Comments from the editor:

Thank you very much for your detailed responses to the very helpful reviewer comments. As you have noted there are several major concerns that have been raised and that need to be addressed in detail in the revised version of the manuscript:

Response: Thanks very much for your careful review and suggestions. We revised the manuscript carefully by following the comments from the editor and all the reviewers. All the changes were highlighted in blue color. Acknowledgements were added in the revision.

(1) As pointed by all reviewers, and I fully agree, a model with >100 parameters does need some careful considerations and efficient methods to effectively constrain the feasible parameter space. In principal I commend you for you approach to identify sensitive parameters and to only calibrate these parameters for each model component. However, I am not convinced that this will be sufficient to robustly identify parameter sets that provide plausible model internal dynamics. I am surprised that you did neither (1) follow a more consistent multi-objective and multi-criteria calibration strategy nor (2) apply some kind of regularization or stronger prior information (e.g. relational parameter or process constraints). Making use of several orthogonal (!!) objective functions has in the past proven very valuable to identify unfeasible parameter sets. Similarly calibration against multiple variables/signatures (for the hydro module e.g. flow, flow duration curve, autocorrelation function of flow, etc.) has recently received considerable attention. To move away from pure curve fitting, hydrology and related sciences need to start to move away from curve fitting towards more process consistency and efficient model selection techniques. I would therefore encourage you to dig a bit into the recent literature and amend you calibration strategy accordingly, so as to ensure an efficiently constrained model.

Response: Thanks very much for your careful review and suggestions. Yes indeed, carful considerations must be taken for models with a large number of parameters when only limited observations are available for model calibration. In fact, how to calibrate complex model is still a hot topic due to possible equifinality and internal competition among different processes.

The parameter analysis tool (PAT) was revised to provide more details as follows. (1) PAT is designed for model performance evaluation, parameter sensitivity analysis and autocalibration (Figure 5). This module is independent of the integrated water system model. As there are many different ways to calibrate the model (different objective functions and different optimization techniques), we cannot include all possibilities in the PAT module. However, we believe that a PAT module is still useful for users who want to use only the common objective functions without any program coding. The users who want other objective functions or different calibration techniques can write the codes to extend the function of PAT which interact with the integrated model by simulating all the processes for each given set of parameter values. (2) In the current

PAT module, five widely used traditional indices are included (i.e., *bias*, relative error *re*, root mean square error *RMSE*, correlation coefficient *r* and coefficient of efficiency *NS*) and their detailed equations are given in the Appendix D. Furthermore, the flow duration curve and cumulative distribution function which are usually adopted to capture multiple signatures of calibrated processes are also produced. An objective function is formed by considering single or multiple indices of simulated components, and can be selected by users on the basis of their specific requirements. (3) The parameter analysis algorithms 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).

In order to obtain sensible parameter values, the following treatments are adopted in the PAT module. 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 if exist. The constraints on parameters are also considered in both parameter sensitive analysis and autocalibration. In particular, 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 types are set in ascending order (that is, 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 water impounding, the settling rates of water quality variables (K_{set}) are greater than the resuspension rates (K_{scu}) and the settling rates in channels (R_{set}). 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 based on the best of our knowledges by referring the literatures (e.g., SWAT, EPIC, and DNDC). The multi-objective and multi-criteria calibration methods are also included in PAT to support our further study (see the manuscript below).

As the main objectives of this paper are to present the integrated water system model and demonstrate its applicability in complex basins, we used a simply weighted sum method to handle multi-criteria in model autocalibration, and also used the cumulative distribution function (similar to flow duration curve) to evaluate the simulation performance in high and low flows. The complicated multi-objective calibration and multi-criteria methods were also conducted to capture multiple signatures and the results were submitted for possible publications elsewhere; see our manuscripts below:

- Zhang Yongyong, Shao Quanxi, Taylor John A.: A balanced calibration of water quantity and quality by multi-objective optimization for integrated water system model. (under review)
- Zhang Yongyong, Shao Quanxi, Zhang Shifeng, Zhai Xiaoyan, She Dunxian: Multi-metric calibration of hydrologic model to capture overall flow regimes in the arid

Hexi Corridor of northwest China. (under review)

The following paragraphs were added in the section of 2.1.5 Parameter analysis tool (PAT) in the revision to address the comments and suggestions (P11 L14-P12 L33):

"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. 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 to help users in the use of our proposed model. It contains 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 coefficient of efficiency (*NS*). 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 by users on the basis of their specific requirements.

In order to obtain 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 types are set in ascending order (from water body, paddy land, urban area, forest, dryland agriculture, unused land to 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 in channels (R_{set}). 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 based on the best of our knowledges by referring the literatures (e.g., SWAT, EPIC, and DNDC) or 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."

Moreover, we really appreciated your comments about the different model calibration. Some related studies were discussed in the section of discussion (Sub-section 4.2 Equifinality) (see P23 L10-17). "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). "

(2) One reviewer also pointed out that it may be over-ambitious to publish such a comprehensive model together with a case study in one paper. In principal I agree with that, as in the original manuscript the descriptions of the model but also of the calibration strategy remained rather superficial. If you prefer to keep the entire manuscript in one, that is fine, but in the sense of reproducibility and dissemination of your model I would strongly encourage you to provide more detailed model descriptions (incl. the full set of equations) in the supplementary material.

Response: Thanks very much for your suggestions. Yes we prefer to keep the entire manuscript in one as we explain in the responses to the reviewers. Following your suggestion, we provided the detailed model descriptions in the supplementary material (See equations S7, S8, S30-S32, S34, S35-47).

(3) The manuscript lacks an adequate discussion so far. How does the model compare to similar models? What are its limitations? What are its advantages? Where should future improvements be necessary?

Response: Thanks very much for your comments. A discussion section was added in the revision, including comparison with other models (P22 L1-25), equifinality (P22 L27-P23 L28) and model limitations (P 23 L30-P24 L17). The advantages of our model were also presented in P22 L2-7, and possible future improvements were discussed in the last paragraph of the revised manuscript (See P25 L3-12).

(4) It is not clear what the scientific objective of the manuscript is. What are the hypothesis you are intending to test? Please explicitly state this at the end of the introduction section

Response: Thanks very much for your comments. Sorry for not summarizing the purpose of our manuscript at the end of the introduction section.

Yes, we found that we only discussed the importance of building integrated water models and reviewed the current related modelling practice. In fact, the motivation of this piece of research is to simultaneously improve water quantity and quality simulation. The board research background of our collaborative team (including hydrology, water quality soil biogeochemistry, and agriculture ecology) makes us to be interested in the interaction mechanisms and linkages in a broad sense by including soil nutrient, crop and so on, rather than just confine to the water quantity and quality. To achieve those goals, we built integrated model by extending a mathematically based hydrological model DTVGM through coupling the detailed descriptions of the soil biogeochemical processes and ecological processes, and considering the prevalent regulations by water projects, because DTVGM performs well in many basins (particularly in China).

By following your suggestion, we added the following paragraphs at the end of introduction section (See P4 L24-2 P5 L17) in the revision:

“In this study, we tend 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 hydrological data 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. DTVGM is currently used in many basins with different scales and climate zones to investigate the effect of human activities and climate change on runoff, and shows good simulation performances (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 is 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 includes popular objective functions, autocalibration approaches and summary statistics, is also developed. To demonstrate the model performances, we simulate 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.”

(5) The manuscript (also the newer version attached to the reviewer responses) requires detailed proof reading by a native speaker. Grammar and spelling errors are plenty.

Response: Thanks for your careful review and sorry for the grammar and spelling errors. During the revision, we tried hard to correct as many errors as possible and then asked a colleague (who is a native English speaker) to check the grammar and spelling. Finally, we paid a professorial editorial service to check and correct the errors. Hopefully, the grammar and spelling errors in the revision are minimal. All the revisions were highlighted by blue color in the manuscript.

Comments from anonymous Referee #1

GENERAL COMMENTS:

Thank you for the opportunity to read an interesting paper about a new integrated water and water quality model. Although this type of model is not entirely new, it adds an alternative to existing model formulations, particularly the inclusion of carbon and crop growth modules, and has indeed a number of novel process considerations. In particular, it is interesting to see the simulation of regulation in a water quality context.

Response: Thanks very much for your careful review and suggestions. We revised the manuscript carefully by following the comments from all the reviewers. All the changes were highlighted in blue color. Acknowledgements were added in the revision.

I do, however, have concerns with how the model calibration was performed and how the model evaluation is presented. The paper goes to great lengths to explain the model setup and which processes are described, but given the manner in which the automated calibration was used, can you really ensure that processes are reproduced in the correct balance?

Response: Thanks very much for your comments. We agree that it was unbalance between the model framework (Section 2.1) and parameter analysis (Section 2.2) in the original manuscript. Actually, the parameter sensitivity analysis and autocalibration are important steps in the model applications. Given that the model contains several modules which make the model framework lengthy and that PAT can be treated as a module, we presented the context of PAT as a part of the proposed model and re-numbered its section as section 2.1.5. The detailed descriptions of PAT were introduced from P11 L6 to P12 L33.

As the PAT is not new but necessary for the completion of model development, we slightly change the title of this manuscript by adding word “with”, that is, the new title became “Integrated water system simulation by considering hydrological and biogeochemical processes: model development, with parameter sensitivity and autocalibration” (See P1 L1-4). We found that the section (2.1.2.1) of soil biochemical module presented the decomposition, nitrification and denitrification processes of N in too detail. It seemed to be also unbalance with other sections. We reorganized this section (See P7 L14-P8 L18).

I also miss a scientific question that the paper addresses, a discussion of whether the study achieves these aims and a clear conclusion regarding this.

Response: Thanks very much for your comments. The main scientific question is to improve the modelling practice of water related components by integrating hydrological, biogeochemical, water quality and ecological processes, and considering the prevalent regulations by water projects. The integrated water system modelling has been a popular scientific research topic to simultaneously simulate different water related processes. However, it is often that a model usually performs well for the oriented processes, and only approximate results for other processes outside of the model’s focus. In particular, the applications of some typical models were not well in our case study area (See P22 L8- L25 in the discussion section). Furthermore, our collaborative team (including hydrology, water quality soil biogeochemistry, and agriculture ecology) is also interested in the interaction mechanisms and linkages in a broad sense by including soil nutrient, crop and so on, rather than just confines to the water quantity and quality. We also adopted a mathematics based hydrological model DTVGM as it has had good performances in many basins (particularly in China). All of these considerations motivated us to extend DTVGM to an integrated water system model by coupling the detailed descriptions of the soil biogeochemical processes and ecological processes, and considering the prevalent regulations by water projects.

The scientific questions and motivation were clearly stated in the revision (See P4 L22-P5 L17):

“In this study, we tend 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 hydrological data 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. DTVGM is currently used in many basins with different scales and climate zones to investigate the effect of human activities and climate change on runoff, and shows good simulation performances (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 is 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 includes popular objective functions, autocalibration approaches and summary statistics, is also developed. To demonstrate the model performances, we simulate 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."

Following the next comment, we also added a discussion section in the revision to present the comparison with other models, model limitations and possible direction for further development, discussion on equifinality (See P22 L1-P24 L17).

The lack of discussion makes we wonder if the authors have considered any of the limitations in their approach or compared their approach to existing studies so that the reader might conclude whether this new approach provides any added value. Could the model be applied in ungauged catchments, other regions easily?

Response: Thanks very much for your comments and suggestions. A discussion section was added in the revision, including comparison with other models, model limitations and equifinality (See P22 L1-P24 L17). We are not sure if this model is applicable in ungauged catchments easily because the choice of parameter values and availability of input data are the keys in the studies of ungauged basin. It would be an interesting research topic but we prefer not to discuss this in this manuscript, in order not to confuse the readers.

What scales is the model suitable for?

Response: Thanks very much for your comments. In this study, we did not specify the exact scale ranges in the model setting. Three levels of spatial calculation units were designed in the model, i.e. sub-basin unit, land-use unit and crop unit from largest to smallest (See P13 L6-29). In our case study, the areas of sub-basin, land-use and crop units ranged from 46.48 to 3771.15 km^2, from 0.04 to 2762.5 km^2, and from 3.73 to 2762.5 km^2, respectively (See P15 L30-32).

We believe that the model performance at different scale ranges will be case by case and reasonable scale ranges would be established after a large number of case studies. As the first paper to introduce the model, we do not made any recommendation on the scales. Due to the continuity of our research and applications, more cases will be conducted in the future.

I commend the authors in testing the model in such a heavily polluted and regulated area for which it was probably difficult to describe inputs for; however limitations, such as accurately describing input data, are not mentioned at all in the paper.

Response: Thanks very much for your careful review and sorry that the model inputs, particularly for the case study, were not clearly stated in the main manuscript although the related data sets for model setup and calibration were given in Table 1 and the detailed information of input data for our case study in Table S2-3 in the supplement material. In the revision, we added some statements and quick summary for the readers to follow the model setting easily (See P15 L15-17; P S16-S17). Moreover, in order to keep the consistent of input data descriptions, we moved the table 2 to table S3 in the supplement material (See P S17).

SPECIFIC COMMENTS

1. I have many technical comments regarding the use of citations in the paper (see below). Please consider carefully whether or not a citation is needed and please don't cite recent papers for old, accepted knowledge.

Response: Thanks very much for your careful review and suggestions. We checked the manuscript carefully. Some citations of recent papers for old, accepted knowledge were removed from the revision (See P2 L17-19). We also updated the references by following your suggestions in other comments (P31 L25-P32 L3; P32 L11, L29-30; P33 L13-18; P34 L 1-3, L6-9, L13-15, L26-28; P35 L4-9; P36 L19-21, L25-28; P37 L3-8).

2. Introduction: The paper lacks a scientific question. I deduce from the introduction that you would like to say that integrated models can improve simulation of the integrated factors, but do you actually show this? Have others shown this?

Response: Thanks very much for your comments.

The main scientific question is to improve the modelling practice of water related components by integrating hydrological, biogeochemical, water quality and ecological processes, and considering the prevalent regulations by water projects. The integrated water system modelling has been a popular scientific research topic to simultaneously simulate different water related processes. However, it is often that a model usually performs well for the oriented processes, and only approximate results for other processes outside of the model's focus. In particular, the applications of some typical models were not well in our case study area (See P22 L8- L25 in the discussion section). Furthermore, our collaborative team (including hydrology, water quality soil biogeochemistry, and agriculture ecology) is also interested in the interaction mechanisms and linkages in a broad sense by including soil nutrient, crop and so on, rather than just confines to the water quantity and quality. We also adopted a mathematics based hydrological model DTVGM as it has had good performances in many basins (particularly in China). All of these considerations motivated us to extend DTVGM to an integrated water system model by coupling the detailed descriptions of the soil biogeochemical processes and ecological processes, and considering the prevalent regulations by water projects.

The scientific questions and motivation were clearly stated in the revision (See P4 L18-20; P5 L21-P6 L2). "In this study, we tend 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 hydrological data 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. DTVGM is currently used in many basins with different scales and climate zones to investigate the effect of human activities and climate change on runoff, and shows good simulation performances (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 is 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 includes popular objective functions, autocalibration approaches and summary statistics, is also developed. To demonstrate the model performances, we simulate 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."

3. On pp 5001, lines 2-10, you mention a number of relevant theories, but can you show that including more and more processes actually improves model results? I'm not sure it does. Given the concerns I have about the calibration process, isn't it likely you end up with an overparameterised model with insufficient data to drive and test such a model? Please discuss.

Response: Thanks very much for your comments. We fully understand your concerns. In fact, the following thoughts guided our model development, Firstly, the integrated water system model is a popular tendency to extend individual process models to other fields in order to deal with multidisciplinary research topics. Given that the hydrological cycle has been widely accepted as a critical linkage among physical (e.g., runoff), biogeochemical (e.g., nutrient, water quality) and ecological processes (e.g., plant growth), energy fluxes at the basin scale (Wigmosta et al., 1994; Singh and Woolhiser, 2002; Burt and Pinay 2005), we extend DTVGM, a mathematical based hydrological model to an integrated water system model (See P2 L30- P3 5; P4 L24-25). Secondly, we did not arbitrarily collect processes in the integration. The mentioned mature theories provided the scientific foundations for the development of the integrated water system model (See P3 L5-14).

Indeed, a large number of parameters was an inevitable issue in the integrated water system model. It was also impossible to well-calibrate all the related processes according the insufficient data. In practice, the parameter sensitivity analysis would be an effective way to reduce the dimensionality in model calibration and focus only on the critical processes and their parameters which are most sensitive to the model outputs (See P22 L27- P23 L28). We develop a parameter analysis tool which includes sensitivity analysis, autocalibration and model performance evaluation to improve modelling efficiency. The detailed descriptions are introduced from P11 L6 to P12 L33. Moreover, we collect several observations of representative water related components (e.g., runoff and water quality observations at different stations, the diffuse pollutant load and crop yield data), to calibrate each subsystem (See P23 L20-23).

The equifinality, model limitations and the comparison with other models are discussed in the new discussion section (See P22 L1-P24 L17).

4. pp5001: What about DAISY, SOILNP, ICECREAM models? It should also be mentioned that there is a tradition of coupling traditional rainfall-runoff models with field scale nutrient models, e.g. Arheimer and Brandt 2000 and 1998

Response: Thanks very much for your comments. These models were discussed in the literature review (See P4 L5-7; P31 L25- P32 L5; P34 L13-15; P37 L3-5).

5. pp5002: What about SWIM and HYPE which are also integrated WQ/Q models?

Response: Thanks very much for your comments. Yes indded, both SWIM and HYPE are the integrated WQ/Q model. SWIM is an improved version of SWAT and HYPE is developed based on a famous hydrological model (HBV) for hydrological and water quality simulation. Thus, it is better to categorize HYPE into the hydrological based model. These models were mentioned in P3 L26; P4 L23; P32 L1-5; P34 L26-28; P 35 L4-6.

6. pp5011: Regarding dams: Maybe you could mention the different functionality of dams and note that methods you try to reproduce are those common for flood control or water supply dams? Hydropower dams are likely to show completely different behavior.

Response: Thanks very much for your comments. A sentence was added in this section, i.e., "Given that different types of dams and sluices are likely to show completely different behaviors of regulation, we try to reproduce the common functionalities for the flood control or water supply dams in this module." (See P10 L23-26).

7. pp5012, Lines 3-5 ': : :usually considered to take place at daily scale' – You use the

term ‘usually’, yet you refer to only 1 study to demonstrate that things are usually done this way? Actually, erosion, overland flow and phosphorous processes could very much be improved with subdaily time-scales. Lake turnover processes would be sufficient at coarser time-scales. Why do you really use a daily time-scale? Is it actually because this has traditionally been considered sufficient for rainfall-runoff modeling and consistent with data availability? I would write instead that it is practical to use a daily time-step as this is consistent with the underlying rainfall-runoff module. You should also discuss at what timescale you believe the results are realistic!

Response: Thanks very much for your comments and suggestion. The sentence was revised following your suggestions. “The sub-daily scale may improve the performance in some modules (e.g., SEM, 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” (See P13 L30-P14 L4).

8. Calibration: I am concerned that the calibration may in fact just be model tuning and not take into account the correct balance between processes and landuse types. Given the very large number of parameters, and the fact that some processes occur in the soil and some processes occur in surface water resulting in the same downstream concentration, how do you ensure calibrated parameters don’t compensate for each other and you don’t get the right answer for the wrong reasons? For example, is it possible that you overestimate retention in groundwater and underestimate retention in surface water? Are the relative contributions of different processes realistic (for example surface runoff, overland flow etc.)? This topic requires more explanation and discussion

Response: Thanks very much for your comments and suggestion. Like other hydrological, water quality or biogeochemistry models, high parameterization and equifinality are inevitable in the integrated water system model. However, in practice, the parameter sensitivity analysis would be effective to reduce the dimensionality in model calibration, and focus on the critical processes and their parameters which are most sensitive to the model outputs (See P22 L27- P23 L28). In this study, we develop a parameter analysis tool which includes sensitivity analysis, autocalibration and model performance evaluation to improve modelling efficiency. Several parameter constraints are also considered in both parameter sensitive analysis and autocalibration. The detailed descriptions are introduced from P11 L6 to P12 L33. We restrict SCE-UA to calibrate only the sensitive parameters as defined by LH-OAT, while other parameters remained constant which are given based on the best of our knowledges by referring the literatures or similar basins (See P16 L21-24). Moreover, we collect some observations of representative water related components (e.g., runoff and water quality observations at different stations, the diffuse pollutant load and crop yield data), to calibrate each subsystem (See P23 L20-23).

The comparison with other models, equifinality, and model limitations are discussed in the section of discussion (See P22 L1-P24 L17).

9. Table 6: It would be worth mentioning in the text the large change in bias for Nitrogen concentrations between calibration and validation at Fuyang and Yinshang. Why do you think this occurs? Could this be related to the limitations in calibration mentioned above?

Response: Thanks very much for your careful review and comments. The biases of runoff and water quality concentration simulations change in the opposite ways. For examples, the runoff underestimation usually causes the overestimation of water quality concentration, while the runoff overestimation usually causes the underestimation of water quality concentration. Therefore, the large changes in bias for NH_4-N concentrations between calibration and validation at Fuyang and Yingshang are probably caused by the changes in bias for the corresponding runoff simulations. However, the model performances of both calibration and validation are acceptable. We do not think the limitations in calibration is related to these large changes (See P19 L25 – 28).

10. How well is the input data described and how do uncertainties in the input data affect the calibration and evaluation of the model?

Response: Thanks very much for your comments. The input data are specified in Tables S2-S3 in the supplementary material. All the data sets for model setup and calibration are collected from the government bureaus, official books or scientific references (See P15 L15-17; P S16-S17). In this study, we do not consider the uncertainties in the input data. In fact, this is also a very important issue and hot topic in the research of hydrological or environmental modeling. Given that this topic itself can be a lengthy paper and is beyond the scope of this study, we discuss this topic in the revision as a future research (See P25 L6-12).

TECHNICAL COMMENTS

Please do a more thorough English language check of the paper. There are many smaller grammatical errors which occasionally make reading difficult. (I have not listed all of these as I believe a language edit is required).

Response: Thanks you very much for your careful review, in the revision, we thoroughly checked the language by ourselves and the final version was double-checked by a professional editorial service. All the changes were highlighted in blue color.

pp5000, Lines 1-5. It seems unfair to attribute knowledge of flooding, water shortages and ecological degradation to a few recent papers. I recommend removing these citations as these are commonly known facts.

Response: Thanks very much for your suggestion. We removed these citations accordingly (See P2 L17-19).

pp5000, Lines 13-16. Again you are referring to recent papers for a longstanding conclusion

Response: Thanks very much for your comments. We also removed these citations accordingly (See P2 L24- 27).

pp5000, Line 20. Also availability of open data contributes.

Response: Thanks very much for your comments. The availability of open data was also added in the sentence (See P2 L29).

pp5001, Line5. Spelling: Darcy's law

Response: Thanks very much for your careful review. It was revised accordingly (See P3 L8).

pp5001, Line 12. "Several models:" – Here you give no examples and quote a single paper –does this paper summarise all these models?

Response: Thanks very much for your careful review. Some references were added (See P3 L15-17), i.e., "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".

pp5004, Line 7. "Non-point pollution" is also commonly referred to as "diffuse pollution".

Response: Thanks very much for your careful review and comments. It was revised accordingly. We also kept the expression "nonpoint pollution source" as an alternative when the diffuse source was first mentioned (See P2 L2; P5 L15).

pp5004, Line 21. What is meant by dry land? I presume you mean dryland agriculture however this is not clear.

Response: Thanks very much for your careful review and comments. It was revised to be dryland agriculture (See P6 L19-20; P12 L14; P14 L14, 18, 19, 25; P15 L19, 24; P20 L27; P21 L20).

pp5005, Lines 14-15. This sentence doesn't make sense. Please reword.

Response: Thanks very much for your careful review. It was revised to "The ecological processes are described by the soil biochemical module and the crop growth module. The crop growth and soil biochemical processes directly affect the soil moisture, evapotranspiration, and nutrient transformation and loss from soil layers. Therefore, our model incorporates the water cycle, nutrient cycle, crop growth, and their key linkages.” (See P7 L9-13).

pp5009, Line 21 This sentence is more complex that it needs to be. What about “Point sources of pollutants are directly added to surface water in the model. Common point sources are urban water treatment plants or industrial plants.

Response: Thanks very much for your suggestions. The sentence was revised following your suggestion (See P10 L3-5).

P5013, Line 10. What is PAT?

Response: Thanks very much for your review. PAT was the parameter analysis tool, which was mentioned in P5 L25. This section was renumbered as section 2.1.5, in order to keep the balance of individual sections (See P11 L6-P12 L33).

P5017, Line 28. “little worse”, This should probably be “a little worse”, otherwise you are in fact saying that the results are not very different.

Response: Thanks very much for your careful review. It was “a little worse”. The sentence was revised accordingly (See P18 L28, P19 L13).

Comments from A. Slaughter (Referee)

Abstract

Some errors with Grammar:

Line 2 of the abstract: ‘crises’ plural.

Response: Thanks very much for your careful review. It was revised accordingly (See P1 L20).

Line 6 of the abstract: use ‘obtained’ rather than ‘gotten’.

Response: Thanks very much for your careful review. This sentence was a lengthy description of TVGM in the abstract and deleted in the revision.

Line 21 of the abstract: NH4 is ammonium, not ammonia.

Response: Thanks very much for your careful review. It was revised accordingly (See P2 L8).

Introduction

Change first line to: Severe water crises are global issues that have emerged as a consequence of the rapid development of the social economy, and include flooding, water shortages, water pollution and ecological degradation.

Response: Thanks very much for your careful review. It was revised accordingly (See P2 L17-19).

Line 11: The integrated river basin management paradigm:

Response: Thanks very much for your careful review. It was revised accordingly (See P2 L24).

Line 16 page 5001: ‘As a result, these models generally show satisfactory performance in simulating major hydrological processes.’

Response: Thanks very much for your careful review. It was revised accordingly (See P3 L22- 23).

Line 25 page 5001: You mention QUAL2E but not QUAL2K which is the updated version of QUAL2E.

Response： Thanks very much for your careful review and suggestion. In the current model version, we used QUAL2E for water quality simulation. We also made some improvements. For examples, we solved the model at the sub-basin scale rather than the fine grid scale in order to keep spatial consistent with the hydrological cycle module. Like QUAL2K, multiple loadings and abstractions can also be the inputs to any reach according to the geographic positions (See P10 L2-5; P10 L10-12). In the future version, the improvements of QUAL2K will be included, especially the reaction simulation in the anoxic conditions.

Introduction

It may be worth mentioning that there is a conflict of temporal scales when considering an integrated water system model: It is common for hydrology and flow to be simulated at a monthly time scale, as in most cases, this is considered a sufficient resolution for water resource planning and management. However, water quality must be simulated at a daily time scale or less, as water quality is affected by transient flow events such as rainfall-runoff events.

Response： Thanks very much for your comments. In this revision, the scale issue was mentioned in the introduction (See P3 L18), and analyzed in the Section 2.2.1 from P13 L4 to P14 L10.

There are some major issues with grammar in the introduction.

Response: Thanks you very much for your careful review. In the revision, we thoroughly checked the language by ourselves and the final version was double-checked by a professional editorial service. All the changes were highlighted in blue color.

Materials and methods

A better description of the parameter analysis tool is required. Is this a parameter estimation tool? How does it work?

Response: Thanks very much for your comments. Following the comments of reviewer 1, we moved the description of parameter analysis tool (PAT) to Section 2.1.5 because this part belongs to the framework of the proposed model (Figure 1). PAT includes parameter sensitivity analysis, autocalibration and model performance evaluation. The detailed descriptions of PAT were introduced in P11 L14-P12 L33.

You refer to nonpoint pollutant sources. Why not just call them diffuse sources?

Response: Thanks very much for your comments. It was revised accordingly. We also kept the expression “nonpoint pollution source” as an alternative when the diffuse source was first mentioned (See P2 L2; P5 L15).

It is not clear how baseflow separation is achieved: i.e. how surface flow, interflow and baseflow are calculated. You say: ’The interflow and baseflow are considered as linear relationships between storage and outflow’, but I am not sure what this means. Is there a specific method that was used to separate flow fractions? I have done this in the past using a statistical baseflow separation method.

Response: Thanks very much for your careful review and comments. The calculations of surface flow, interflow and baseflow were presented from P6 L25 to L26, and from P26 L4-L5 in the Appendix A. In fact, in this model, we supposed that the interflow and baseflow were from the upper and lower soil layers, respectively. The yields were linear correlated with soil moisture of the upper and low layers, respectively. Therefore, they were calculated individually. This approach is widely adopted in the hydrological models (e.g., VIC, SWAT, and DTVGM).

The hydrological modelling and the statistical method are two commonly used methods to separate the baseflow. However, the statistical method is usually used to separate the baseflow from the observed runoff series, which could be useful to calibrate the baseflow parameters of hydrological cycle models. In this paper, we did not investigate this topic, and only selected the total runoff to demonstrate the performance of hydrological cycle module.

You mention various decomposition, denitrification and nitrification processes simulated. How do you account for the temperature effects on the rates of these processes? Do you simulate or read in temperature (both air and water)?

Response: Thanks very much for your comments. We read in the maximum and minimum air temperature. The water and soil temperature are calculated in the model (See P10 L1; P14 L12; P27 L1-10 in the manuscript; P S8 L15-20 in the supplementary material).

In the WQM what is the time resolution at which water quality is simulated? You describe using QUAL2E for water quality in rivers, but why not QUAL2K which is the updated model?

Response: Thanks very much for your comments. We use a daily time-step in the model, as this is consistent with the underlying rainfall-runoff module (See P13 L30-31). In the current model version, we used QUAL2E for water quality simulation and made some improvements. For examples, we solved the model at the sub-basin scale rather than the fine grid scale in order to keep spatial consistent with the hydrological cycle module. Like QUAL2K, multiple loadings and abstractions can also be the input to any reach according to the geographic positions (See P10 L2-5). In the future version, the improvements of QUAL2K will be included, especially for the reaction simulation in the anoxic conditions.

Not enough description is provided of the method for representing water quality variable fate in dams. Processes affecting water quality in dams and lakes are very complicated and must consider stratification, sedimentation and algal uptake. In that respect, how does the model account for uptake of nutrients by algae and macrophytes?

Response: Thanks very much for your comments. We were very sorry that the detailed descriptions were not provided in the manuscript for the simulation of water quality variables in water bodies (rivers, reservoirs or lakes). The main reason was that we adopted the mature water quality models, which has presented the detail descriptions of fate and degradation of water quality variables (e.g., nitrification and denitrification, sedimentation, resuspension, decay, algal uptake). Thus, we just cited the references of the adopted model (See P10 L8) in order to avoid the redundant descriptions in the manuscript. However, we provided a clear flowchart of the processes of water quality variables in Figure 4. The detailed equations were also provided in the supplementary material (P S8 L12- S11 L1).

We did not considered the stratification of water impounding in this study. The main reasons were that, on the one hand, the high resolution bathymetric data of individual dams or lakes should be needed, on the other hand, the extended model focused on the processes of water and nutrients at the basin scale. This issue was mentioned in the model limitations in the discussion section (See P24 L7-9).

In regards to landuse units considered, does agriculture consider rain-fed agriculture as opposed to irrigated agriculture, or is this not relevant to the catchment studied?

Response: Thanks very much for your comments. The agriculture irrigation is relevant to the ecological processes modules and has considered in the proposed model (See Fig.3a). However, in the case study, we did not consider the irrigation because the data were hard to collect in our study area. However, we considered the human water withdraw in the dam regulation module (See P30 L8-15).

In regards to parameter analysis and calibration, it would be good to discuss equifinality as well as the need for independent calibration and validation (confirmation) data sets.

Response: Thanks very much for your suggestions. In the revision, the equifinality was discussed in Section 4.2 (See P22 L27-P23 L28). The calibration and validation data sets for the individual subsystems and their necessity were also presented in Table 2.

This paper requires editing by a professional copy editor. There are just too many grammatical mistakes.

Response: Thanks you very much for your careful review. In the revision, we thoroughly checked the language by ourselves and the final version was double-checked by a professional editorial service. All the changes were highlighted in blue color.

What concerns me about this paper is the description of a large model has been condensed into one publication. Consequently, conceptual descriptions of the model components are too brief and don’t provide sufficient information. A common strategy for publication of this sort of work is to write several publications, with later publications building on the earlier publications. So for example, one could start with a discussion of the hydrological modelling, and in a later publication deal with the water quality modelling within the same catchment.

Response: Thanks very much for your good suggestions. We agreed that the length of this paper was long. However, we believe that it will benefit the readers to introduce the whole model framework with a case study in one publication because one of main objectives was to present an integrated water system model, rather than individual models. We have tried to reduce the length of the paper by presenting the detailed equations in some appendices and supplementary material. Moreover, the detailed descriptions of some modules were also added, such as section 2.1.5 from P 11 L6-P12 L33. We also provided the detailed model descriptions in the supplementary material (See equations S7, S8, S30-S32, S34, and S35-47).

Comments from anonymous Referee #3

The hydrological cycle and surface water quality are closely related to vegetation, soil and biogeochemical elements, which are strongly influenced by human activities. It is essential to quantify interactions among these components for watershed management. This research developed a comprehensive model, which is in great need to provide a tool for better understanding of system function.

Response: Thanks very much for your comments and careful review. We revised the manuscript carefully by following the comments from all the reviewers. All the changes were highlighted in blue color. Acknowledgements were added in the revision.

One of my concern is the optimisation of the model structure, as sub-models were developed separately for different proposes. For example the biogeochemical module and crop growth module are site-specific, which may need detailed soil input. It is quite difficult to obtain in current soil datasets.

Response: Thanks very much for your comments. Indeed, the ecological process modules are site-specific while the hydrological cycle module and water quality process module are at the basin scale. The solution of different spatial scales is one of important issues for the integration of different modules. In this study, we designed three levels of spatial calculation units, i.e., sub-basin, land-use and crop. The crop and land-use units were approximate to the site or field scale for the ecological process module, while the sub-basin unit was suitable for the hydrological cycle module and water quality process modules. The outputs of different levels of units were exchanged based on the area percentage of units (See P13 L6-29).

The detail input datasets of underlying surface are helpful to improve the model performance, especially for HCM, SBM, CGM, and SEM. We can still obtain the main inputs of soil characteristics from the current soil datasets although the spatial resolution is not high. The other data were used the default values which are given based on the best of our knowledges by referring the literatures (e.g., SWAT, EPIC, and DNDC) or similar basins.

It may be helpful to show how to conduct model calibration and validation. The model is very comprehensive, and there 182 parameters in the model. It may have difficulty in determination of parameter values in practice. I hope the authors can add one paragraph in the discussion and show how the model is used in practice, and your perspectives in the model's optimisation and application.

Response: Thanks very much for your comments. As usual, for complicated models, sensitivity analysis needs to be conducted first before calibration so that only limited number of parameters need to be calibrated while the rest parameters remain default values which are given based on the best of our knowledges by referring the literatures (e.g., SWAT, EPIC, and DNDC) or similar basins. (See P12 L9-13; P17 L5-7). Moreover, the constraints on parameters are also considered in both parameter sensitive analysis and autocalibration. The model calibration and validation were specified in the Section 2.1.5 from P11 L6 to P12 L33, and Fig.5. Moreover, the equifinality was discussed in section 4.2 from P22 L27- P23 L28).

In supplement 2.1, the ‘accumulated heat’ is actually effective temperature, i.e., average temperature minus a base temperature. The ‘heat unit index’ is actually the thermal time, which may be more understandable. I cannot see HUI ranges from 0 to 1 from S7, as PHUj may not equal to the accumulated HU over growing seasons.

Response: Thanks very much for your careful review and comments. The terminologies were revised accordingly. PHU_j is the required cumulative thermal time for crop j from sowing to maturity, and HU_k is the actual cumulative thermal time in each simulation year. Therefore, when HU_k equals PHU_j (HUI = 1), Crop *j* will be maturity (See P S2 L6-L15).

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 faced 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) is extended to an integrated water system model by coupling multiple water-related processes in hydrology, biogeochemistry, water quality and ecology, and considering the interference of human activities. A parameter analysis tool, which includes sensitivity analysis, autocalibration and model performance evaluation, is 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, is selected as the case study area.

The model performances are evaluated on the key water-related components including runoff, water quality, diffuse pollution load (or nonpoint source) and crop yield. Results show that our proposed model simulates most components reasonably well. In particular, the simulated daily runoff series at most regulated and less-regulated stations match well with the observations. The average correlation coefficient and coefficient of efficiency between the simulated and observed runoffs are 0.85 and 0.70, respectively. Both the simulated low and high flow events at most stations are improved when the dam regulation is considered. The daily ammonium-nitrogen (NH_4-N) concentration, which is used as a key index in the water quality evaluation, is also well captured with the average correlation coefficient of 0.67. Furthermore, the diffuse source load of NH_4-N and the corn yield are reasonably simulated for each administrative region. 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 they 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) and HYPE (Lindström *et al.*, 2010). (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 systems 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). Biogeochemistry 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 biogeochemistry 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.

Unlike the above-mentioned models, 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 simulations of 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 biochemistry models (Gassman *et al.*, 2007). Several modified versions have been developed, such as SWIM (Krysanova *et al.*, 1998), and SWAT-N (Polhert *et al.* 2006, 2007).

In this study, we tend 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 hydrological data 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. DTVGM is currently used in many basins with different scales and climate zones to investigate the effect of human activities and climate change on runoff, and shows good simulation performances (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 is 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 includes popular objective functions, autocalibration approaches and summary statistics, is also developed. To demonstrate the model performances, we simulate 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 the 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 process 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 (HCM) module (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 water quality 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 yield calculation is the core of hydrological simulation. TVGM is adopted to calculate the surface runoff yields for different land-use 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 widely available daily maximum and minimum temperature data 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 layer). Plant transpiration is also linked to the soil biochemical module (to drive the vertical migration of nutrients in the soil layer). The surface runoff is linked to the soil erosion module, while the overland flow 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 by the soil biochemical module and the crop growth module. The crop growth and soil biochemical processes directly affect the soil moisture, evapotranspiration, and nutrient transformation and loss from 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 pollution sources to water bodies (Figure 3a).

Soil C and N cycle. We adopt the sub-models of daily step decomposition and denitrification in DNDC (Li *et al.*, 1992) 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^--N 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 soil organic carbon 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 based on 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, and nutrient uptake, etc. (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 system. 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 sources of pollutants 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 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 loss and migration load 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, 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 sources of pollutant 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 to maintain spatial consistent 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 dams or sluices 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. 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 to help users in the use of our proposed model. It contains 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 coefficient of efficiency (*NS*). 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 by users on the basis of their specific requirements.

In order to obtain 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 types are set in ascending order (from water body, paddy land, urban area, forest, dryland agriculture, unused land to 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 in channels (R_{set}). 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 based on the best of our knowledges by referring the literatures (e.g., SWAT, EPIC, and DNDC) or 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 in the model, namely, sub-basin, land-use and crop from largest to smallest. These units are defined as the minimum polygons with similar hydrological properties, land-use types 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 units of each sub-basin are partitioned by the land-use 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 units (paddy land and dryland agriculture, forest, 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 units are divided into 10 specific categories of crop units as fallow for all these land-use 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 category of a specific land-use 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 unit scale, or sub-basin scale based on the percentages of area 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, 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 and rivers based on DEM, land-use data. The sub-basin attributes (e.g., location, evaluation, area, land surface slope and slope length, land-use 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 is applied to a highly regulated and heavily polluted catchment (the Shaying River Catchment) in China. The simulated water-related components contains daily runoff and water quality concentrations at river cross-sections, spatial patterns of diffuse source pollutant 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 are 8.30 million (big animals: cattle, pigs and sheep) and 178.42 million (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^{o}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 are collected from the government bureaus, official books or scientific references. The detailed descriptions were presented in Tables S2 and S3 of the supplementary material. The Shaying River Catchment are divided into 46 sub-basins. According to the land-use classification standard of China (CNS,2007), the main land use types are 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) are calculated based on the percentage of soil types in each sub-basin. The main crops are early rice and late rice in the paddy land, and winter wheat and corn in the dryland agriculture. The main agricultural management schemes (fertilize, plant, harvest and kill) are summarized by field investigation in the studies of Wang *et al*., (2008) and Zhai *et al.* (2014) (Table S3). Crop rotation and its management scheme are considered in the model by setting the start time, the duration of management and the fertilizer amounts. Two fertilizations (base and additional fertilization) are considered in the model during the complete growth cycle of a certain crop. The areas of sub-basin, land-use and crop units ranged from 46.48 km^2 to 3771.15 km^2, from 0.04 km^2 to 2762.5 km^2, and from 3.73 km^2 to 2762.5 km^2, respectively.

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

Moreover, 5 reservoirs, 12 sluices and over 200 wastewater discharge outlets are considered in the model according to their geographical positions. The farm manure from rural living and livestock farming are considered in the model 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 are used to calibrate the model parameters. There are 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 are hard to collect for the whole catchment, only the statistical results from official reports or statistical yearbooks (Wang, 2011; Henan Statistical Yearbook, 2003, 2004 and 2005) are collected to validate the model performances.

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

The model calibration is conducted by the following steps. Hydrological parameters are 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 are from 2003 to 2005 and from 2006 to 2008, respectively. The weighted sum method is usually used to comprehensively handle multi-objectives (Efstratiadis and Koutsoyiannis, 2010). In this study, single objective functions are formed by equally weighting the evaluation indices as (f_{runoff} and $f_{NH4\text{-}N}$)

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

because the case study is only a demonstration of the model performance.

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

3. Results

3.1 Parameter sensitivity analysis

Nine sensitive parameters are 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 control the main hydrological processes, in which soil water and evapotranspiration processes are distinctly important and explain 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 are explained by 14 sensitive parameters which are 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 is that hydrological processes provide the hydrological boundaries that affect 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 is further influenced by the settling and biological oxidation processes. Moreover, it is 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 fit the observations well at all the stations (Figure 7 and Table 3). The *biases* are 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 are caused by the average objective function of all three evaluation rather than the *bias* only. The *r* values range 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 are 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 can see that the high and low flows are usually overestimated at all stations 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 are simulated well when the dam and sluice regulation is considered (Table 4). The best fitting is 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} range from -0.08 (Zhoukou) to -0.29 (Huaidian) for high flow simulation, from -0.05 (Zhoukou) to -0.31 (Huaidian) for average flow simulation, and from -1.97 (Fuyang) to -3.91 (Yingshang) for low flow simulation except Zhoukou (1.28). The improvements in the low flow simulations are very obvious. However, their performances still need to be improved further, particularly for the underestimation at Zhoukou and Huaidian. The possible reasons are as follows. On one hand, the applied evaluation indices (*r* and *NS*) are known to emphasize the high flow simulation rather than the low flow simulation (Pushpalatha *et al.*, 2012) and the objective of autocalibration is to obtain the optimal solution for the average of three evaluation indices rather than the *bias* only. The slight sacrifice of *bias* improves 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 are even better, particularly for *r* and *NS*. The *r* values range 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 range 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 are improved although they become 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*) are well captured.

3.3 Water quality simulation

The simulated concentrations of NH_4-N match well with the observations according to the evaluation standard recommend by Moriasi *et al.* (2007) (Figure 9 and Table 5). The *r* values are 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 is 0.67. The *bias* are considered as “acceptable” with a range from -0.27 (Fuyang for validation) to 0.29 (Zhoukou for calibration). The best simulation are at Luohe Station. The obvious discrepancies between the simulations and observations often appear in the period from January to May because of the poor simulation performance on the low flows. Although the *biases* change markedly from calibration to validation at Fuyang and Yingshang stations, the model performances are still acceptable. The possible explanation is that the *biases* for corresponding runoff simulations at these two stations also change.

Compared with the results without the consideration of regulation, the simulation results are obviously improved when the regulation is considered except for the calibration at Fuyang Station. The decreases in $f_{NH4\text{-}N}$ value range from 0.10 (Huaidian for calibration) to 0.49 (Zhoukou for validation) although there is a slight increase at

Fuyang for the calibration (0.02). Therefore, it is concluded that the consideration of dam and sluice regulation plays an important role in the water quality simulation. In the upper stream of Shaying River, the flow is small and the NH_4-N concentration decrease obviously because of the degradation and settlement of large water storage. In the downstream of Shaying River, the NH_4-N concentration increases 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 are usually overestimated or are higher than the simulation with regulation at the upstream stations (Luohe and Zhoukou). However, the concentrations are underestimated at the downstream stations (Huaidian, Fuyang and Yingshang). The largest difference between the simulations with and without the consideration of regulation appears at Zhoukou.

The spatial pattern of average annual load of diffuse source NH_4-N is shown in Figure 10a. The estimated annual yield rates range 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 is 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 and fertilizer amount in the official report (Wang, 2011), the *bias* of simulated diffuse source load in the whole region is 21.31% when the two regions with the biggest *biases* (Fuyang and Pingdingshan) are excluded as outliers. The high load regions are in the middle of Pingdingshan, Xuchang, Zhengzhou, Fuyang and Zhoukou regions. The spatial pattern is 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 is just 30%-70% in the paddy areas, which is 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 is approximately 4.70×10^4 t $year^{-1}$ in the Shaying River Catchment. The diffuse source contributes 38.57% of the overall NH_4-N load on average from 2003 to 2005, and this value is slightly higher than the statistical results (29.37%) given in the official report (Wang, 2011). Moreover, the diffuse source contributions at the stations range from 31.72% (Huaidian) to 47.13% (Shenqiu). Compared with the diffuse source loads in the individual administrative regions in 2000, the simulated loads tend to increase from 2003 to 2005 except in Kaifeng region. The yields in Fuyang and Pingdingshan regions increase at highest rates. The primary pollution source in the Shaying River Catchment is still the point source, but the diffuse pollution is also an important concern. In term of spatial variation, the contribution of diffuse source to the pollutant load is high in the upstream and is low in the middle and downstream because the point source emission is 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 is also improved remarkably by considering the detailed processes of nutrient in the soil layers in our model.

3.4 Crop yield simulation

The simulated corn yield and its spatial pattern are shown in Figure 11. The average annual yields are summarized at sub-basin scale and range 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 is 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 are Luohe, Fuyang and Zhoukou in the middle and downstream where the primary land use is the dryland agriculture (93.12%, 95.87% and 93.18%, respectively). The crop yields in Luohe, Nanyang, Kaifeng regions are well simulated. The total yield is 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 (Beven, 2006). Our proposed model is developed in this direction and tends to benefit integrated river basin management. Therefore, in comparison with most existing models, our proposed model considers all the water-related processes as an integrated system rather than isolated systems for individual processes.

Our model provides competitive simulation results in the Huai River Basin (Figures 7-9; Tables 3-5). Several typical models have also been 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). Different models have generally comparable performances on the runoff or water quality simulations. For SWAT, the f_{runoff} values are from 0.11 to 0.20 with the average of 0.16 at the daily scale at the unregulated stations (Shi *et al.*, 2013), and from 0.09 to 0.75 with the average of 0.32 at the monthly scale at the regulated stations (Zhang *et al.*, 2012). The $f_{NH4\text{-}N}$ values range from 0.18 to 0.86 with the average of 0.47 (Zhang *et al.*, 2012). For Xinganjiang model, the f_{runoff} values are from 0.13 to 0.21 with the average of 0.16 at the daily scale at the unregulated stations (Shi *et al.*, 2013). For DTVGM, the f_{runoff} values are 0.14 and 0.21 at the daily scale in the calibration and verification periods, respectively at Bengbu station. Our model performs better than SWAT, especially for the regulated runoff and water quality simulations. Moreover, both the Xinanjiang model and DTVGM can only simulate the flow series at the unregulated or less-regulated stations because they do not consider the dam regulation in their model frameworks.

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, in our model, nine and 14 sensitive parameters are detected for runoff and NH_4-N simulation, respectively (Table 2). 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. 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 are specified in Table 1 and most widely-used measures of model performances are also provided in the PAT. In the case study, we also employ several observation sources (e.g., runoff and water quality observations at different stations, the diffuse pollution load and crop yield data), and use 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 are still inaccurate. The current version focuses 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 are 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 considers the main water-related processes in the hydrological, ecology and water quality subsystems but numerous processes are 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 is extended primarily to an integrated water system model to address the complex water issues emerging in the basins. The model performance is demonstrated in the Shaying River Catchment, China. The model provides 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 load, water quality concentrations in water bodies), and ecological components (e.g., crop yield) which could be calibrated if observations are available. The case study shows that the simulated runoffs at most stations fit the observations well in the highly regulated Shaying River Catchment. All the evaluation criteria are acceptable for both the daily and monthly simulations at most stations. This model well simulates the discontinuous daily NH_4-N concentration and properly captures 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 are acceptable and several processes are still not well calibrated (such as low flow events, diffuse pollution load, and crop yield). 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.

Appendix A: Hydrological cycle module

The basic water balance equation is

$$P_i + SW_i = SW_{i+1} + Rs_i + Ea_i + Rss_i + Rbs_i + In_i \tag{A1}$$

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

$$\begin{cases} E_a = E_t + E_s \le E_0 \\ E_p = \begin{cases} LAI \cdot E_0/3 & 0 \le LAI \le 3.0 \\ E_0 & LAI > 3.0 \end{cases} \\ E_s = E_0 \cdot exp(-5.0\times10^{-5} \cdot rsd) \end{cases} \tag{A2}$$

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

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

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}, T_{overl} and T_{rch} are the routing times of the total flow, 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.

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

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

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

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.

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

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

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

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

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

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.

Pohlert, T., L. Breuer, J.A. Huisman, and H.-G. Frede.: Assessing the model performance of an integrated hydrological and biogeochemical model for discharge and nitrate load predictions. *Hydrol. Earth Syst. Sci.* 11:997–1011, 2007.

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.

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 map	Land use 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).

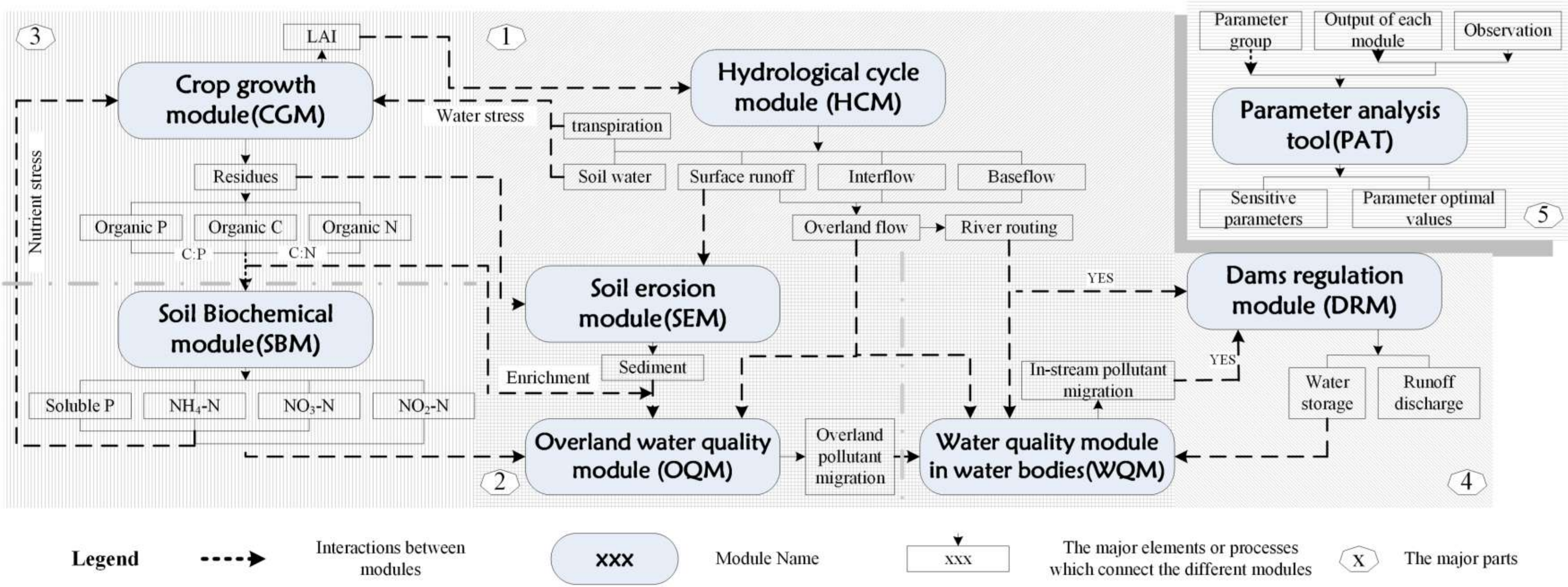
Figure 7. The daily runoff simulation at all stations.

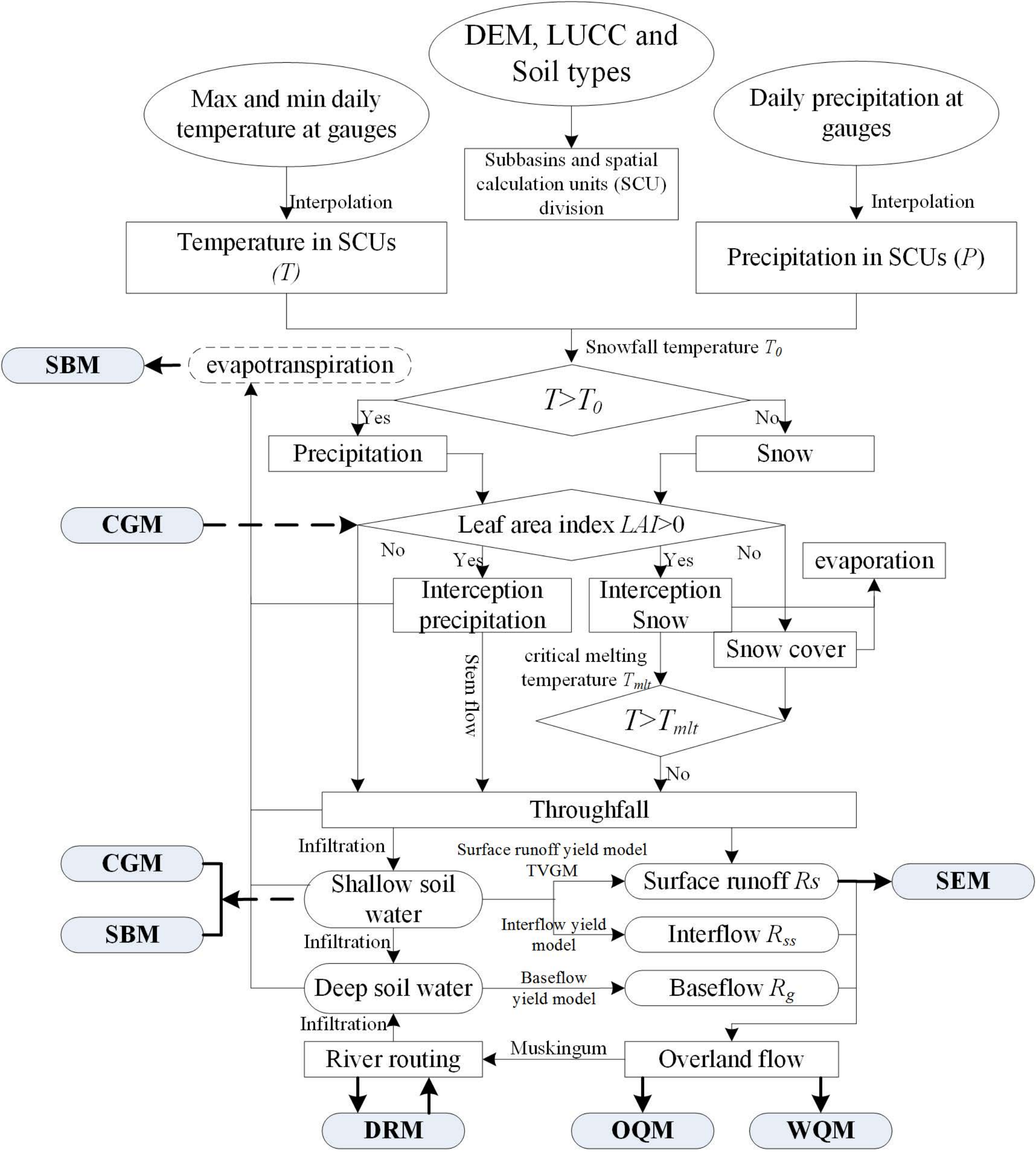
Figure 8. The cumulative distributions of simulated and observed daily runoff at all stations

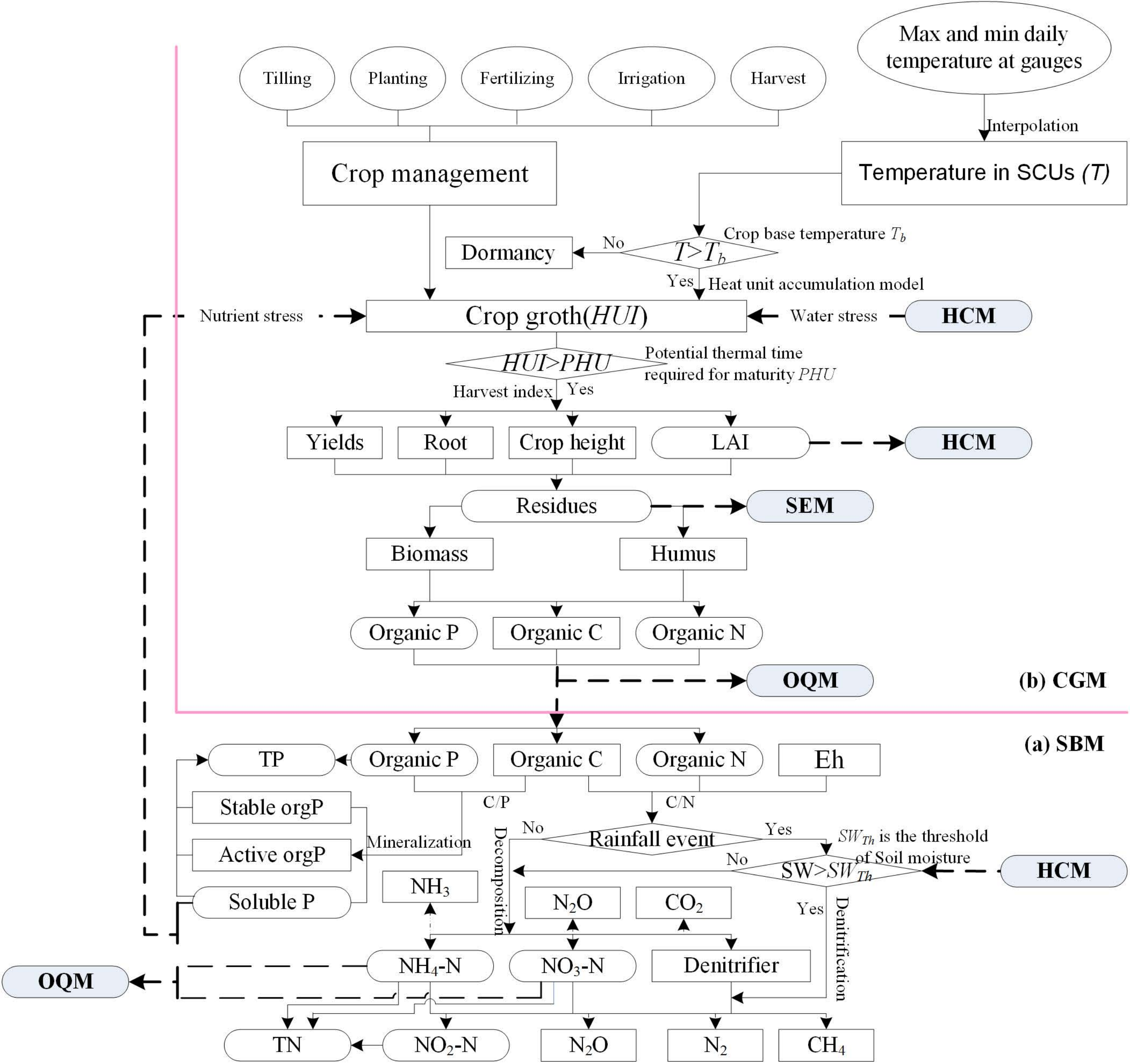
Figure 9. The simulated NH_4-N concentration variation at all stations.

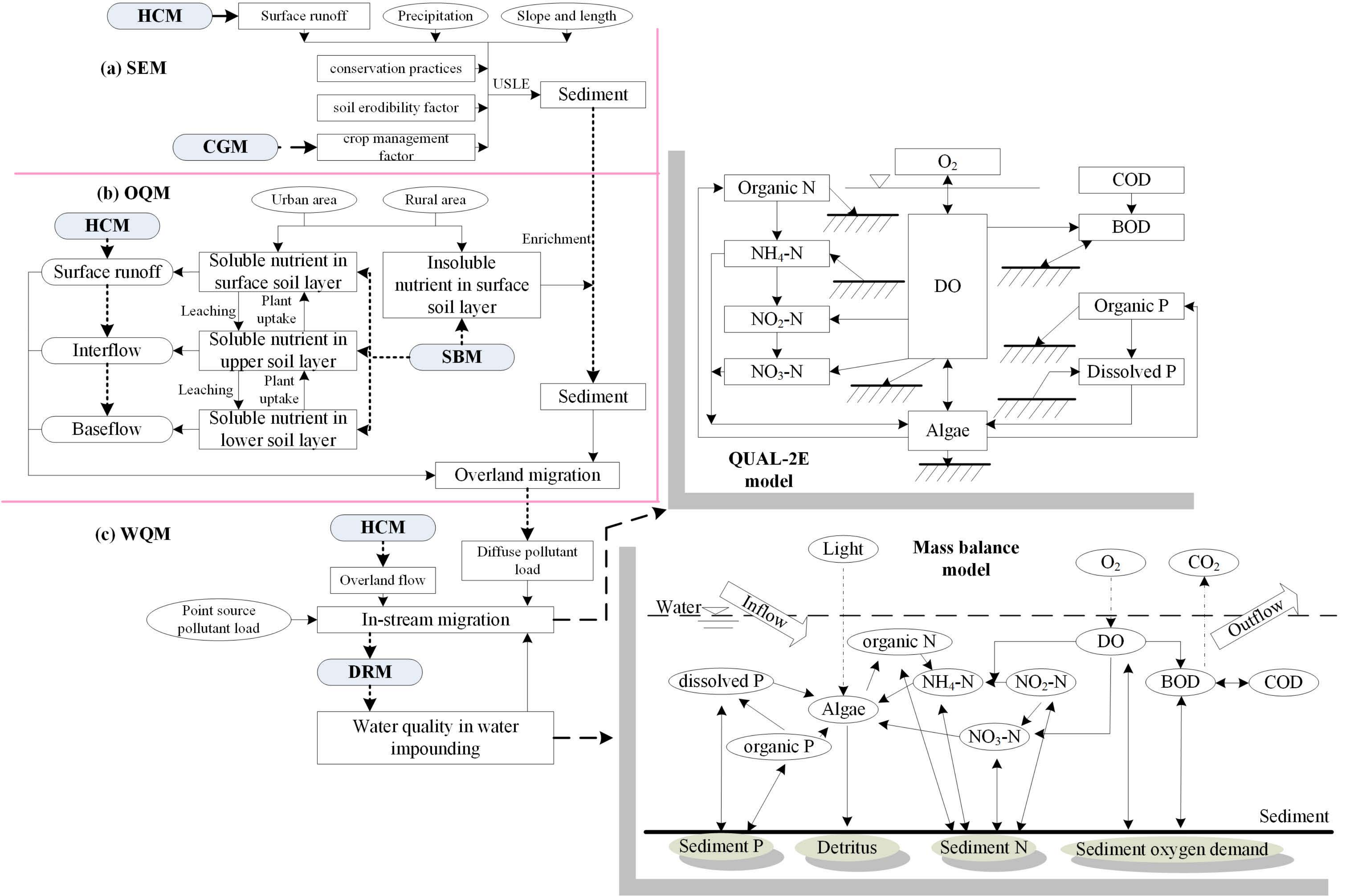
Figure 10. The spatial pattern of diffuse source NH_4-N load (a) and its relationship with paddy area (b) and rice yield (c) at the sub-basin and regional scale in the Shaying River Catchment.

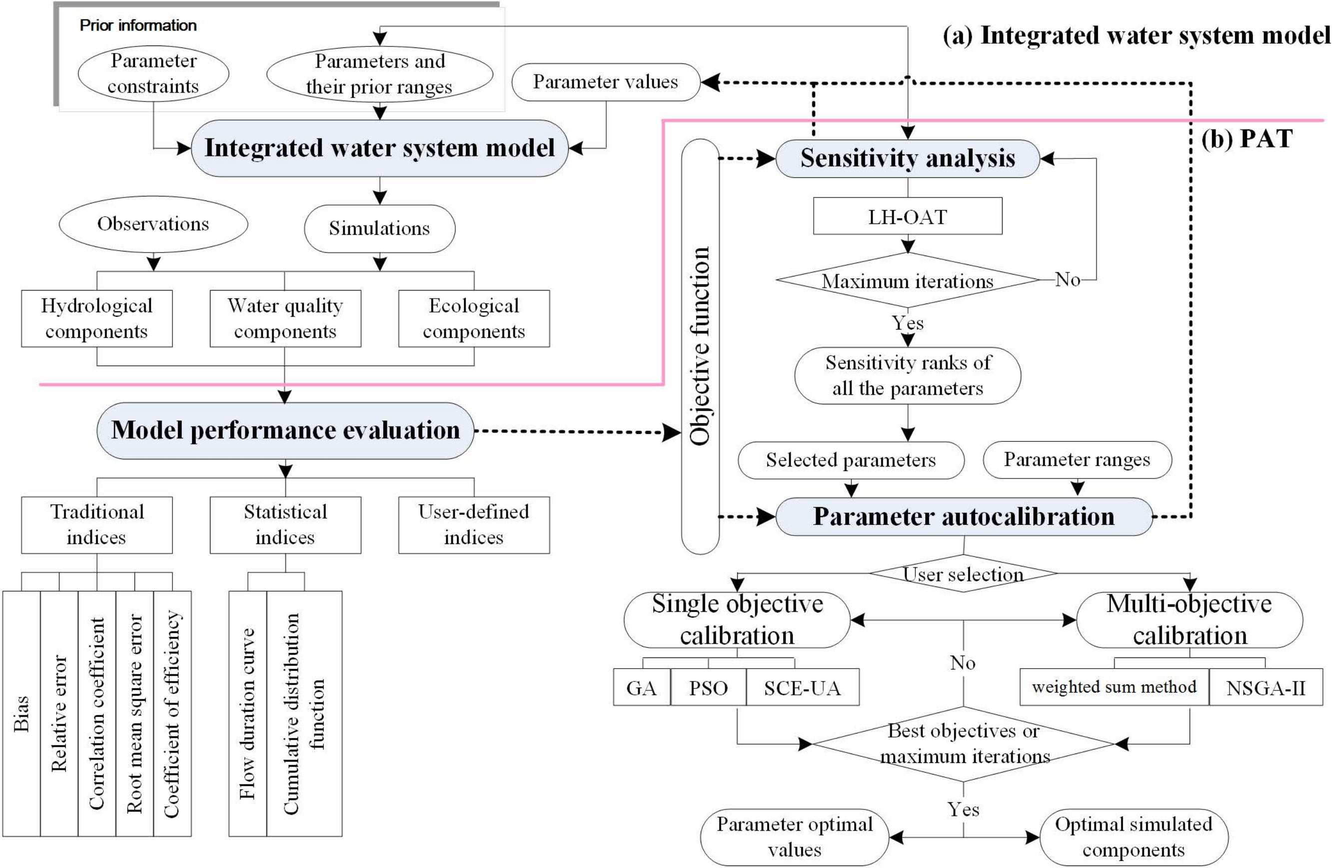
Figure 11. The spatial pattern of corn yield at the sub-basin and regional scale in the Shaying River Catchment.

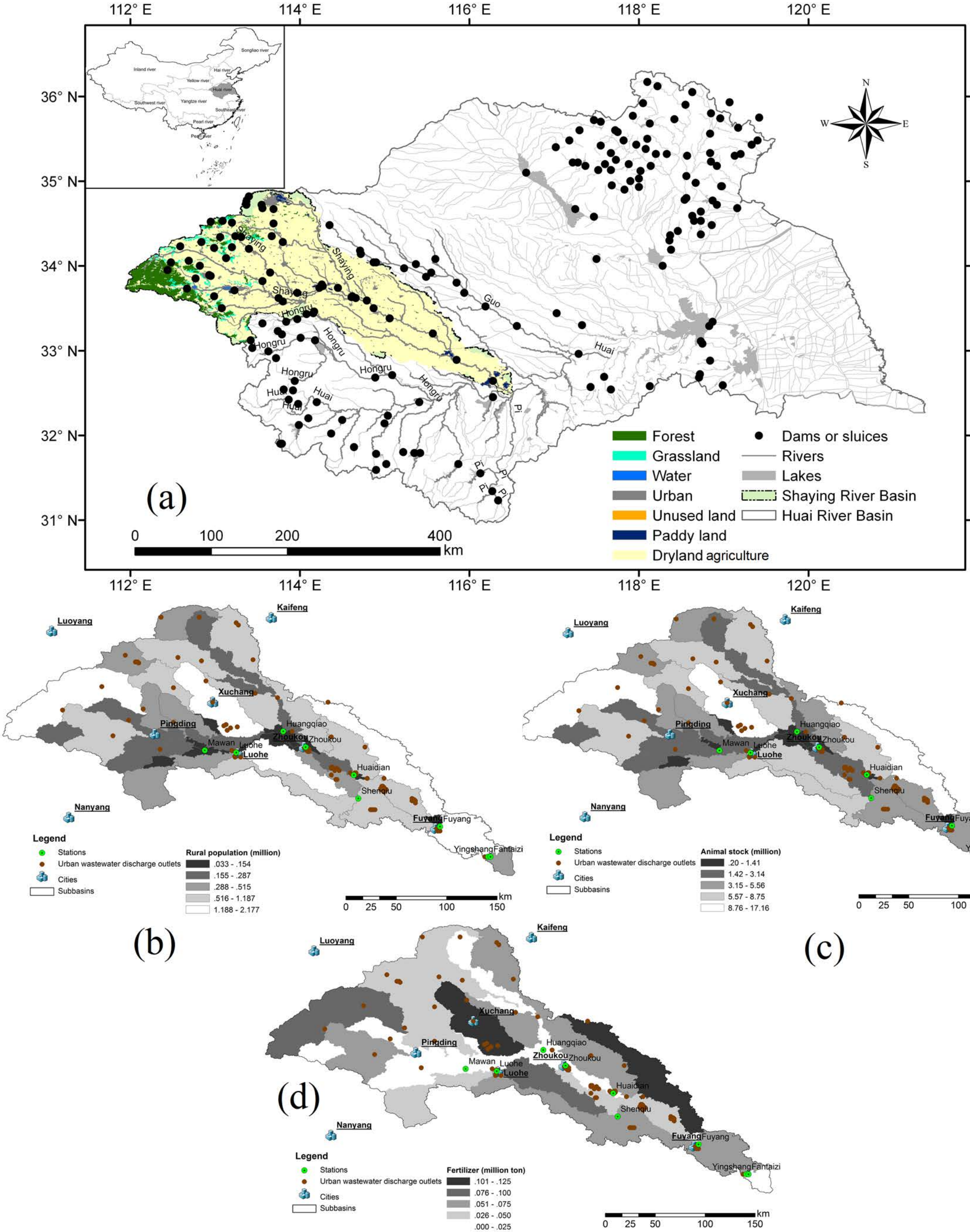

③
LAI
①
Hydrological cycle module (HCM)
Crop growth module(CGM)
Water stress
transpiration
Nutrient stress
Residues
Soil water
Surface runoff
Interflow
Baseflow
Organic P
Organic C
Organic N
Overland flow
River routing
C:P
C:N
Soil erosion module(SEM)
Soil Biochemical module(SBM)
Sediment
Enrichment
Soluble P
NH4-N
NO3-N
NO2-N
Overland water quality module (OQM)
Overland pollutant migration
Water quality module in water bodies(WQM)
②
In-stream pollutant migration
YES
YES
Dams regulation module (DRM)
Water storage
Runoff discharge
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

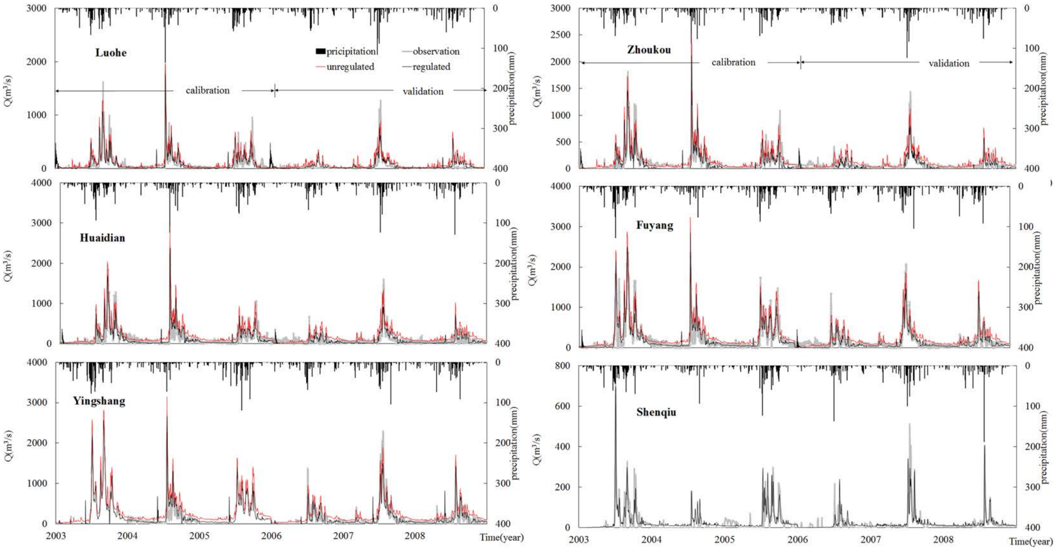

DEM, LUCC and Soil types
Max and min daily temperature at gauges
Daily precipitation at gauges
Subbasins and spatial calculation units (SCU) division
Interpolation
Interpolation
Temperature in SCUs (T)
Precipitation in SCUs (P)
Snowfall temperature T0
SBM
evapotranspiration
T>T0
Yes
No
Precipitation
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
SBM
Shallow soil water
Surface runoff Rs
SEM
Interflow yield model
Interflow Rss
Infiltration
Deep soil water
Baseflow yield model
Baseflow Rg
Infiltration
River routing
Muskingum
Overland flow
DRM
OQM
WQM

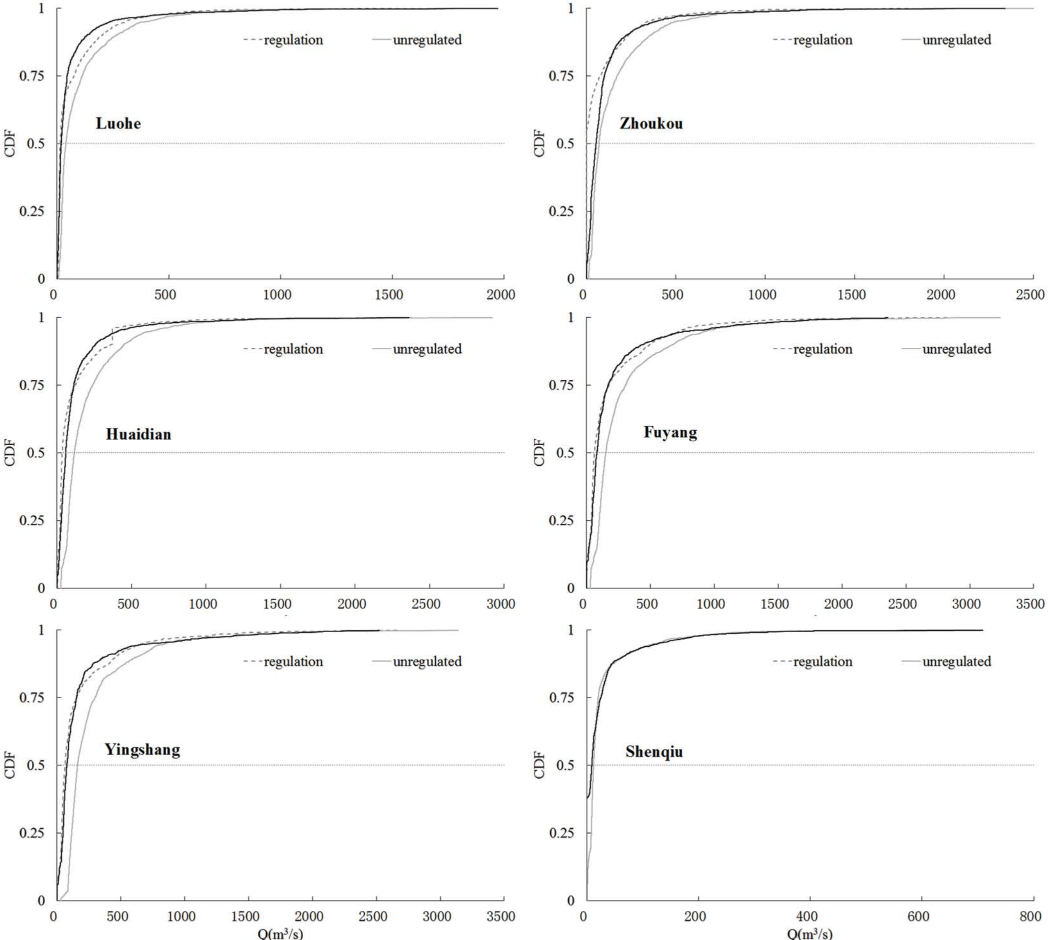

Max and min daily temperature at gauges
Tilling
Planting
Fertilizing
Irrigation
Harvest
Interpolation
Crop management
Temperature in SCUs (T)
Crop base temperature T_b
Dormancy
No
$T>T_b$
Yes
Heat unit accumulation model
Nutrient stress
Crop groth(HUI)
Water stress
HCM
$HUI>PHU$
Potential thermal time 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
Active orgP
Mineralization
No
Rainfall event
Yes
SW_{Th} is the threshold of Soil moisture
Decomposition
No
$SW>SW_{Th}$
HCM
NH_3
Soluble P
N_2O
CO_2
Yes
Denitrification
NH_4-N
NO_3-N
Denitrifier
OQM
TN
NO_2-N
N_2O
N_2
CH_4

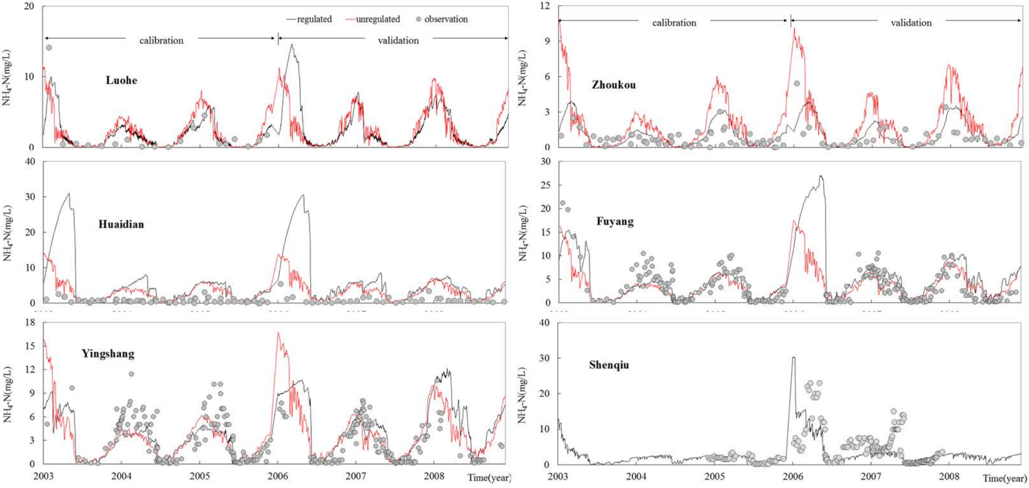

HCM
Surface runoff
Precipitation
Slope and length
(a) SEM
conservation practices
USLE
Sediment
soil erodibility factor
CGM
crop management factor
(b) OQM
Urban area
Rural area
HCM
Enrichment
Surface runoff
Soluble nutrient in surface soil layer
Insoluble nutrient in surface soil layer
Leaching
Plant uptake
Interflow
Soluble nutrient in upper soil layer
SBM
Leaching
Plant uptake
Sediment
Baseflow
Soluble nutrient in lower soil layer
Overland migration
(c) WQM
HCM
Overland flow
Diffuse pollutant load
Point source pollutant load
In-stream migration
DRM
Water quality in water impounding
O$_2$
Organic N
COD
BOD
NH$_4$-N
DO
Organic P
NO$_2$-N
NO$_3$-N
Dissolved P
Algae
QUAL-2E model
Light
Mass balance model
O$_2$
CO$_2$
Water
Inflow
Outflow
organic N
DO
dissolved P
NH$_4$-N
NO$_2$-N
BOD
COD
Algae
organic P
NO$_3$-N
Sediment
Sediment P
Detritus
Sediment N
Sediment oxygen demand

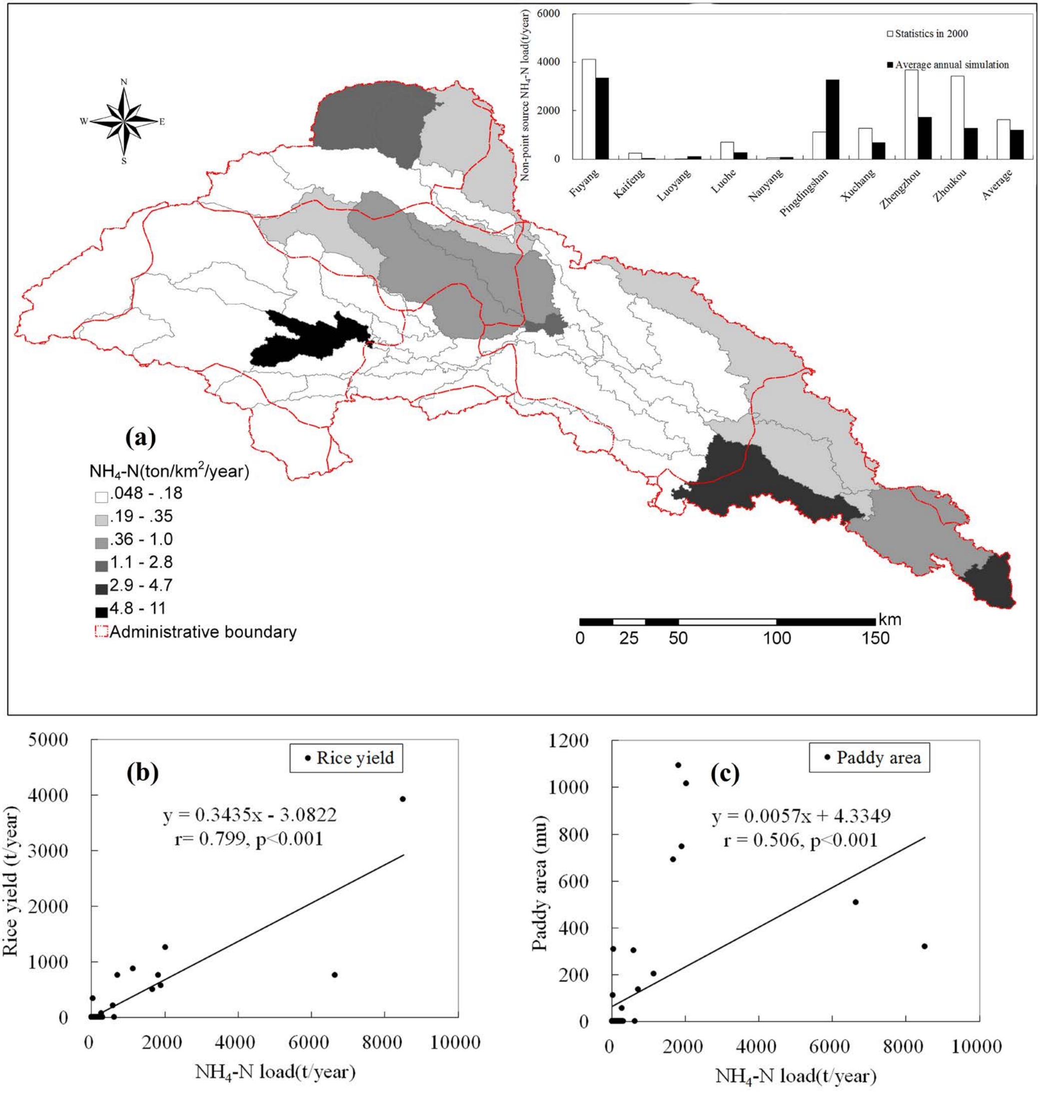

(a) Integrated water system model
Prior information
Parameter constraints
Parameters and their prior ranges
Parameter values
Integrated water system model
Observations
Simulations
Hydrological components
Water quality components
Ecological components
Model performance evaluation
Traditional indices
Statistical indices
User-defined indices
Bias
Relative error
Correlation coefficient
Root mean square error
Coefficient of 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
weighted sum method
NSGA-II
No
Best objectives or maximum iterations
Yes
Parameter optimal values
Optimal simulated components

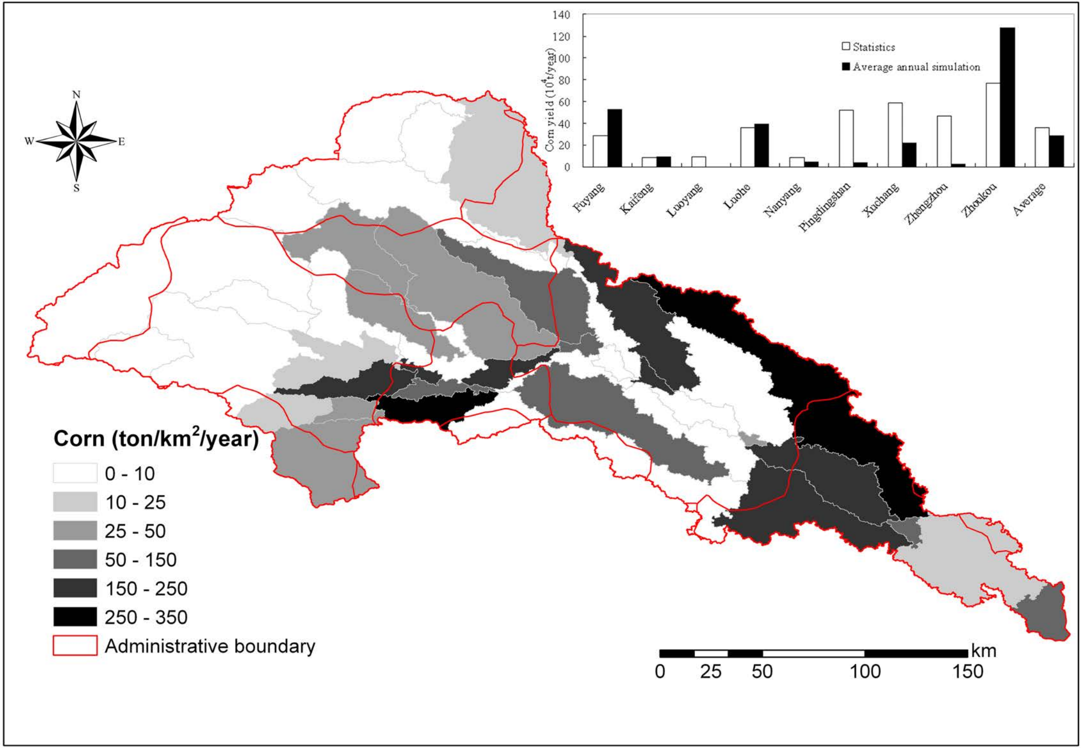

112° E
114° E
116° E
118° E
120° E
36° N
35° N
34° N
33° N
32° N
31° N
Songliao river
Inland river
Hai river
Yellow river
Huai river
Southwest river
Yangtze river
Southeast river
Pearl river
N
S
W
E
Shaying
Hongru
Huai
Guo
Pi
(a)
0
100
200
400
km
Forest
Grassland
Water
Urban
Unused land
Paddy land
Dryland agriculture
Dams or sluices
Rivers
Lakes
Shaying River Basin
Huai River Basin
Luoyang
Kaifeng
Xuchang
Pingding
Huangqiao
Zhoukou
Mawan
Luohe
Huaidian
Shenqiu
Nanyang
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
0
25
50
100
150
km
(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
0
500
1000
1500
2000
2500
3000
4000
200
400
600
800
100
300


Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
Shenqiu
regulation
unregulated
CDF
0
0.25
0.5
0.75
1
0
500
1000
1500
2000
2500
3000
3500
0
200
400
600
800
Q(m³/s)


regulated
unregulated
observation
calibration
validation
Luohe
Zhoukou
Huaidian
Fuyang
Yingshang
Shenqiu
NH$_4$-N(mg/L)
Time(year)
2003
2004
2005
2006
2007
2008


N
W
E
S
Non-point source NH4-N load(t/year)
6000
4000
2000
0
Statistics in 2000
Average annual simulation
Fuyang
Kaifeng
Luoyang
Luohe
Nanyang
Pingdingshan
Xuchang
Zhengzhou
Zhoukou
Average
(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
(b)
Rice yield
y = 0.3435x - 3.0822
r= 0.799, p<0.001
Rice yield (t/year)
5000
4000
3000
2000
1000
0
0
2000
4000
6000
8000
10000
NH4-N load(t/year)
(c)
Paddy area
y = 0.0057x + 4.3349
r = 0.506, p<0.001
Paddy area (mu)
1200
1000
800
600
400
200
0
0
2000
4000
6000
8000
10000
NH4-N load(t/year)


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