interactions with other modules.

Figure 4. The flowchart of SEM (a), OQM (b) and WQM (c) in the water quality part and the interactions with other modules.

Figure 5. The flowchart of PAT and its interactions with other modules.

Figure 6. The location of study area (a) and the digital delineation of sub-basin, point source pollutant outlets, rural population (b), animal stock (c) and fertilization (d).

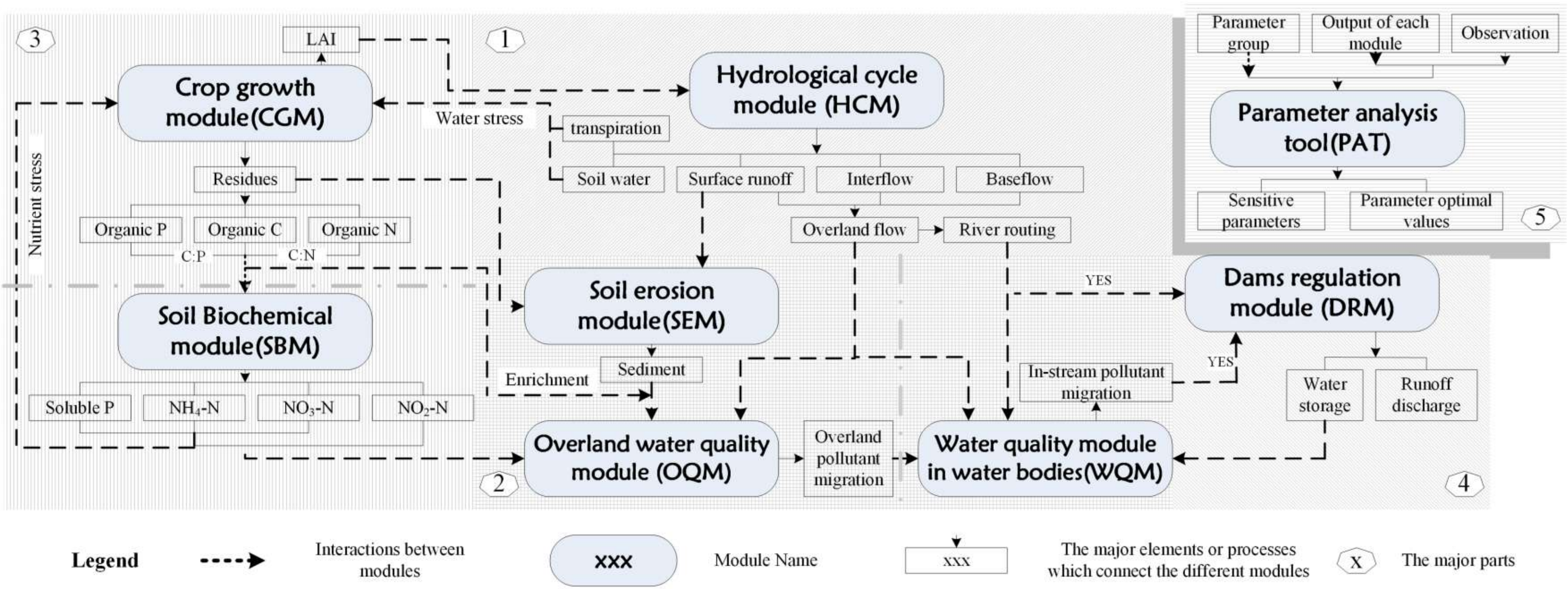
Figure 7. The daily runoff simulation at all stations.

Figure 8. The cumulative distributions of simulated and observed daily runoff at all stations

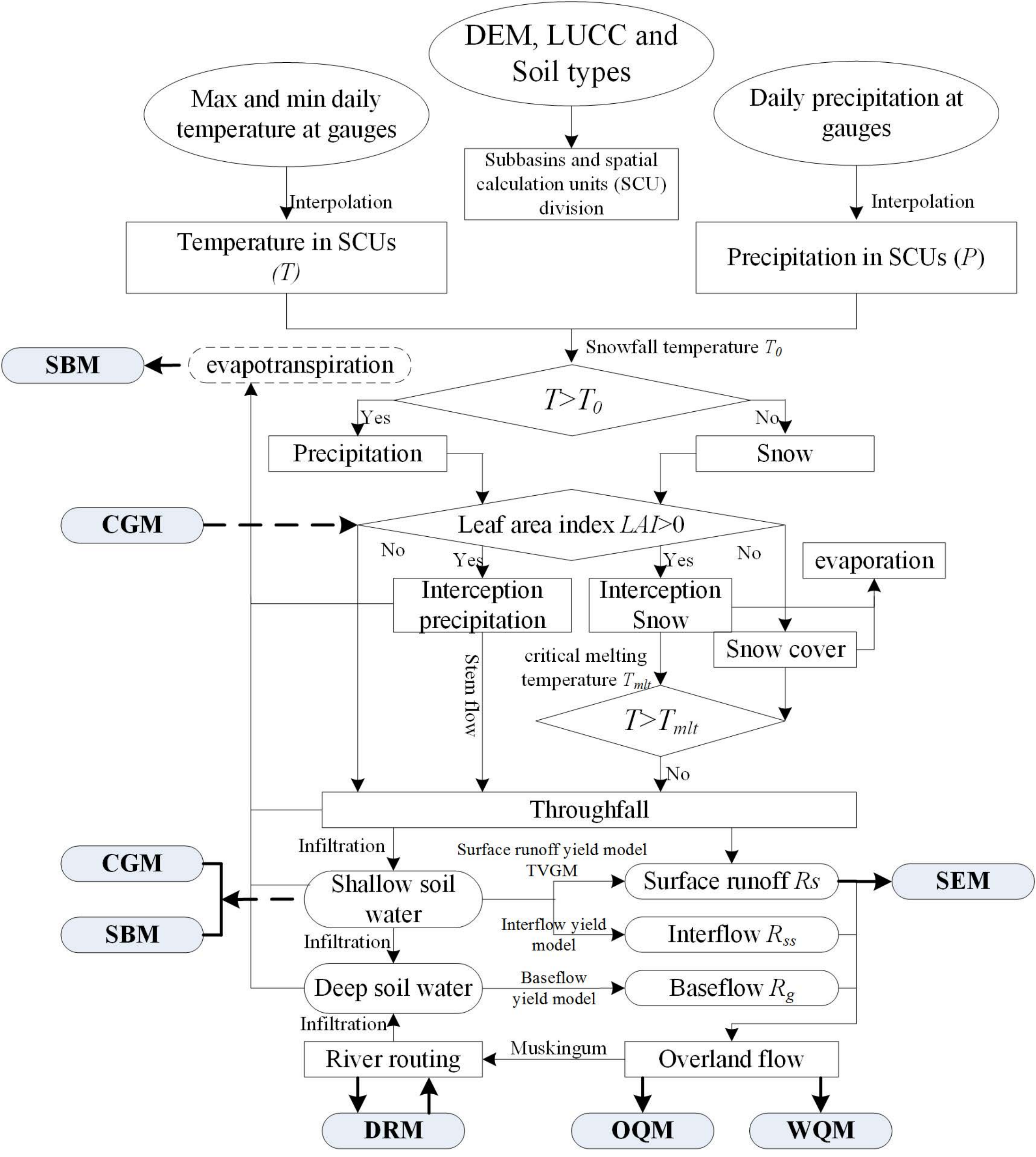
Figure 9. The simulated NH_4-N concentration variation at all stations.

Figure 10. The spatial pattern of diffuse source NH_4-N load (a) and its relationship with paddy area (b) and rice yield (c) at the sub-basin and regional scale in the Shaying River Catchment.

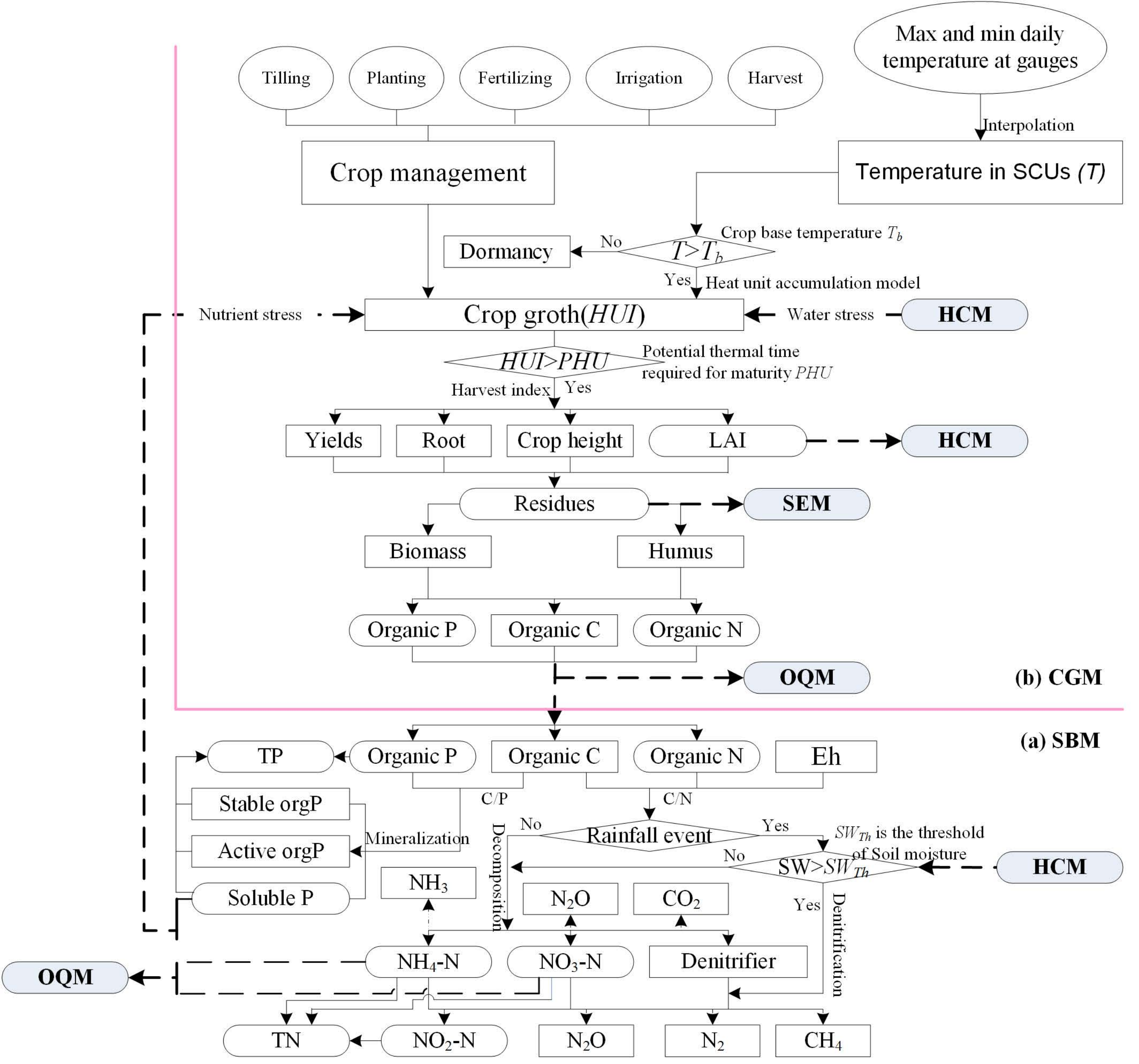
Figure 11. The spatial pattern of corn yield at the sub-basin and regional scale in the Shaying River Catchment.

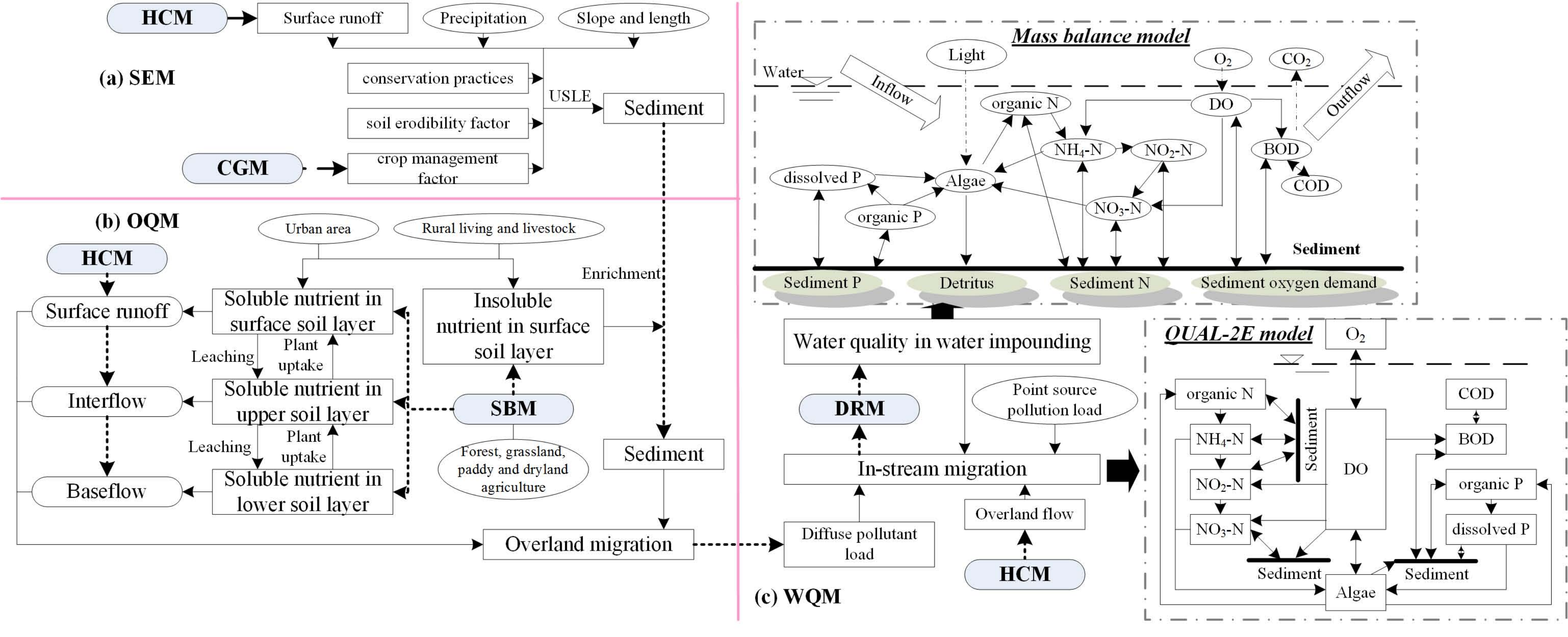


3
LAI
1
Crop growth module(CGM)
Hydrological cycle module (HCM)
Water stress
transpiration
Nutrient stress
Residues
Soil water
Surface runoff
Interflow
Baseflow
Organic P
Organic C
Organic N
Overland flow
River routing
C:P
C:N
Soil erosion module(SEM)
YES
Soil Biochemical module(SBM)
Dams regulation module (DRM)
Enrichment
Sediment
In-stream pollutant migration
YES
Water storage
Runoff discharge
Soluble P
NH_4-N
NO_3-N
NO_2-N
Overland water quality module (OQM)
Overland pollutant migration
Water quality module in water bodies(WQM)
2
4
Parameter group
Output of each module
Observation
Parameter analysis tool(PAT)
Sensitive parameters
Parameter optimal values
5
Legend
Interactions between modules
xxx
Module Name
xxx
The major elements or processes which connect the different modules
X
The major parts

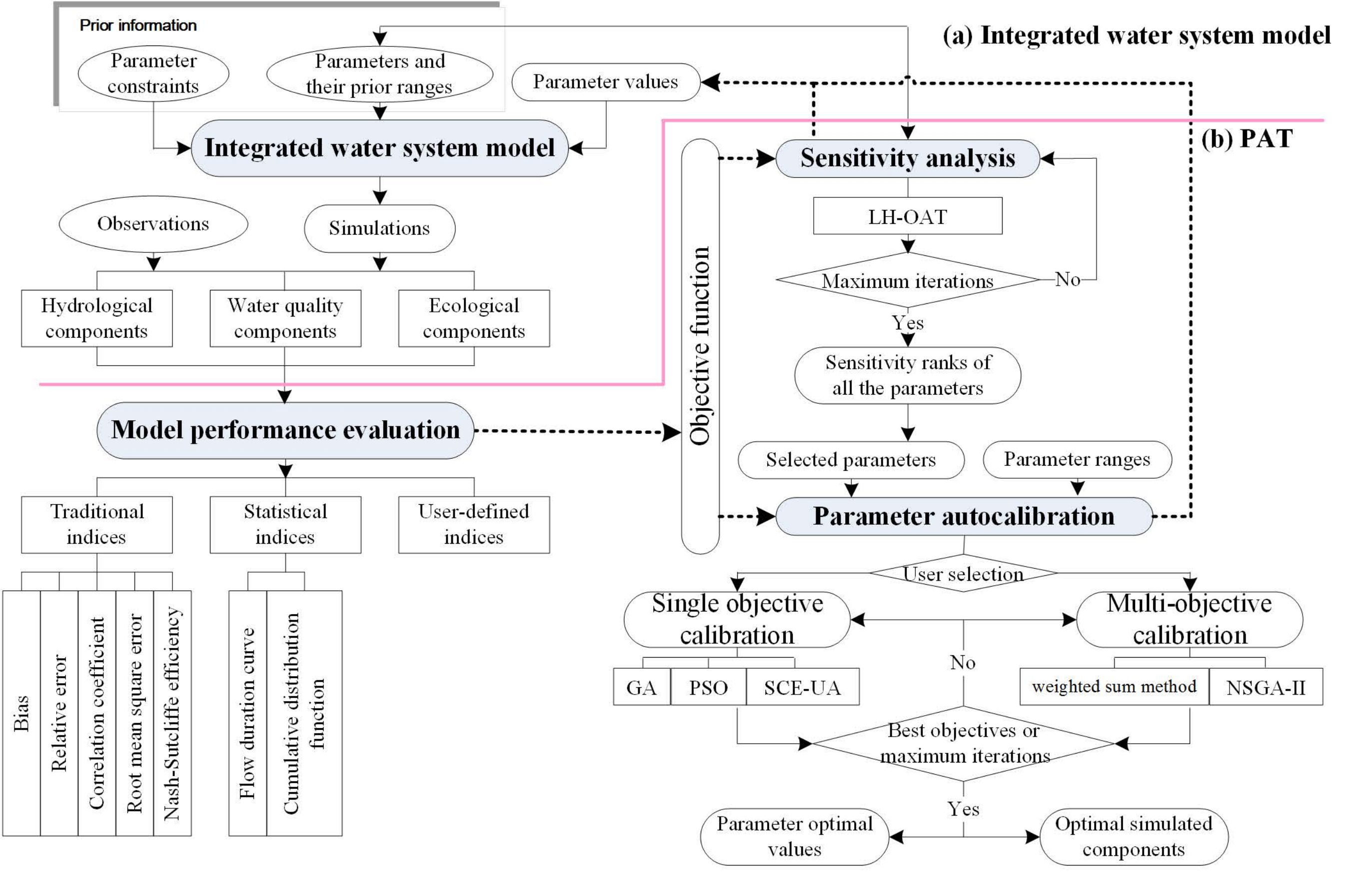


DEM, LUCC and Soil types
Max and min daily temperature at gauges
Daily precipitation at gauges
Subbasins and spatial calculation units (SCU) division
Interpolation
Interpolation
Temperature in SCUs (T)
Precipitation in SCUs (P)
Snowfall temperature T0
SBM
evapotranspiration
T>T0
Yes
No
Precipitation
Snow
CGM
Leaf area index LAI>0
No
Yes
Yes
No
evaporation
Interception precipitation
Interception Snow
Snow cover
Stem flow
critical melting temperature Tmlt
T>Tmlt
No
Throughfall
Infiltration
Surface runoff yield model TVGM
CGM
Shallow soil water
Surface runoff Rs
SEM
SBM
Interflow yield model
Interflow Rss
Infiltration
Deep soil water
Baseflow yield model
Baseflow Rg
Infiltration
River routing
Muskingum
Overland flow
DRM
OQM
WQM

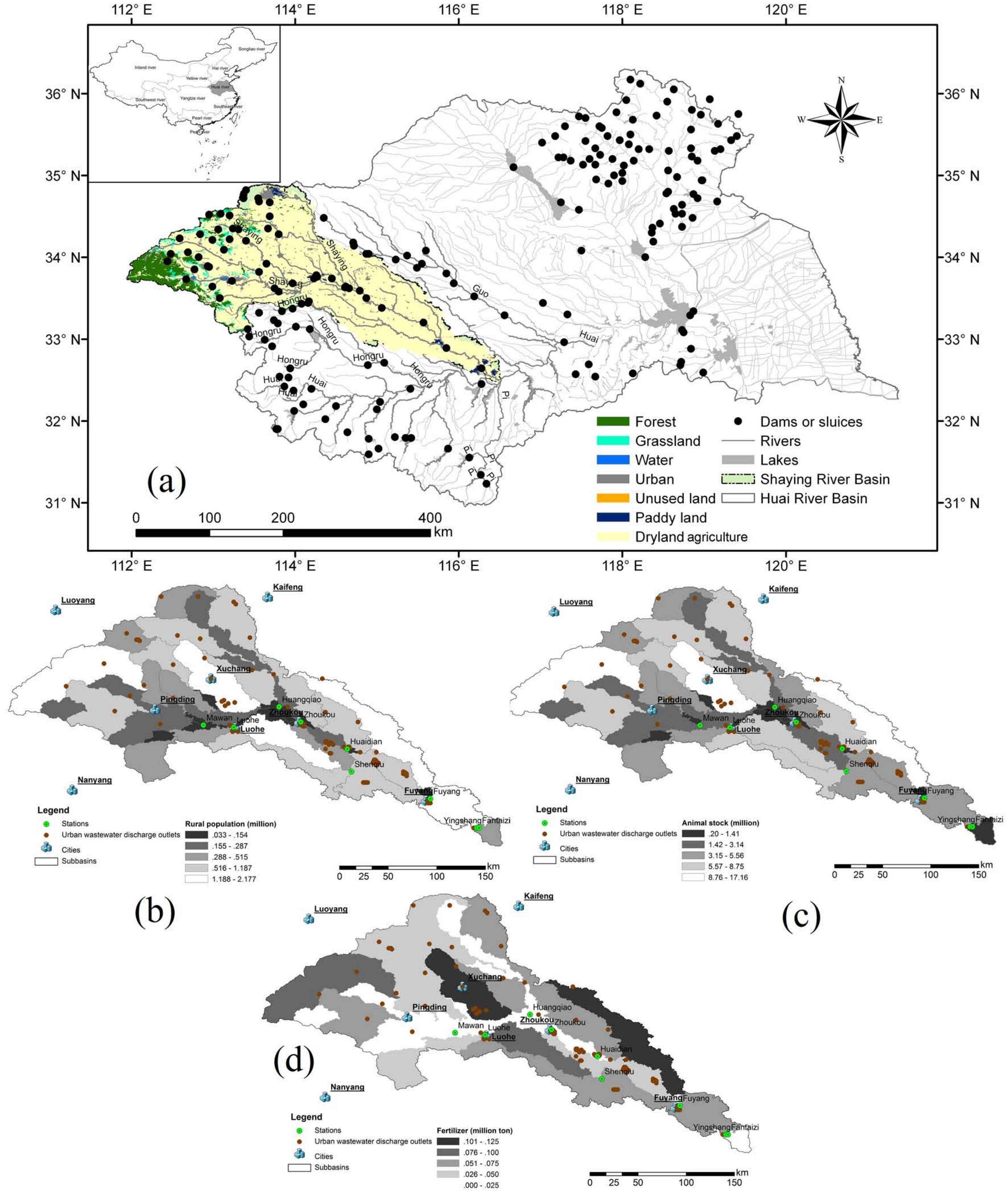

Max and min daily temperature at gauges
Tilling
Planting
Fertilizing
Irrigation
Harvest
Interpolation
Crop management
Temperature in SCUs (T)
Crop base temperature T_b
Dormancy
No
$T>T_b$
Yes
Heat unit accumulation model
Nutrient stress
Crop groth(HUI)
Water stress
HCM
$HUI>PHU$
Potential thermal time
required for maturity PHU
Harvest index
Yes
Yields
Root
Crop height
LAI
HCM
Residues
SEM
Biomass
Humus
Organic P
Organic C
Organic N
OQM
(b) CGM
(a) SBM
TP
Organic P
Organic C
Organic N
Eh
Stable orgP
C/P
C/N
No
Yes
Rainfall event
SW_{Th} is the threshold
of Soil moisture
Active orgP
Mineralization
Decomposition
No
$SW>SW_{Th}$
HCM
NH_3
Soluble P
N_2O
CO_2
Yes
Denitrification
OQM
NH_4-N
NO_3-N
Denitrifier
TN
NO_2-N
N_2O
N_2
CH_4



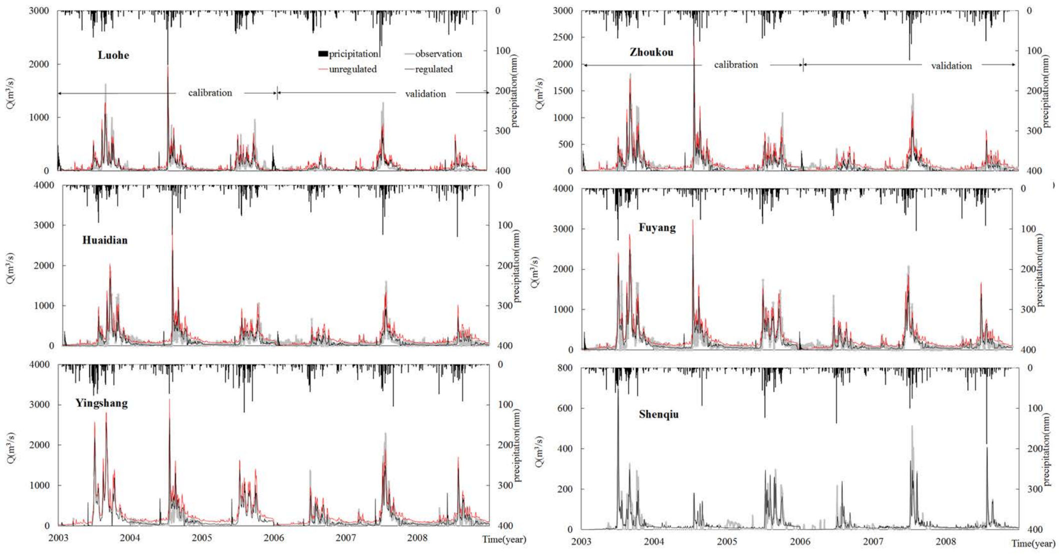
(a) SEM
HCM
Surface runoff
Precipitation
Slope and length
conservation practices
soil erodibility factor
CGM
crop management factor
USLE
Sediment
(b) OQM
Urban area
Rural living and livestock
HCM
Surface runoff
Interflow
Baseflow
Soluble nutrient in surface soil layer
Soluble nutrient in upper soil layer
Soluble nutrient in lower soil layer
Leaching
Plant uptake
Insoluble nutrient in surface soil layer
Enrichment
SBM
Forest, grassland, paddy and dryland agriculture
Sediment
Overland migration
(c) WQM
Mass balance model
Water
Inflow
Light
Outflow
O_2
CO_2
organic N
DO
NH_4-N
NO_2-N
BOD
COD
dissolved P
Algae
organic P
NO_3-N
Sediment
Sediment P
Detritus
Sediment N
Sediment oxygen demand
Water quality in water impounding
DRM
Point source pollution load
In-stream migration
Overland flow
Diffuse pollutant load
HCM
QUAL-2E model
organic N
NH_4-N
NO_2-N
NO_3-N
Sediment
DO
COD
BOD
organic P
dissolved P
Algae



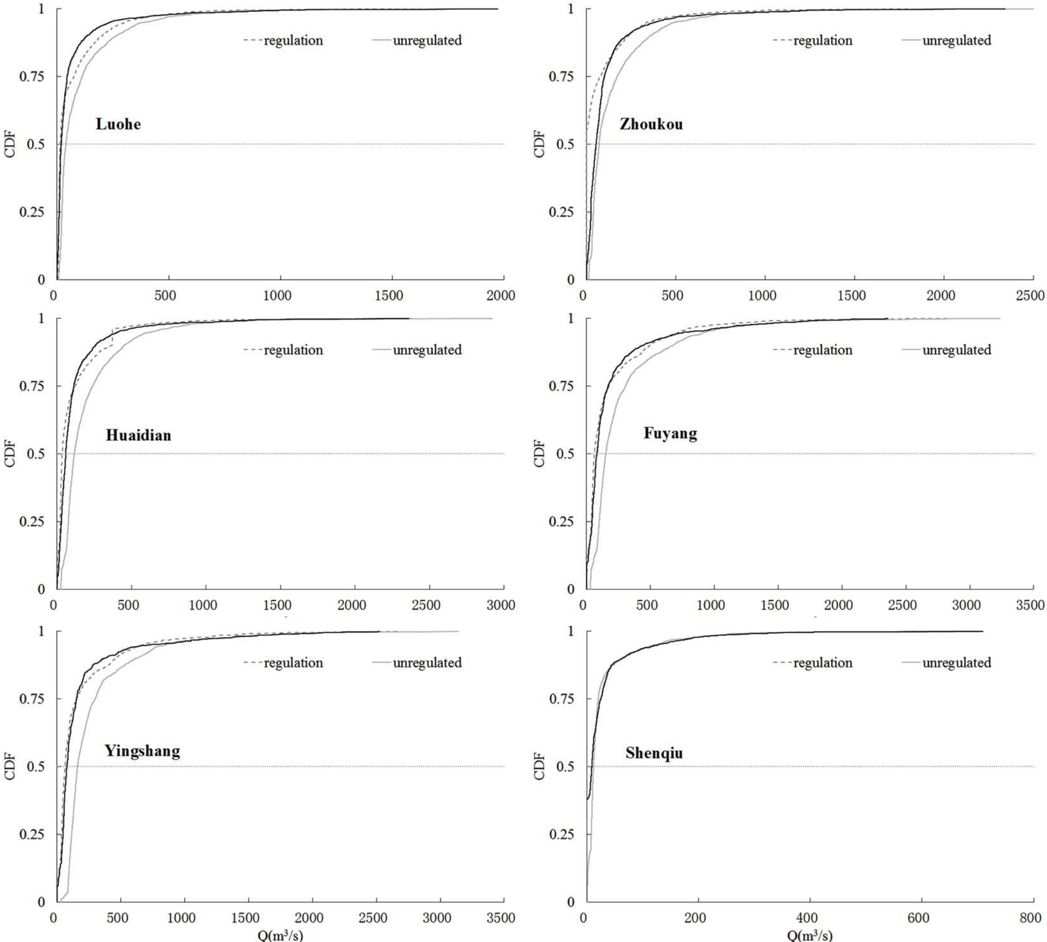
(a) Integrated water system model
Prior information
Parameter constraints
Parameters and their prior ranges
Parameter values
Integrated water system model
Observations
Simulations
Hydrological components
Water quality components
Ecological components
Model performance evaluation
Traditional indices
Statistical indices
User-defined indices
Bias
Relative error
Correlation coefficient
Root mean square error
Nash-Sutcliffe efficiency
Flow duration curve
Cumulative distribution function
(b) PAT
Objective function
Sensitivity analysis
LH-OAT
Maximum iterations
No
Yes
Sensitivity ranks of all the parameters
Selected parameters
Parameter ranges
Parameter autocalibration
User selection
Single objective calibration
Multi-objective calibration
GA
PSO
SCE-UA
No
weighted sum method
NSGA-II
Best objectives or maximum iterations
Yes
Parameter optimal values
Optimal simulated components



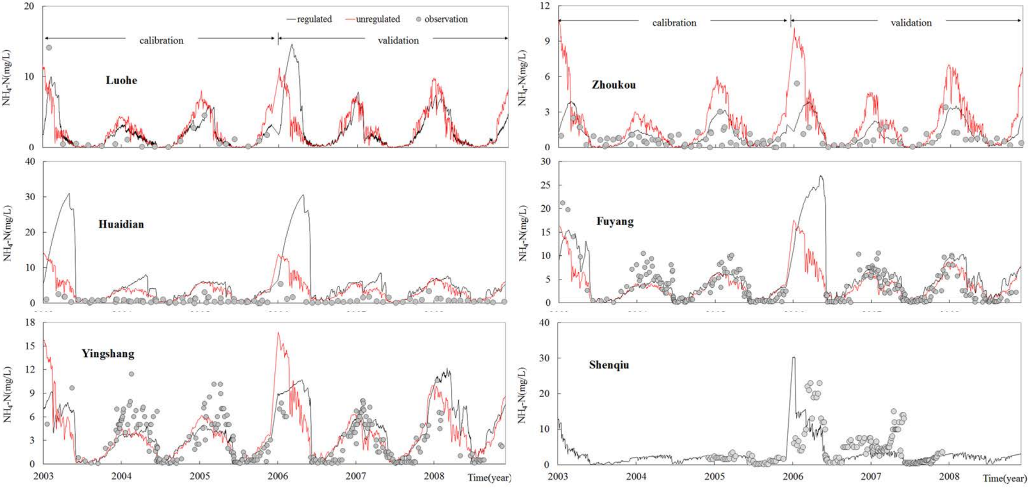
112° E
114° E
116° E
118° E
120° E
36° N
35° N
34° N
33° N
32° N
31° N
Songliao river
Inland river
Hai river
Yellow river
Huai river
Yangtze river
Southwest river
Southeast river
Pearl river
N
S
W
E
Shaying
Shaying
Shaying
Hongru
Hongru
Hongru
Hongru
Hongru
Hongru
Huai
Huai
Huai
Huai
Guo
Pi
Pi
Pi
Pi
Pi
Forest
Grassland
Water
Urban
Unused land
Paddy land
Dryland agriculture
Dams or sluices
Rivers
Lakes
Shaying River Basin
Huai River Basin
(a)
0
100
200
400
km
Luoyang
Kaifeng
Xuchang
Pingding
Huangqiao
Zhoukou
Zhoukou
Mawan
Luohe
Luohe
Huaidian
Shenqiu
Nanyang
Fuyang
Fuyang
Yingshang
Fantaizi
Legend
Stations
Urban wastewater discharge outlets
Cities
Subbasins
Rural population (million)
.033 - .154
.155 - .287
.288 - .515
.516 - 1.187
1.188 - 2.177
km
0
25
50
100
150
(b)
Animal stock (million)
.20 - 1.41
1.42 - 3.14
3.15 - 5.56
5.57 - 8.75
8.76 - 17.16
(c)
(d)
Fertilizer (million ton)
.101 - .125
.076 - .100
.051 - .075
.026 - .050
.000 - .025



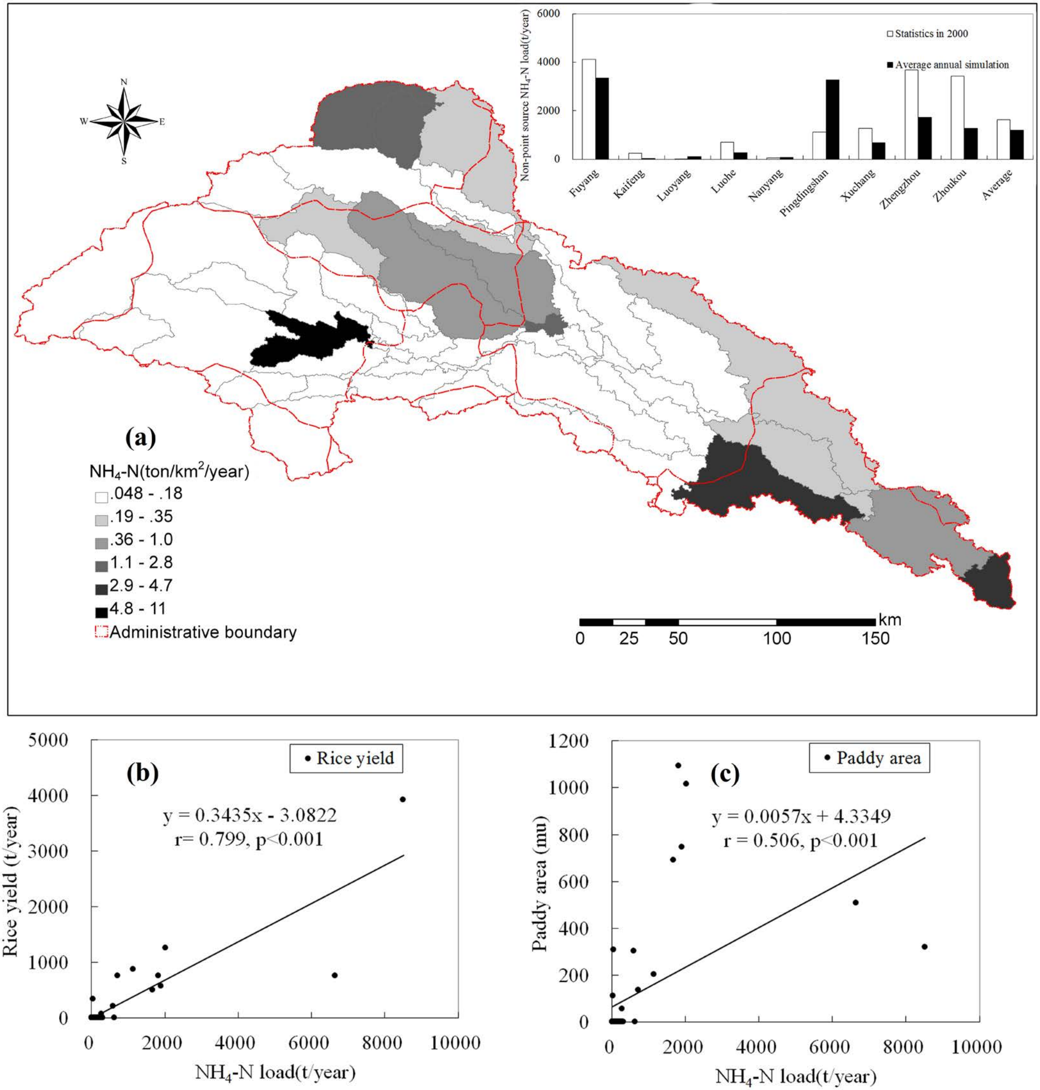
Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
Shenqiu
pricipitation
observation
unregulated
regulated
calibration
validation
Q(m³/s)
precipitation(mm)
Time(year)
2003
2004
2005
2006
2007
2008



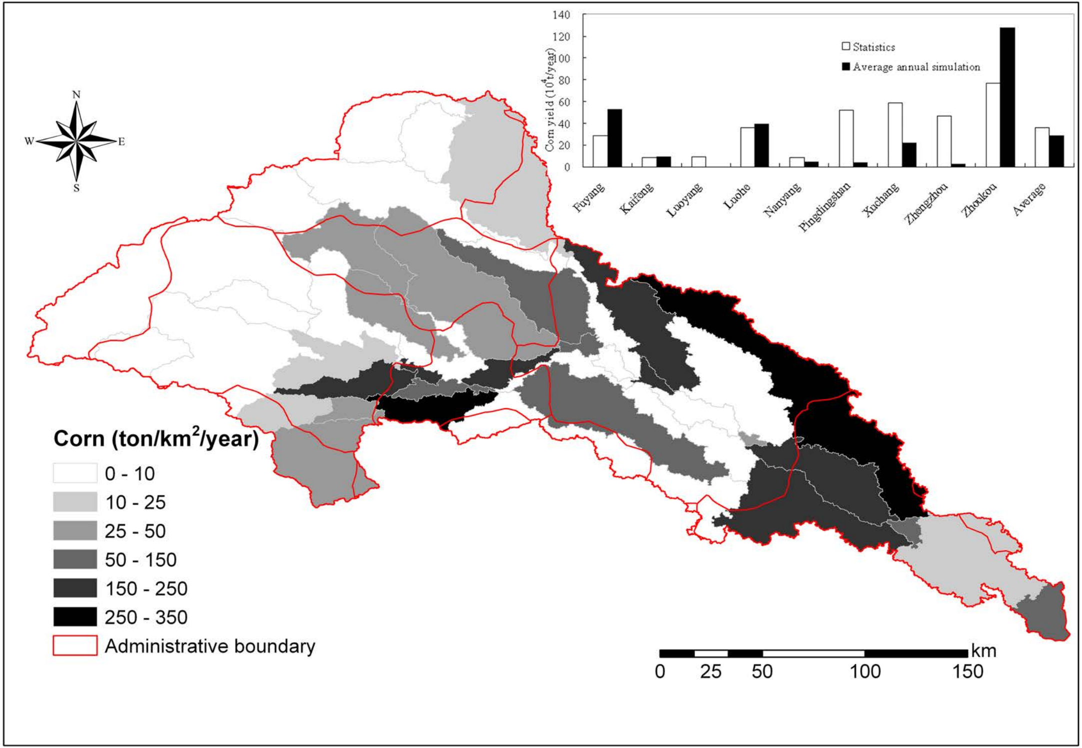
Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
Shenqiu
regulation
unregulated
CDF
Q(m³/s)



regulated
unregulated
observation
calibration
validation
Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
Shenqiu
NH_4-N(mg/L)
2003
2004
2005
2006
2007
2008
Time(year)



(a)
NH4-N(ton/km2/year)
.048 - .18
.19 - .35
.36 - 1.0
1.1 - 2.8
2.9 - 4.7
4.8 - 11
Administrative boundary
km
0 25 50 100 150
N
S
W
E
Non-point source NH4-N load(t/year)
Statistics in 2000
Average annual simulation
Fuyang
Kaifeng
Luoyang
Luohe
Nanyang
Pingdingshan
Xuchang
Zhengzhou
Zhoukou
Average
(b)
Rice yield
y = 0.3435x - 3.0822
r= 0.799, p<0.001
Rice yield (t/year)
NH4-N load(t/year)
(c)
Paddy area
y = 0.0057x + 4.3349
r = 0.506, p<0.001
Paddy area (mu)
NH4-N load(t/year)



140
120
100
80
60
40
20
0
Corn yield (10^4t/year)
Statistics
Average annual simulation
Fuyang
Kaifeng
Luoyang
Luohe
Nanyang
Pingdingshan
Xuchang
Zhengzhou
Zhoukou
Average
N
S
W
E
Corn (ton/km^2/year)
0 - 10
10 - 25
25 - 50
50 - 150
150 - 250
250 - 350
Administrative boundary
km
0
25
50
100
150