

Interactive comment on “Integrated water system simulation by considering hydrological and biogeochemical processes: model development, parameter sensitivity and autocalibration” by Y. Y. Zhang et al.

Anonymous Referee #1

Received and published: 11 August 2015

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. Acknowledgement was 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 extended model and re-numbered its section as section 2.1.5 (See P11 L29-P13 L2).

Moreover, the interconnections between PAT and other modules were specified in section 2.1.5 (See P12 L250-P13 L2). 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 P8 L1-P9 L6).

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 L5- P23 L2 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).

Motivated by the requirements for the integrated river basin management, integrated water system model should be further developed to produce reasonable simulation in both water quantity and quality processes in the real basins with possibility to simulate more water-related processes, such as soil biogeochemistry, crop growth. In this study, we extend the simulation functions of DTVGM as an integrated water system model and improves the modeling practice of water-related components. Our specific objectives are as follows: (1) to integrate the detailed interactions and linkages among hydrological, water quality, soil biogeochemical and ecological processes, as well as the prevalent regulations of water projects at the basin scale; (2) to couple robust parameter analysis approaches with the integrated water system model to improve model performances; (3) to examine the applicability of the extended model on key water-related components in complex basins, e.g., flow regimes, nonpoint source (or called diffuse source) pools of nutrients, water quality variables in water body and crop yield.

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 L5-P24 L22).

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 L5-P24 L22). 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 L9-31). In our case study, the areas of sub-basin, land-use and crop units ranged from 46.48 to 3771.15 km², from 0.04 to 2762.5 km², and from 3.73 to 2762.5 km², respectively (See P16 L1-3).

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 L18-20; P S14-S15). 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 S15).

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 L20-22; P2 L30-P3 L2). We also updated the references by following your suggestions in other comments (P32 L8-14; P34 L22-24;

P35 L 6-8, L15-17; P37 L12-14).

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 regulation 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 processes from which the model is oriented but inadequately simulates other processes. In particular, we have used some typical models in our case study area but found that they did perform well enough on the concerned water quantity and quality processes (See P22 L5- P23 L2 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).

Motivated by the requirements for the integrated river basin management, our study makes an attempt to extend the simulation functions of DTVGM as an integrated water system model and improve the modelling practice of water related components. Our specific objectives are: (1) to integrate the detailed interactions and linkages among hydrological, water quality, soil biogeochemical and ecological processes, as well as the prevalent regulations of water projects at the basin scale; (2) to couple the robust parameter analysis approaches with the integrated water system model to improve the model performances; (3) to examine the applicability of the extended model on some key water-related components in complex basins, e.g., flow regimes, nonpoint source (or called diffuse source) pools of nutrients, water quality variables in water body and crop yield.

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 (See P2 L5-6). 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 L6-9; P5 L7-20). 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 L15-23).

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 P12 L9-13; P23 L30- P24 L2). Moreover, we also collected several observations of representative water related components (e.g., runoff and water quality observations at different stations, the nonpoint pollutant source load and crop yield data), to calibrate each subsystem (See P23 L26-29).

The equifinality, model limitations and the comparison with other models were discussed in the new discussion section (See P22 L5-P24 L22).

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 L16-18; P32 L8-14; P34 L22-24; P37 L12-14)

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

Response: Thanks very much for your comments. Yes indeed, 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 P4 L3, 4; P5 L2-5; P32 L10-14; P 35 L6-8, 15-17.

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 P11 L14-16).

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 time scale might improve the performance in some modules (e.g., SEM, WQM). However, most observations (e.g., climate data sets, soil nutrient availability, water quality concentrations in the field measurements), are provided at daily time scale, thus leading to potential uncertainty or inability to downscale the observation into a sub-daily time scale. Moreover, the sub-daily module will increase model complexity compared to regional simulations, as it is quite hard to obtain the information at the regional scale. Thus most processes need to be simplified to fit regional research” (See P14 L1-9).

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 were inevitable in the integrated water system model. However, 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 the discussion in P23 L4- P24 L2). In this study, we restricted SCE-UA to calibrate only the sensitive parameters as defined by LH-OAT, while other parameters remained constant. Moreover, we also collected some observations of representative water related components (e.g., runoff and water quality observations at different stations, the nonpoint pollutant source load and crop yield data), to calibrate each subsystem (See P16 L21-32).

The comparison with other models, equifinality, and model limitations were discussed in the section of discussion (See P22 L5-P24 L22).

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 changed in the opposite ways. For examples, the runoff underestimation usually cause the overestimation of water quality concentration, while the runoff overestimation usually cause the underestimation of water quality concentration. Therefore, the large change in bias for NH₄-N concentrations between calibration and validation at Fuyang and Yingshang was probably caused by the change in bias for the corresponding runoff simulations. However, the model performances of both calibration and validation were acceptable. We did not think the limitations in calibration was related to these large changes (See P19 L32 – P20 L3).

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 were specified in Tables S2-S3 in the supplementary material. All the data sets for model setup and calibration were collected from the government bureaus, official books or scientific references (See P15 L18-20; P S14-S15). In this study, we did 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 L13-14).

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 L20-22).

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 L30- P3 L2).

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 P3 L5).

pp5001, Line5. Spelling: Darcy's law

Response: Thanks very much for your careful review. It was revised accordingly (See P3 L18).

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 L25-27), 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; for an overview see 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. The expression "nonpoint source pollutant" is also widely used in the water related sciences. In the revision, we kept this terminology, and also added the expression "diffuse sources" as an alternative when the nonpoint source was first mentioned (See P2 L4; P6 L1).

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 P7 L4; P13 L18, 20, 27; P15 L22, 27; P21 L1, L29).

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 process modules contain the soil biochemical module and the crop growth module. The crop growth and soil biochemical processes directly affect the soil moisture, evapotranspiration, the nutrient transformation and loss from soil layers. Therefore, the model incorporates the water cycle, nutrient cycle, crop growth, and their key linkages." (See P7 L25-29).

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 L20-23).

P5013, Line 10. What is PAT?

Response: Thanks very much for your review. PAT was the parameter analysis tool, which was mentioned in P6 L9. This section was renumbered as section 2.1.5, in order to keep the balance of individual sections (See P11 L29-P13 L2).

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 L24, P19 L19).

1 **Integrated water system model considering**
2 **hydrological and biogeochemical processes: model**
3 **development, with parameter sensitivity and**
4 **autocalibration**

5

6 **Y.Y. Zhang^{1,*}, Q.X. Shao^{2,*}, A.Z. Ye³, H.T. Xing^{1,4}, and J.Xia¹**

7 [1] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of
8 Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences,
9 Beijing, 100101, China.

10 [2] CSIRO Digital Productivity Flagship, Leeuwin Centre, 65 Brockway Road,
11 Floreat Park, WA 6014, Australia.

12 [3] College of Global Change and Earth System Science, Beijing Normal University,
13 100875, Beijing, China.

14 [4] CSIRO Agriculture Flagship, GPO BOX 1666, Canberra, ACT 2601, Australia

15 Correspondence to: Y.Y. Zhang (zhangyy003@igsnr.ac.cn) and

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

17

18 **Abstract**

19 Integrated water system modeling is a reasonable approach **for obtaining scientific**
20 **understanding of severe water crises faced in the world and for promoting** the
21 implementation of integrated river basin management. The time variant gain model
22 (TVGM), which is a classic hydrological model, is based on the complex Volterra
23 nonlinear formulation. TVGM has **obtained** good performance **on** runoff simulation in
24 numerous basins, but is disadvantageous **in predicting** other water-related components.
25 In this study, TVGM was extended to an integrated water system model by coupling
26 multiple water-related processes in hydrology, biogeochemistry, water quality and
27 ecology, and by considering the interference of human activities. Parameter sensitivity
28 and autocalibration modules were developed to improve simulation efficiency. The

1 Shaying River Catchment, which is the largest, highly regulated and heavily polluted
2 tributary in the Huai River Basin of China, was selected as the case study area. The
3 key water-related components were simulated, including runoff, water quality,
4 nonpoint source (or diffuse source) pollutant load, and crop yield. Results showed that
5 the extended model simulated most components reasonably well. In particular, the
6 simulated daily runoff series at most regulated and less-regulated stations matched
7 well with the observations. The average correlation coefficient and coefficient of
8 efficiency between the simulated and observed runoffs were 0.85 and 0.70,
9 respectively. Both the simulated low and high flow events were improved when the
10 dam regulation was considered, except the low flow simulation at Zhoukou and
11 Huaidian Stations. The daily ammonium-nitrogen (NH₄-N) concentration, which is
12 used as a key index in the water quality evaluation in China, was well captured with
13 the average correlation coefficient of 0.67. Furthermore, the nonpoint source NH₄-N
14 load and the corn yields were simulated for each administrative region and the results
15 were reasonable compared with the data from the official report and statistical
16 yearbooks, respectively. This study is expected to provide a scientific basis for the
17 implementation of such a modeling practice for integrated river basin management.

18

19 **1. Introduction**

20 Severe water crises are global issues that have emerged as a consequence of the rapid
21 development of the social economy, and include flooding, water shortages, water
22 pollution and ecological degradation. These issues hinder the equitable development
23 of regions by compromising the sustainability of vital water resources and ecosystems
24 (Gleick, 1997). It is impossible to address these water-related problems within a
25 single scientific discipline (e.g., hydrology, hydraulics, water quality or aquatic
26 ecology) because of the complicated interconnections among the physical, chemical
27 and ecological components of an aquatic ecosystem (Kindler, 2000; Biswas, 2004;
28 Paola *et al.*, 2006). The paradigm of integrated river basin management may be a
29 sensible solution at basin scale by focusing on the coordinated management of water
30 resources in terms of social-economy, water quality and ecosystems. Correspondingly,
31 integrated water system modeling is a reasonable practice for the simultaneous
32 simulation of water-related components (flow regimes, nutrient loss, sediment and

1 water quality), and is also an effective tool for supporting water resource allocation,
2 environmental flow management and river restoration.

3 Integrated water system modeling has gained popularity in the last decades due to the
4 rapid development of water-related sciences, computer sciences, earth observation
5 technologies and the availability of open data. Moreover, models naturally tend to
6 grow in complexity (Beven, 2006). The hydrological cycle has been widely accepted
7 as a critical linkage among physical, biogeochemical and ecological processes, and
8 energy fluxes at the basin scale (Wigmosta *et al.*, 1994; Singh and Woolhiser, 2002;
9 Burt and Pinay 2005). For example, the physiological and ecological processes of
10 vegetation affect evapotranspiration, soil moisture distribution and infiltration, and
11 nutrient absorption and movement. On the contrary, soil moisture and nutrient content
12 directly affect crop growth. Overland flow is a carrier of the pollutants to water bodies.
13 Therefore, the variation patterns of water-related components and their causes at the
14 basin scale should be analyzed by coupling all of these processes to capture the
15 interactions and feedback between individual cycles. Furthermore, multidisciplinary
16 research provides an effective way to enable breakthroughs in water system modeling
17 by integrating the mature theories of water-related disciplines (e.g., accumulated
18 temperature law for phenological development, Darcy's law for groundwater flow,
19 Saint-Venant equation for flow routing, balance equation for mass and momentum,
20 Richards' equation for unsaturated zone, Horton theory for infiltration,
21 Penman-Monteith equation for evapotranspiration), with support from abundant data
22 sources (e.g., high-resolution spatial information data, chemical and isotopic data
23 from field experiments) (Singh and Woolhiser, 2002; Kirchner, 2006).

24 Several models have been developed by using mature models of different disciplines
25 since the 1980s (Di Toro *et al.*, 1983; Brown and Barnwell 1987; Johnsson *et al.*,
26 1987; Hamrick, 1992; Li *et al.*, 1992; Abrahamsen and Hansen, 2000; Tattari *et al.*,
27 2001). A general review and discussion can be found in Singh and Woolhiser (2002).
28 Owing to the complexity of the integrated system and the scale conflicts between
29 different models, most existing models focus only on one or two major water-related
30 processes. According to the model orientation, existing models can be categorized into
31 three major classes. (i) Hydrological models emphasize the rainfall-runoff relationship
32 and link with some dominating water quality and biogeochemical processes. As a
33 result, these models generally show satisfactory performance in simulating major

1 hydrological processes. Examples of widely accepted models include TOPMODEL
2 (Beven and Kirkby, 1979), SHE (Abbott *et al.*, 1986), HSPF (Bicknell *et al.*, 1993),
3 VIC (Liang *et al.*, 1994), ANSWERS (Bouraoui and Dillaha, 1996), HBV-N
4 (Arheimer and Brandt 1998, and 2000), and HYPE (Lindström *et al.*, 2010). (ii) Water
5 quality models focus on the migration and transformation processes of pollutants in
6 water bodies. The models can obtain the high spatial and temporal resolutions of
7 water quality variables in river systems by adopting multi-dimensional dynamic
8 equations. However, these models simulate the overland processes of water and
9 pollutants with difficulty. Typical models include WASP (Di Toro *et al.*, 1983),
10 QUAL2E (Brown and Barnwell 1987) and EFDC (Hamrick, 1992). (iii)
11 Biogeochemistry models have advantages in the simulation of the physiological and
12 ecological processes of vegetation, and the vertical movements of nutrients and water
13 in soil layers at the field or experimental catchment scales. Nevertheless, these models
14 lack accurate hydrological features (Deng *et al.*, 2011). Thus these models are hard to
15 simulate the movements of water, nutrients and their losses along flow pathways in
16 the basin. Examples of the biogeochemistry models include SOILN (Johnsson *et al.*,
17 1987), EPIC (Sharpley and Williams, 1990), DNDC (Li *et al.*, 1992), Daisy
18 (Abrahamsen and Hansen, 2000), and ICECREAM (Tattari *et al.*, 2001). Therefore,
19 most models usually achieve good performances only on the oriented processes, and
20 only approximate results for other processes outside of the model's focus.

21 SWAT is a typical integrated water system model that can simulate most water-related
22 processes over long period at large scales (Arnold *et al.*, 1998). The model structure
23 and functions of SWAT are considered landmark in the field of water system modeling.
24 However, not all water-related processes can be well captured in practice because of
25 the inaccurate descriptions of certain processes, such as the daily simulations of
26 extreme flow events (Borah and Bera, 2004), soil nitrogen and carbon (Gassman *et al.*,
27 2007), and applicability in regulated basins (Zhang *et al.*, 2012). Particularly, SWAT
28 applies two alternative approaches to simulate surface runoff, namely, the empirical
29 soil conservation service (SCS) curve number method and the conceptual
30 Green-Ampt infiltration model. The SCS equation is usually prioritized, but the
31 applicability of the curve number is questioned (Rallison and Miller 1981). The
32 Green-Ampt infiltration model is usually limited to the simulation of flow events at
33 micro-scales (King *et al.*, 1999). Furthermore, SWAT has difficulty in accurately

1 capturing the complicated dynamic processes of soil nitrogen and carbon compared
2 with other biochemistry models (Gassman *et al.*, 2007). Therefore, several modified
3 versions of SWAT were developed, such as SWIM which is based on the hydrological
4 components from SWAT and the nutrient cycle components from the MATSALU
5 model (Krysanova *et al.*, 1998), and SWAT-N which extend SWAT by adopting
6 DNDC (Polhert *et al.* 2006, 2007).

7 The time variant gain model (TVGM) proposed by Xia (1991) is a lumped
8 hydrological model based on the hydrological data from many different scale basins
9 all over the world. In TVGM, the rainfall-runoff relationship is considered nonlinear
10 with surface runoff coefficient that varies over time and is significantly affected by
11 antecedent soil moisture. TVGM has strong mathematical basis because this nonlinear
12 relationship is transformed into a complex Volterra nonlinear formulation. Wang *et al.*
13 (2002) extended TVGM to a distributed model (DTVGM) by taking advantages of
14 better computing facilities and available data sources. DTVGM is currently widely
15 used in many basins with different scales and different climate zones to investigate the
16 effect of human activities and climate change on runoffs, and obtained good
17 simulation performances (Xia *et al.*, 2005; Wang *et al.*, 2009). However, DTVGM is
18 confined to hydrological cycle studies and cannot be applied to the integrated river
19 basin management because other water-related processes are not included, such as the
20 water quality processes, ecological processes, soil biogeochemical processes.

21 Motivated by the requirements for the integrated river basin management, integrated
22 water system model should be further developed to produce reasonable simulation in
23 both water quantity and quality processes in the real basins with possibility to
24 simulate more water-related processes, such as soil biogeochemistry, crop growth. In
25 this study, we extend the simulation functions of DTVGM as an integrated water
26 system model and improves the modeling practice of water-related components. Our
27 specific objectives are as follows: (1) to integrate the detailed interactions and
28 linkages among hydrological, water quality, soil biogeochemical and ecological
29 processes, as well as the prevalent regulations of water projects at the basin scale; (2)
30 to couple robust parameter analysis approaches with the integrated water system
31 model to improve model performances; (3) to examine the applicability of the
32 extended model on key water-related components in complex basins, e.g., flow

1 regimes, nonpoint source (or called diffuse source) pools of nutrients, water quality
2 variables in water body and crop yield.

3

4 **2. Methods and material**

5 **2.1 Model framework**

6 The proposed model includes seven major modules, namely the hydrological cycle
7 module (HCM), soil biochemical module (SBM), crop growth module (CGM), soil
8 erosion module (SEM), overland water quality module (OQM), water quality module
9 of water bodies (WQM) and dam regulation module (DRM). The parameter analysis
10 tool (PAT) is also designed for model calibration. The exterior exchange components
11 connecting different modules are given in Figure 1. The more detailed descriptions of
12 each module and its interactions with other modules are given in sub-sections 2.1.1 to
13 2.1.5. The main equations of each process are deferred to the appendix and
14 supplementary materials for readers who are interested in the mathematical details.

15 The extended model is based on the hypothesis that the cycles of water and
16 nutrients (N, P and C) are inseparable and act as the critical linkages among all the
17 modules. The model takes full advantages of the existing models, i.e., the powerful
18 interconnection of the hydrological model with other processes at a large spatial scale,
19 the elaborative description of the ecological model on nutrient vertical movement in
20 soil layers, and the elaborative description of the water quality model on nutrient
21 movements along river networks. First, several key components that are simulated by
22 the hydrological module, such as evapotranspiration, soil moisture, and flow, serve as
23 critical linkages in all the modules (Section 2.1.1). Second, the soil biochemical
24 processes determine the nutrient loads absorbed in the crop growth process (CGM)
25 and migrated into water bodies as the nonpoint pollutant source (OQM and WQM).
26 The accurate descriptions of soil biochemical processes are helpful in improving the
27 simulation of water quality processes in responding to agricultural management
28 (Section 2.1.2). Third, the hydrological module provides a function to describe the
29 spatial connections among spatial calculation units to simulate the overland and
30 in-stream movements of water and nutrients at the basin scale (Sections 2.1.1 and
31 2.1.3).

1 **2.1.1 Hydrological cycle module (HCM)**

2 TVGM is adopted to calculate the surface runoff yields for different land-use areas,
3 such as forests, grasslands, water bodies, urban areas, unused land, paddy land, and
4 dryland agriculture. The potential evapotranspiration is calculated by using
5 Hargreaves' method (Hargreaves and Samani, 1982) because only the widely
6 available daily maximum and minimum temperature data are used. The actual plant
7 transpiration is expressed as a function of potential evapotranspiration and leaf area
8 index, whereas soil evaporation is expressed as a function of potential
9 evapotranspiration and surface soil residues (Neitsch *et al.*, 2011). The interflow and
10 baseflow have linear relationships with the soil moisture in the upper and lower layers,
11 respectively (Wang *et al.*, 2009). The infiltration from the upper to the lower soil
12 layers is calculated by using storage routing methodology (Neitsch *et al.*, 2011). The
13 Muskingum method or kinetic wave equation is used for river flow routing.

14 Figure 2 shows that shallow soil water from the hydrological cycle module is one of
15 the major factors that connect the crop growth module (to control crop growth) and
16 the soil biochemical module (to control the vertical migration and reaction of nutrients
17 in the soil profile). Plant transpiration is also linked to the soil biochemical module (to
18 provide energy for vertical migration of nutrients in the soil profile). The surface
19 runoff is linked to the soil erosion, while the overland flow is connected to the
20 overland water quality modules (to drive migration of nutrients and sediment along
21 flow pathways), and water quality module for runoff routing in water bodies (rivers
22 and lakes). Moreover, the hydrological cycle module provides the inflows of
23 individual dams or sluices for the dam regulation module.

24 **2.1.2 Ecological process modules**

25 The ecological process modules contain the soil biochemical module and the crop
26 growth module. The crop growth and soil biochemical processes directly affect the
27 soil moisture, evapotranspiration, the nutrient transformation and loss from soil layers.
28 Therefore, the model incorporates the water cycle, nutrient cycle, crop growth, and
29 their key linkages.

1 2.1.2.1 Soil biochemical module (SBM)

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

9 **Soil C and N cycle.** The daily step decomposition and denitrification sub-models in
10 DNDC are adopted to simulate the biogeochemical processes of C and N in the soil
11 profile at field scale according to the crop pattern in actual situations (Li *et al.*, 1992).
12 The decomposition and other oxidation processes are the dominant microbial
13 processes in the aerobic condition. The three conceptual organic C pools are: the
14 decomposable residue, microbial biomass and stable (humus). The decomposition of
15 each C pool is treated as the first-order decay process with the individual
16 decomposition modified by soil temperature and moisture, clay content, and C: N
17 ratio. The major simulated processes of decomposition under aerobic condition are
18 mineralization, immobilization, ammonia (NH₃) volatilization and nitrification.
19 Mineralization or immobilization of mineral N (NH₄⁺ and NO₃⁻) is determined by the
20 flow rate of SOC pools. NH₃ volatilization is controlled by the simulated NH₄⁺
21 concentration, clay content, pH, soil moisture and temperature. NH₄⁺ is oxidized to
22 NO₃⁻-N during nitrification and nitrous oxide (N₂O) is emitted into the air during
23 nitrification. Denitrification occurs under anaerobic condition, which is controlled by
24 soil moisture, temperature, pH, and dissolved soil organic carbon content. The
25 detailed descriptions are given in Appendix B and Li *et al.*, 1992.

26 **Soil P cycle.** The major processes of soil P cycle are simulated based on the study of
27 Horst *et al.* (2001). Six P pools are considered, namely, three organic pools (stable and
28 active pools for plant uptake, fresh pool associated with plant residue) and three
29 mineral pools (soluble mineral, stable and active pools) as the consequence of
30 mineralization, decomposition and sorption (Horst *et al.*, 2001). The P dynamics
31 processes were considered in Horst *et al.* (2001) and Neitsch *et al.* (2011) through

1 modeling the P release from fertilizer, manure, residue, microbial biomass, humic
2 substances, and P sorption by plant uptake.

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

7 2.1.2.2 Crop growth module (CGM)

8 The crop growth module is developed based on EPIC crop growth model (Hamrick,
9 1992), which uses the concepts of daily accumulated heat units on phenological crop
10 development, Monteith's approach for potential biomass, harvest index for
11 partitioning grain yield, and stress adjustments for water, temperature and nutrient (N
12 and P) availability in the root zone of the soil layers. It simulates total dry matter, leaf
13 area index, root depth and density distribution, harvest index, and nutrient uptake, etc
14 (Williams *et al.*, 1989; Sharpley and Williams, 1990). The crop respiration and
15 photosynthesis drive the vertical movements of water and nutrient. In the crop growth
16 module, the output of leaf area index is the main factor connecting the hydrological
17 cycle module (to control the transpiration) and the crop residue left in the fields is the
18 main source of organic matters (C, N and P) connecting to the soil biochemical
19 module for soil biochemical processes, to the overland water quality module, and to
20 the soil erosion module as one of the five constraint factors (Figure 3b).

21 **2.1.3 Water quality process modules**

22 The water quality process modules focus on the migration and transformation of water
23 quality variables (e.g., sediment, different forms of nutrients, chemical oxygen
24 demand: COD) along the flow pathways in the land surface and river system. The
25 main modules are the soil erosion module for the sediment yield, the overland water
26 quality module for the nonpoint source pollutant loss and migration from the soil
27 layers to water bodies, and the water quality module for the migration and
28 transformation of pollutants in water bodies (point and nonpoint source loads).

1 2.1.3.1 Soil erosion module (SEM)

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

8 2.1.3.2 Overland water quality module (OQM)

9 This module simulates the overland loss and migration load of nonpoint source
10 pollutants (e.g., sediment, insoluble and soluble nutrients, COD) for the water quality
11 module of water bodies (Figure 4b). The main sources include the nutrient loss from
12 soil layers and urban areas, the farm manure from livestock in rural areas. The
13 nutrient loss from soil layers, as the primary nonpoint source in most catchments, is
14 determined by the overland flow and sediment yield (Williams *et al.*, 1989) and the
15 other sources are estimated by using the export coefficient method (Johnes, 1996).
16 The overland migration processes contain the soluble pollutant migration with
17 overland flow and the insoluble pollutant migration with sediment. All of these
18 processes occur along the overland transport paths.

19 2.1.3.3 Water quality module of water bodies (WQM)

20 Point and nonpoint pollution sources are considered in the extended model. Point
21 sources are directly added to the surface water in the model according to their
22 geographic positions. Common point sources are urban water treatment plants or
23 industrial plants.

24 Two modules are designed for different types of water bodies, such as, the in-stream
25 water quality module and the water quality module of water impounding (reservoir or
26 lake). The enhanced stream water quality model (QUAL-2E) (Brown and Barnwell
27 1987), is a comprehensive and versatile stream model that simulates the longitudinal
28 movement and transformation of water quality variables in the branch streams. The
29 model is centered at dissolved oxygen (DO) and can simulate up to 15 water quality
30 variables including water temperature, DO, sediment, different forms of nutrients (N

1 and P), COD (Figure 4c). The model is solved at the sub-basin scale rather than at the
2 fine grid scale to maintain spatial consistent with the hydrological cycle module. The
3 water quality outputs provide the water quality boundary of dams or sluices for the
4 dam regulation module. The water quality module of water impounding assumes that
5 water body is at the steady state and focuses on the vertical interaction of water
6 quality. The main processes include water quality degradation and settlement,
7 sediment resuspension and decay.

8 **2.1.4 Dam regulation module (DRM)**

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

14 Given that different types of dams and sluices are likely to show completely different
15 regulation behaviors, we try to reproduce the common functionalities for the flood
16 control or water supply dams in this module. Three methods are proposed for
17 calculating the water storage and outflow of dams or sluices, namely, the measured
18 outflow, controlled outflow with target water storage, and the relationship between
19 outflow and water storage volume. The first method requires users to provide the
20 measured outflow series during the simulation period. The second method simplifies
21 the regulation rule of dam or sluice for long-term analysis by assuming that water is
22 stored according to the usable water level during non-flooding season and the flood
23 control level during flooding season. The redundant water is discharged. This method
24 requires the characteristic parameters of dams or sluices including water storage
25 capacities of dead, usable, flood control and maximum flood levels and the
26 corresponding water surface areas. The third method is based on relationships among
27 water level, water surface area, storage volume and outflow according to the design
28 data of dam, or long-term observed data (Zhang *et al.*, 2013) (Appendix C).

29 **2.1.5 Parameter analysis tool (PAT)**

30 Parameter sensitivity analysis and calibration are important steps in the applications of
31 highly parameterized models and are treated seriously, particularly for integrated

1 water system models (Mantovan and Todini, 2006; Mantovan *et al.* 2007; McDonnell
2 *et al.*, 2007). In the model, 78 lumped and 104 distributed parameters involve the
3 hydrological, ecological and water quality processes. The distributed parameters are
4 divided into 46 overland parameters, 18 stream parameters and 40 parameters of water
5 projects (only for the sub-basin with reservoir or sluice) according to their spatial
6 distribution. These parameter values are determined by the properties of overland
7 landscape and soil, stream patterns, and water projects, respectively. Different spatial
8 calculation units share many common parameter values if their properties are the
9 same. The sensitive parameters and their ranges are determined first to reduce the
10 parameter dimensions by parameter sensitivity analysis or according to user
11 experiences. Their values are then calibrated by auto-optimization methods or manual
12 adjustment to achieve optimal model performances, whereas the insensitive
13 parameters remain constant.

14 Owing to high parameterization, an optimum result is hard to achieve by subjective
15 selection and judgment; therefore, PAT is designed for parameter sensitivity analysis,
16 autocalibration and model performance evaluation. Moreover, PAT is a part of the
17 extended model (Figure 5). The algorithms include the parameter sensitivity method
18 (latin hypercube one factor at a time: LH-OAT) (van Griensven *et al.*, 2006),
19 auto-optimization methods such as particle swarm optimization (PSO) (Kennedy,
20 2010), genetic algorithm (GA) (Goldberg 1989) and shuffled complex evolution
21 (SCE-UA) (Duan *et al.*, 1994). The five indices are provided to evaluate model
22 performance including bias (*bias*), relative error (*re*), root mean square error (*RMSE*),
23 correlation coefficient (*r*), and coefficient of efficiency (*NS*). These methods and
24 indices can be selected by users on the basis of their specific requirements.

25 The interconnections between PAT and other modules are the parameter groups that
26 need to be analyzed in sensitivity analysis and optimization, and the objective
27 functions described by the evaluation indices of model performance. PAT randomly
28 samples the parameter values from the multi-dimensional parameter spaces to the
29 extended model to obtain the values of the objective function. For parameter
30 sensitivity analysis, the sensitivity index of individual parameters is evaluated by
31 comparing the variation of the objective function value along with the changes of
32 parameter values. For parameter autocalibration, the good parameter groups are kept

1 or updated following the particular criteria provided by the auto-optimization methods
2 until the convergence criteria or maximum iterations are achieved.

3

4 **2.2 Model operation**

5 **2.2.1 Multi-scale solution**

6 The spatial heterogeneities of basin attributes and the different time scales used in
7 individual processes cause inconsistent spatial and temporal scales in model
8 integration (Blöschl and Sivapalan, 1995; Sivapalan and Kalma, 1995; Singh and
9 Woolhiser, 2002). For the spatial scale, three levels of spatial calculation units are
10 designed in the model, namely, sub-basin, land-use and crop from largest to smallest.
11 These units are minimum polygons with similar hydrological properties, land-use
12 types and agriculture crop cultivation patterns. The sub-basins are defined on the basis
13 of DEM, the positions of gauges and water projects (dams or sluices), and are used in
14 the hydrological cycle module (e.g., flow routing in both land and in-stream),
15 overland water quality module, water quality module of water bodies and dam
16 regulation module. Seven specific land-use units of each sub-basin are partitioned by
17 the land-use classification (i.e., forest, grassland, water, urban, unused land, paddy
18 land and dryland agriculture). The related modules are the hydrological cycle module
19 (e.g., water yield, infiltration, interception and evapotranspiration) and soil erosion
20 module. Moreover, several specific land-use units (paddy land and dryland agriculture,
21 forest, grassland), where agricultural activities usually occur, are divided further into
22 the crop units for detailed analysis of the impact of agricultural management on water
23 and nutrient cycles. In the current version of the model, these four land-use units are
24 divided into 10 specific categories of crop units: fallow for all these land-use units;
25 grass for grassland unit; fruit tree and non-economic tree for forest unit; early rice and
26 late rice for paddy unit; spring wheat, winter wheat, corn, and mixed dry crop for
27 dryland agriculture unit. The crop unit category of a specific land-use pattern varies
28 depending on crop cultivation structure and timing. The related modules are the soil
29 biochemical module and the crop growth module. All of the outputs of the crop unit
30 are summarized at the land-use unit scale, or sub-basin scale on the basis of the area
31 percentage of different units.

1 For the temporal scale, it is practical to use a daily time-step as this is consistent with
2 the underlying rainfall-runoff module and the data availability. The sub-daily time
3 scale may improve the performance in some modules (e.g., SEM, WQM). However,
4 most observations (e.g., climate data sets, soil nutrient availability, and water quality
5 concentrations), are at daily time scale, thus leading to potential uncertainty or
6 inability to downscale the observations into a sub-daily time scale. Moreover, the
7 sub-daily module will increase model complexity compared to regional simulations,
8 as it is quite hard to obtain the information at the regional scale. Thus most processes
9 need to be simplified to fit regional research. Linear or nonlinear aggregation
10 functions are used to transform different time scales to daily scale (Vinogradov *et al.*,
11 2011), such as exponential relation for flow infiltration and overland flow routing
12 processes, soil erosion processes (A5, A6 and S32 in the appendices), and
13 accumulative relation for the crop growth process (S7 in the supplementary material).

14 **2.2.2 Basic datasets and spatial delineation**

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

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

26 **2.3 Study area and model testing**

27 In this study, the extended model was applied to a highly regulated and heavily
28 polluted river basin of China. The simulated components contained daily runoff and
29 water quality concentrations at some river cross-sections, spatial patterns of nonpoint
30 source pollutant load and crop yield at sub-basin scale.

1 **2.3.1 Study area**

2 The Shaying River Catchment (112°45'~113°15'E, 34°20'~34°34'N), which is the
3 largest sub-basin of the Huai River Basin in China, is selected as the study area
4 (Figure 6a). The drainage area is 36,651 km² with a mainstream of 620 km. The
5 average annual population (2003-2008) (Figure 6b) is 32.42 million, with rural
6 population of 23.70 million. The average annual stocks (Figure 6c) are 8.30 million
7 (big animals) and 178.42 million (poulties). The average annual use of chemical
8 fertilizer (Figure 6d) is 1.55 million ton (N: 38%-51%, P: 16%-25%, K: 7%-12% and
9 others: 16%-35%). The basin is located in the typical warm temperate, and
10 semi-humid continental climate zone. The annual average temperature and rainfall are
11 14-16°C and 769.5 mm, respectively. The Shaying River is the most seriously polluted
12 tributary with a pollutant load contribution of over 40% in the whole Huai River and
13 is usually known as the water environment barometer of the Huai River mainstream.
14 To reduce flood or drought disasters, 24 reservoirs and 13 sluices, whose regulation
15 capacities are over 50% of the total annual runoff, have been constructed and
16 fragmented river into several impounding pools.

17 **2.3.2 Model setup**

18 All data sets for model setup and calibration were collected from the government
19 bureaus, official books or scientific references. The detailed descriptions were
20 presented in Tables S2 and S3 of the supplementary material. The Shaying River
21 Catchment was divided into 46 sub-basins. According to the land-use classification
22 standard of China (CNS,2007), the main land use types were dryland agriculture
23 (84.04%), forest (7.66%), urban (3.27%), grassland (2.68%), water (1.43%), paddy
24 (0.91%), and unused land (0.01%).The soil input parameters (the contents of sand,
25 clay and organic matter) were calculated on the basis of the percentage of soil types in
26 each sub-basin. The main crops were early rice and late rice in the paddy land, and
27 winter wheat and corn in the dryland agriculture. Their main agricultural management
28 schemes (fertilize, plant, harvest and kill) were summarized by field investigation in
29 the studies of Wang *et al.*, (2008) and Zhai *et al.* (2014) (Table S3). Crop rotation and
30 its management schemes were considered in the model by setting the start time, the
31 duration of management and the fertilizer amounts. Only two fertilizations (base and
32 additional fertilization) were considered in the model during the complete growth

1 cycle of a certain crop. The areas of sub-basin, land-use and crop units ranged from
2 46.48 km² to 3771.15 km², from 0.04 km² to 2762.5 km², and from 3.73 km² to 2762.5
3 km², respectively.

4 The daily data series at 65 precipitation stations and six temperature stations were
5 interpolated to each sub-basin from 2003 to 2008 by using the inverse distance
6 weighting method and the nearest-neighbor interpolation method, respectively. The
7 social and economic data (e.g., population and livestock in the rural area, chemical
8 fertilizer amount) were calculated for each sub-basin on basis of the area percentage.

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

14 **2.3.3 Model evaluation**

15 NH₄-N concentration is one of the widely used indices in assessing water quality
16 condition in China (CSEPA, 2002). Thus, both the observation series of daily runoff
17 and NH₄-N concentration were used to calibrate the model parameters. There were
18 five regulated stations (Luohe, Zhoukou, Huaidian, Fuyang and Yingshang) and one
19 less-regulated station (Shenqiu) which is the downstream station situated far from
20 water projects.

21 We selected LH-OAT for parameter sensitivity analysis and SCE-UA for parameter
22 calibration in PAT. The initial parameter values were randomly preset from the value
23 ranges determined by their physical characteristics. The evaluation indices are *bias*, *r*
24 and *NS* as a demonstration of model performance of the extended model. However,
25 *NS* is sensitive to extreme value, outlier and number of data points and is not
26 commonly applied in environmental sciences (Ritter and Muñoz-Carpena, 2013).
27 Thus *NS* was not used to evaluate the NH₄-N concentration simulation. Furthermore,
28 given that the real observed yields of nonpoint pollutant loads and crops were hard to
29 collect for the whole catchment (Chen *et al.*,2014), their simulations were only
30 evaluated preliminarily by using *bias* according to the statistical results from official
31 reports or statistical yearbooks (Wang, 2011; Henan Statistical Yearbook, 2003, 2004
32 and 2005).

1 The model calibration was conducted step-by-step as follows. Hydrological
 2 parameters were calibrated first against the observed runoff series at each station from
 3 upstream to downstream, and then water quality parameters against the observed
 4 $\text{NH}_4\text{-N}$ concentration series. The calibration and validation periods were from 2003 to
 5 2005 and from 2006 to 2008, respectively. To reduce the dimensions of the calibration
 6 problem, we restricted SCE-UA to calibrate only the sensitive parameters as defined
 7 by LH-OAT, whereas other parameters remained constants. Weighting method was
 8 usually used to comprehensively handle different objectives (Efstratiadis and
 9 Koutsoyiannis, 2010). In this study, these objective functions were simply aggregated
 10 to single objectives (f_{runoff} and $f_{\text{NH}_4\text{-N}}$) as

$$11 \quad \begin{cases} f_{runoff} = \min[(|bias| + 2 - r - NS) / 3] \\ f_{\text{NH}_4\text{-N}} = \min[(|bias| + 1 - r) / 2] \end{cases} \quad (1)$$

12 because the case study was only a demonstration of the model performance.

13 Moreover, the effect of dam regulation in the integrated water system models was
 14 considered because of the high regulation in most rivers. The dam and sluice
 15 regulation usually disturbs the intra-annual distribution of flow events, such as,
 16 flattening high flow and increasing low flow. The simulation performances of high
 17 and low flow were separately evaluated, and the effectiveness of the DRM was tested
 18 by comparing the simulation with and without considering the regulation. The high
 19 and low flow was determined by the cumulative distribution function (CDF). A
 20 threshold of 50% was used for easy presentation, i.e., the flow was treated as high
 21 flow (or low flow) if its percentile was greater than (or smaller than) the threshold.

22

23 **3. Results**

24 **3.1 Parameter sensitivity analysis**

25 Nine sensitive parameters were detected for runoff simulation (Table 2): soil related
 26 parameters W_{fc} (field capacity), W_{sat} (saturated moisture capacity), K_r (interflow yield
 27 coefficient) and K_{sat} (steady state infiltration rate); TVGM parameters g_1 (basic surface
 28 runoff coefficient) and g_2 (influence coefficient of soil moisture) for surface runoff
 29 calculation; ground water recharge parameters K_g (baseflow yield coefficient) and T_g
 30 (delay time for aquifer recharge); and adjusted factor K_{ET} of evapotranspiration. All of

1 these parameters controlled the main hydrological processes, in which soil water and
2 evapotranspiration processes were distinctly important and explain 54.3% and 23.2%
3 of the runoff variation, respectively.

4 For NH₄-N concentration simulation, more than 90% of observed NH₄-N
5 concentration variations were explained by 14 sensitive parameters categorized into
6 hydrological (59.28% of variation), NH₄-N (20.65% of variation) and COD (12.34%
7 of variation) related parameters. The main explanations were that hydrological
8 processes provided the hydrological boundaries that affected the nonpoint source
9 pollutant load into rivers, the degradation and settlement processes of NH₄-N in water
10 bodies (van Griensven *et al.*, 2002). NH₄-N concentration was further influenced by
11 the settling and biological oxidation processes. Moreover, it was a competitive
12 relationship between COD and NH₄-N to consume DO of water bodies in a certain
13 limited level (Brown and Barnwell, 1987).

14 **3.2 Hydrological simulation**

15 The simulations fitted the observations well at all the stations from the midstream to
16 downstream (Figure 7 and Table 3). The *biases* were very close to 0.0 at all the
17 regulated stations except Zhoukou with an underestimation (0.24 for calibration and
18 0.41 for validation) and Luohe with an overestimation (-0.52 for validation). The
19 obvious biases were caused by the calibration for obtaining the optimal solution of the
20 average of three evaluation indices, rather than the *bias* only. The *r* values ranged
21 from 0.75 (Luohe for validation) to 0.92 (Yingshang for calibration) with an average
22 value of 0.85, whereas the *NS* values ranged from 0.51 (Luohe for validation) to 0.84
23 (Yingshang for calibration) with an average value of 0.70. The results of the regulated
24 stations were a little worse than those of the less-regulated station (Shenqiu) owing to
25 the regulation.

26 By comparing the simulation results with the observations from 2003 to 2008, we can
27 see that the high and low flows were usually overestimated at all stations if the model
28 did not consider the regulations (Figure 8). Except the high flow events at Zhoukou,
29 both high and low flow events at all the stations were simulated well when the dam
30 and sluice regulation was considered (Table 4). The best fitting was at Fuyang,
31 particularly for the high flow simulation (*bias*=0.10, *r*=0.89 and *NS*=0.78). From
32 unregulation to regulation settings, the improvements measured by *f_{runoff}* ranged from

1 -0.08 (Zhoukou) to -0.29 (Huaidian) for high flow simulation, from -0.05 (Zhoukou)
2 to -0.31 (Huaidian) for average flow simulation, and from -1.97 (Fuyang) to -3.91
3 (Yingshang) for low flow simulation except Zhoukou (1.28). The improvements in the
4 performance of low flow simulations were the most obvious. However, their
5 performance still need to be improved further, particularly the underestimation at
6 Zhoukou and Huaidian. The possible reasons were that, on the one hand, the applied
7 evaluation indices (r and NS) were known to emphasize high flow simulation rather
8 than the low flow simulation (Pushpalatha *et al.*, 2012) and the objective of
9 autocalibration was to obtain the optimal solution for the average of three evaluation
10 indices rather than the *bias* only. The slight sacrifice of *bias* improves the overall
11 simulation performance evaluated by these three indices. On the other hand, the dam
12 regulation module still could not fully capture the low flow events.

13 Furthermore, the model performances of monthly flows were even better, particularly
14 for r and NS . The values of r ranged from 0.87 (Luohe for both calibration and
15 validation) to 0.95 (Fuyang for calibration) with the mean of 0.92, whereas the values
16 of NS ranged from 0.67 (Luohe for validation) to 0.94 (Shenqiu for validation) with
17 the mean of 0.80. Zhang *et al.* (2013) reproduced the long-term monthly flows by
18 SWAT at the same stations. Compared with existing results, the extended model
19 improved the flow simulations at the downstream stations although it became a little
20 worse at the upstream stations (Luohe and Zhoukou for calibration). In particular, the
21 water volume and agreements with the observations (i.e., *bias* and NS) were well
22 captured.

23 3.3 Water quality simulation

24 The simulated concentrations matched well with the observations according to the
25 evaluation standard recommend by Moriasi *et al.* (2007) (Figure 9 and Table 5). The r
26 values of all stations were over 0.60 except Zhoukou (0.56 for validation), Yingshang
27 (0.49 for validation) and Shenqiu (0.41 for validation) with an average value of 0.67.
28 The *bias* of all stations were considered as “acceptable” with a range from -0.27
29 (Fuyang for validation) to 0.29 (Zhoukou for calibration). The best simulation was at
30 Luohe. The obvious discrepancies between the simulations and observations often
31 appeared in the period from January to May because of the poor simulation
32 performance of low flows. Although the *bias* between calibration and validation

1 changed markedly at Fuyang and Yingshang, the model performances were still
2 acceptable. The probable explanation was that the *bias* for corresponding runoff
3 simulations at these two stations also changed.

4 Compared with the results without the consideration of regulation, the simulation
5 results were significantly improved when the regulation was considered except at
6 Fuyang for calibration. The decrease in f_{NH_4-N} value ranged from 0.10 (Huaidian for
7 calibration) to 0.49 (Zhoukou for validation) although it was increased slightly at
8 Fuyang for calibration (0.02). The regulation of dams and sluices played an important
9 role in the water quality simulation. In the upper stream of Shaying River, the flow
10 was small and the pollutant concentration decreased obviously because of the
11 degradation and settlement of large water storage. In the downstream of Shaying
12 River, the NH_4-N concentration increased because of the pollutant accumulation and
13 the decreasing flow of dams and sluices owing to the regulation (Zhang *et al.*, 2010).
14 Therefore, the simulated concentrations without regulation were usually
15 overestimated or higher than the simulation with regulation at the upstream stations
16 (Luohe and Zhoukou), however, these concentrations were underestimated at the
17 downstream stations (Huaidian, Fuyang and Yingshang). The largest simulation
18 difference between with and without the regulation consideration appeared at
19 Zhoukou.

20 The spatial pattern of average annual nonpoint source NH_4-N loads was shown in
21 Figure 10a. The modeled annual yield rates ranged from 0.048 t km⁻² year⁻¹ to 11.00 t
22 km⁻² year⁻¹ with an average value of 0.73 t km⁻² year⁻¹. The yield of each
23 administrative region was summarized from the sub-basin scale according to the area
24 percentage of sub-basins in each administrative region. Compared with the statistical
25 load of each administrative region based on the soil erosion, land use and fertilizer
26 amount in the official report (Wang, 2011), the *bias* of simulated nonpoint source load
27 in the whole region was 21.31% when the two regions with great *bias* (i.e., Fuyang
28 and Pingdingshan) were excluded as outliers. The high load yield regions were in the
29 middle of Pingdingshan, Xuchang, Zhengzhou, Fuyang and Zhoukou regions. The
30 spatial pattern was significantly correlated with the distribution of paddy area
31 ($r=0.506$, $p<0.001$) and rice yield ($r=0.799$, $p<0.001$) (Figures 10 b and c). The
32 fertilizer loss in the paddy areas might be the primary contributor to the nonpoint
33 source NH_4-N load, possibly because the average nitrogen loss coefficient in China

1 was just 30%-70% in the paddy areas, which was higher than that in dryland
2 agriculture (20%-50%) (Zhu, 2000; Xing and Zhu, 2000).

3 The observed average annual point source NH₄-N loads into rivers were
4 approximately 4.70×10⁴ t year⁻¹ in the Shaying River Catchment. This result was
5 summarized from the collected data for model input. The nonpoint source load
6 contributed 38.57% of the overall NH₄-N load on average from 2003 to 2005; this
7 value was slightly higher than the statistical results (29.37%) given in the official
8 report (Wang, 2011). Moreover, the contributions of the nonpoint source load at the
9 stations ranged from 31.72% (Huaidian) to 47.13% (Shenqiu). Compared with the
10 nonpoint source load of each administrative region in 2000, the simulated annual
11 loads tended to increase from 2003 to 2005 except in the Kaifeng region. The yields
12 of the Fuyang and Pingdingshan regions increased most evidently. The primary
13 pollution source in the Shaying River Catchment was still the point source; however,
14 the nonpoint source was also an important concern. The spatial characteristic of the
15 Shaying River Catchment was that the high nonpoint source contribution in the
16 upstream, and low nonpoint source contribution in the middle and downstream
17 because the point source load emission was usually concentrated in this region.
18 Therefore, compared with the results of Zhang *et al.* (2013), the overall simulation
19 performance of NH₄-N concentration was also improved significantly by considering
20 the detailed processes of nutrient in the soil layers.

21 3.4 Crop yield simulation

22 The simulated corn yield and its spatial pattern were shown in Figure 11. The average
23 annual yields were summarized at sub-basin scale and ranged from 0.08 to 326.95 t
24 km⁻² year⁻¹ with the average value of 76.84 t km⁻² year⁻¹. The yield of each
25 administrative region was further summarized and compared with the data from
26 statistical yearbooks from 2003 to 2005 (Henan Statistical Yearbook, 2003, 2004 and
27 2005) to test the simulation performance (See the inset of Fig.11). The high-yield
28 regions were Luohe, Fuyang and Zhoukou in the middle and down reaches, whose
29 primary land use were dryland agriculture (93.12%, 95.87% and 93.18%,
30 respectively). The yields of Luohe, Nanyang, Kaifeng regions were well simulated.
31 The total yield was underestimated in the whole basin with a bias of 19.93%. The
32 discrepancies might be caused by the boundary mismatch between the administrative

1 region and sub-basin, obvious spatial heterogeneities of human agricultural activities,
2 and the inaccurate cropping patterns in such huge regions. A high-resolution remote
3 sensing image and field investigation might improve the model performance.

4 **4. Discussion**

6 **4.1 Comparison with other models**

7 The evident difference between the extended model and existing models is that the
8 water-related processes are considered as an integrated water system rather than
9 isolated systems for individual processes. Integrated water system models provide
10 significant scientific basis to reasonably describe the complex water-related processes
11 in real basins. The model is also expected to simulate the critical water-related
12 components simultaneously, especially to improve the simulation performance of
13 water quantity and quality. The proposed model will be more beneficial to integrated
14 river basin management than the existing models.

15 The results of the extended model can compete with those of the existing studies
16 in the Huai River Basin. Several typical models have also been applied to simulate
17 runoff and water quality, e.g., SWAT for monthly runoff and water quality simulation
18 at the regulated stations (Zhang *et al.*, 2012), SWAT and Xinganjiang models for daily
19 runoff simulation at the unregulated upstream stations (Shi *et al.*, 2013), and DTVGM
20 for daily runoff simulation (Ma *et al.*, 2014). Different models showed generally
21 comparable performance of runoff or water quality simulations. For SWAT, the f_{runoff}
22 values were from 0.11 to 0.20 with a mean of 0.16 at the daily scale for the
23 unregulated stations (Shi *et al.*, 2013), and from 0.09 to 0.75 with a mean of 0.32 at
24 the monthly scale for the regulated stations (Zhang *et al.*, 2012); and f_{NH-N} values
25 ranged from 0.18 to 0.86 with a mean of 0.47 (Zhang *et al.*, 2012). For Xinganjiang
26 model, the f_{runoff} values were from 0.13 to 0.21 with a mean of 0.16 at the daily scale
27 for the unregulated stations (Shi *et al.*, 2013). For DTVGM, the f_{runoff} values were 0.14
28 and 0.21 at the daily scale in the calibration and verification periods, respectively at
29 Bengbu station. However, the extended model performed better than SWAT,
30 especially for the regulated runoff and water quality simulations (see Tables 4-6).
31 Moreover, both the Xinanjiang model and DTVGM can only simulate the flow series

1 at unregulated or less-regulated stations because they did not consider the dam
2 regulation in the model frameworks.

3

4 **4.2 Equifinality**

5 Untill now, the detailed water-related processes are still ambiguous and hard to be
6 deduced from strong physical foundations (Beven and Freer, 2001; Beven, 2006).
7 Empirical equations are still adopted to approximate the physical processes with
8 numerous unknown parameters, especially in the large scale models. Furthermore,
9 model extension requires additional numbers of parameters to be defined and
10 calibrated without additional observations (Beven, 2006). A single output variable of
11 models is usually associated with multiple processes and many parameters. For
12 instance, SWAT contains over 200 parameters (Arnold *et al.*, 1998) and DNDC has
13 nearly 100 parameters (Li *et al.*, 1992). Pohlert *et al.*, (2006) reported that six
14 hydrological and 12 N-cycle sensitive parameters were detected in SWAT-N for the
15 simulation of water flow and N leaching. In our extended model, nine and 14 sensitive
16 parameters were detected for runoff and NH₄-N simulation, respectively, including the
17 soil related parameters, surface runoff, and baseflow parameters, evapotranspiration
18 parameters, and the degradation and settling parameters of water quality variables
19 (Table 3). Therefore, most existing models are subject to equifinality, which was more
20 serious if more water -related processes were considered, or more sub-basins were
21 delineated for the distributed models.

22 The utilization of more information and multiple performance measures for single
23 predicted component would be helpful for alleviating the equifinality (Her and
24 Chaubey, 2015). In the extended model, the independent calibration and validation
25 data sets were specified in Table 1 and most widely-used measures of model
26 performances were also provided in the PAT. In the case study, we employed several
27 observation sources (e.g., runoff and water quality observations at different stations,
28 the nonpoint pollutant source load and crop yield data), and three measures to
29 evaluate model performance for the individual components (e.g., *bias*, *r* and *NS*).
30 Nonetheless, to make full use of the existing data in practice, parameter sensitivity
31 analysis would be an effective way to reduce dimensionality in model calibration, and
32 to focus only on critical processes and parameters that are sensitive to model outputs

1 (van Griensven *et al.*, 2006). Model autocalibration would be efficient to obtain the
2 optimal simulations from numerous samples in multi-dimensional parameter spaces.

3

4 **4.3 Model limitations**

5 The extended model still has several limitations:

6 (1). The mathematical descriptions of groundwater, crop growth processes and
7 agriculture management practices were still inaccurate. The current model focuses on
8 the detailed descriptions of hydrological and nutrient cycle in the soil layers and water
9 bodies and dam regulations. Satisfactory performances on water quantity and quality
10 simulation were obtained in our case study. However, the applications in the
11 simulation of the groundwater, nonpoint source pollution load, crop yield in the
12 agriculture regions could still be further improved. Moreover, the stratification of
13 water impounding in the water quality module should be considered if the high
14 resolution data of terrain elevation in the dams or lakes is available.

15 (2). High parameterization was an inevitable issue because of its all-inclusive
16 framework. The integrated water system model considers most of water-related
17 processes in the hydrological, ecology and water quality subsystems. All processes
18 were controlled by unmeasurable parameters because of their empirical and/or scale
19 dependent nature (Her and Chaubey, 2015). Although parameter sensitivity analysis
20 and calibration were widely used approaches to resolve the high parameterization
21 issue, the equifinality and parameter uncertainty were still inevitable because of the
22 insufficient observations, and the complex interactions among different subsystems.

23

24 **5. Conclusions**

25 In this study, an integrated water system model was primarily developed on the basis
26 of TVGM hydrological model to address the complex water issues emerging in the
27 basins. The model performance was demonstrated in the Shaying River Catchment of
28 China. The model would provide a reasonable tool for the effective water governance
29 by capturing several indicative components of water-related subsystems
30 simultaneously including the hydrological components (e.g., soil water and
31 evaporation, plant transpiration, runoff and water storage in the dams and sluices),

1 water quality components (e.g., nonpoint source load, water quality concentrations in
2 water bodies), and ecological components (crop yield) which could be calibrated if
3 observations are available. The case study showed that the simulated runoffs at most
4 stations fitted the observations well in the highly regulated Shaying River Catchment.
5 All the evaluation criteria were acceptable for both the daily and monthly simulations
6 at most stations. This model captured the variation of discontinuous daily NH₄-N
7 concentration and properly simulated the spatial patterns of nonpoint source pollutant
8 load and corn yield.

9 Owing to the heterogeneity of spatial data in large basins and insufficient observations
10 of individual subsystems, not all the results were acceptable and several processes
11 were still not well calibrated (low flow events, nonpoint source pollutant load, and
12 crop yield, etc.). The model could also be improved by further considering more
13 accurate humanity activities in the agricultural management, calibration of
14 multi-component and model uncertainty analysis because of the interactions and
15 tradeoffs among different processes. The over-parameterization and the reasonable
16 initial conditions of parameters would also be treated carefully in applications.
17 Advanced mathematic analysis technologies would benefit the future model
18 development, such as multi-objective optimization algorithm.

19

20 **Appendix A: Hydrological cycle module**

21 The basic water balance equation is

$$22 \quad P_i + SW_i = SW_{i+1} + Rs_i + Ea_i + Rss_i + Rbs_i + In_i \quad (A1)$$

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

27 E_s and E_p are determined by potential evapotranspiration (E_0 , mm), leaf area index
28 (LAI , m²/m²) and surface soil residues (rsd , t/ha) (Ritchie, 1972) as.

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

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

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

$$4 \quad R_s = g_1 (SW_u / W_{sat})^{g_2} \cdot (P - In) \quad (A3)$$

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

7 The interflow (R_{ss} , mm) and baseflow (R_{bs} , mm) have linear relationships with the
8 soil moisture of the upper and lower layers, respectively (Wang *et al.*, 2009) as

$$9 \quad \begin{cases} R_{ss} = k_{ss} \cdot SW_u \\ R_{bs} = k_{bs} \cdot SW_l \end{cases} \quad (A4)$$

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

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

$$14 \quad \begin{cases} W_{inf} = (SW_u - W_{fc}) \cdot [1 - \exp(-t/T_{inf})] \\ T_{inf} = (W_{sat} - W_{fc}) / K_{sat} \end{cases} \quad (A5)$$

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

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

$$19 \quad \begin{cases} Q_{overl} = (Q'_{overl} + Q_{stor,i-1}) \cdot [1 - \exp(-T_{retain}/T_{route})] \\ T_{route} = T_{overl} + T_{rch} = \frac{L_{overl}^{0.6} \cdot n_{overl}^{0.6}}{18 \cdot slp_{overl}^{0.3}} + \frac{0.62 \cdot L_{rch} \cdot n_{rch}^{0.75}}{A^{0.125} \cdot slp_{rch}^{0.375}} \end{cases} \quad (A6)$$

20 where Q_{overl} is the overland flow discharged into main channel (mm); Q'_{overl} is the
21 lateral flow amount generated in the sub-basin (mm), $Q_{stor,i-1}$ is the lateral flow in the
22 previous day (mm); T_{retain} is the retain time of flow (days); T_{route} , T_{overl} and T_{rch} are the

1 routing times of the total flow, overland flow and river flow, respectively (days); L_{overl}
 2 and L_{rch} are the lengths of sub-basin slope and river, respectively (km); slp_{overl} and
 3 slp_{rch} are the slopes of sub-basin and river, respectively (m/m); n_{overl} and n_{rch} are the
 4 Manning's roughness coefficients for sub-basin and river, respectively (m/m); A is the
 5 sub-basin area (km²).

6

7 **Appendix B: Soil biochemical module**

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

$$9 \quad T(Z,t) = \bar{T} + (AM/2 \cdot \cos[2\pi \cdot (t - 200)/365] + TG - T(0,t)) \cdot \exp(-Z/DD) \quad (B1)$$

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

$$14 \quad \begin{cases} DD = DP \cdot \exp\left\{\ln(500/DP) \cdot [(1 - \xi)/(1 + \xi)]^2\right\} \\ DP = 1000 + 2500BD/[BD + 686 \exp(-5.63BD)] \\ \xi = SW/[(0.356 - 0.144BD) \cdot Z_M] \\ TG_{IDA} = (1 - AB) \cdot (T_{mx} + T_{mn})/2 \cdot (1 - RA/800) + T_{mx} \cdot RA/800 + AB \cdot TG_{IDA-1} \end{cases} \quad (B2)$$

15 where DP is maximum damping depth of soil temperature (mm); BD is soil bulk
 16 density (t/m³); ξ is scale parameter; IDA is day of the year; AB is surface albedo; RA
 17 is daily solar radiation (ly).

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

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

$$21 \quad dC/dt = \mu_{CLAY} \cdot \mu_{C:N} \cdot \mu_{t,n} \cdot [S \cdot k_1 + (1 - S) \cdot k_2] \quad (B3)$$

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

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

$$2 \quad FIX_{NH_4} = [0.41 - 0.47 \cdot \log(NH_4)] \cdot (CLAY / CLAY_{max}) \quad (B4)$$

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

$$5 \quad \begin{cases} \log(K_{NH_4} / K_{H_2O}) = \log(NH_{4m} / NH_{3m}) + pH \\ NH_{3m} = 10^{\{\log(NH_4) - (\log(K_{NH_4}) - \log(K_{H_2O})) + pH\} \cdot (CLAY / CLAY_{max})} \\ AM = 2 \cdot (NH_3) \cdot (D \cdot t / 3.14)^{0.5} \end{cases} \quad (B5)$$

6 where K_{NH_4} and K_{H_2O} are dissociation constants for $NH_4^+ : NH_3$ equilibrium, $H^+ : OH^-$
7 equilibrium, respectively; NH_{4m} and NH_{3m} are NH_4^+ and NH_3 concentrations (mol/L)
8 in the liquid phase, respectively; AM and D are accumulated NH_3 loss (mol/cm²) and
9 diffusion coefficients (cm²/d²), respectively.

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

$$13 \quad \begin{cases} dNNO = NH_4 \cdot [1 - \exp(-K_{35} \cdot \mu_{t,n} \cdot dt)] \cdot \mu_{sw,n} \\ N_2O = (0.0014 \cdot NH_4 / 30.0) \cdot (0.54 + 0.51 \cdot T) / 15.8 \end{cases} \quad (B6)$$

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

16 *Denitrification:* The growth rate of denitrifier ($(dB/dt)_g$, kg/ha/day) is proportional to
17 their respective biomass, which is calculated with double Monod kinetics equation

$$18 \quad \begin{cases} (dB/dt)_g = \mu_{DN} \cdot B(t) \\ \mu_{DN} = \mu_{t,dn} \cdot (u_{NO_3} \cdot \mu_{PH,NO_3} + u_{NO_2} \cdot \mu_{PH,NO_2} + u_{N_2O} \cdot \mu_{PH,N_2O}) \\ u_{N_xO_y} = u_{N_xO_y,max} \cdot (C / K_{C,1/2} + C) \cdot (N_xO_y / K_{N_xO_y,1/2} + N_xO_y) \end{cases} \quad (B7)$$

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

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

2 The death rate of denitrifier ($(dB/dt)_d$, kg/ha/hr) is proportional to denitrifier biomass,
3 viz.

$$4 \quad (dB/dt)_d = M_C \cdot Y_C \cdot B(t) \quad (B9)$$

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

7 The consumption rates of soluble C and CO_2 production are calculated as

$$8 \quad \begin{cases} dC_{con}/dt = (\mu_{DN}/Y_C + M_C) \cdot B(t) \cdot \mu_{sw,d} \\ dCO_2/dt = dC_{con,t}/dt - (dB/dt)_d \end{cases} \quad (B10)$$

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

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

$$11 \quad dN_xO_y/dt = (u_{N_xO_y}/Y_{N_xO_y} + M_{N_xO_y} \cdot N_xO_y/N) \cdot B(t) \cdot \mu_{PHN_xO_y} \cdot \mu_{t,dn} \quad (B11)$$

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

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

$$16 \quad (dN/dt)_{ass} = (dB/dt)_g \cdot (1/CNR_{D:N}) \quad (B12)$$

17 The emission rates are the functions of adsorption coefficients of the gases in soils
18 and to the air filled porosity of the soil, given as.

$$19 \quad \begin{cases} P(N_2) = 0.017 + ((0.025 - 0.0013 \cdot AD) \cdot PA \\ P(N_2O) = [30.0 \cdot (0.0006 + 0.0013 \cdot AD) + (0.013 - 0.005 \cdot AD)] \cdot PA \\ P(NO) = 0.5 \cdot [(0.0006 + 0.0013 \cdot AD) + (0.013 - 0.005 \cdot AD) \cdot PA] \end{cases} \quad (B13)$$

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

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

$$6 \quad Leach_{NO_3} = W_{inf} \cdot \mu_{CLAY} \cdot \mu_{soc} \quad (B14)$$

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

9 **B.3 P cycle**

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

12

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

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

$$16 \quad \Delta V = V_{flowin} - V_{flowout} + V_{pcp} - V_{evap} - V_{seep} - V_{withd} \quad (C1)$$

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

22 According to the design data of dams and sluices in China, there is a particular
 23 relationship among water level, storage and outflow. The outflow is determined by
 24 the water level or water storage volume. Thus, the relationships are described by
 25 equations.

$$26 \quad \begin{cases} V_{flowout} = f'(V, H) \\ SA = f''(V, H) \end{cases} \quad (C2)$$

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

5

6 **Appendix D: Evaluation indices of model performance**

7 Bias:
$$bias = \frac{\sum_{i=1}^N (O_i - S_i)}{\sum_{i=1}^N O_i} \quad (D1)$$

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

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

10 Correlation coefficient:
$$r = \frac{\sum_{i=1}^N (O_i - \bar{O}) \cdot (S_i - \bar{S})}{\sqrt{\sum_{i=1}^N (O_i - \bar{O})^2 \cdot \sum_{i=1}^N (S_i - \bar{S})^2}} \quad (D4)$$

11 Coefficient of efficiency:
$$NS = 1 - \frac{\sum_{i=1}^N (O_i - S_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (D5)$$

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

15

16 **Acknowledgements**

17 This study was supported by the Natural Science Foundation of China (No.
 18 41271005), the China Youth Innovation Promotion Association CAS (No. 2014041),
 19 the Key Project for the Strategic Science Plan in IGSNRR, CAS (No. 2012ZD003),
 20 the Endeavour Research Fellowship, China Visiting Scholar Project from China
 21 Scholarship Council, and the CSIRO Computational and Simulation Sciences
 22 Research Platform. The authors would like to thank Dr. Yongqiang Zhang, Mr. James
 23 R Frankenberger for their participation in our internal review procedure, Dr. Markus
 24 Hrachowitz and Dr. Christian Stamm for improving the quality and presentation of

1 the manuscript, and the anonymous reviewers for their valuable comments and
2 suggestions.

3

4 **References**

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

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

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

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

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

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

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

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

25 Biswas, A.K. Integrated water resources management: A reassessment-A water forum
26 contribution. *Water Inter.* 29:248 - 256, 2004.

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

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

- 1 Bouraoui, F., and Dillaha, T. A.: ANSWERS-2000: Runoff and sediment transport
2 model. *J. Environ. Eng.*, 122(6): 493-502, 1996.
- 3 Brown, L. C. and Barnwell, T. O.: The enhanced stream water quality models
4 QUAL2E and QUAL2E-UNCAS: documentation and user manual. Env. Res.
5 Laboratory. US EPA, 1987.
- 6 Burt, T. P. and Pinay, G.: Linking hydrology and biogeochemistry in complex
7 landscapes. *Prog. Phys. Geog.*, 29(3): 297-316, 2005.
- 8 Chen, Y., Song, X., Zhang, Z., Shi, P., and Tao, F.: Simulating the impact of flooding
9 events on non-point source pollution and the effects of filter strips in an
10 intensive agricultural watershed in China, *Limnology*, 16, 91–101, 2015. doi:
11 10.1007/s10201-014-0443-2.
- 12 China's national standard (CNS): *Current land use condition classification*
13 (GB/T21010-2007), General administration of quality supervision, inspection and
14 quarantine of China and Standardization administration of China, Beijing, China,
15 2007.
- 16 China State Environmental Protection Administration (CSEPA), (2002).
17 *Environmental quality standards for surface water -GB 3838-2002*. Beijing:
18 China Environmental Science Press. (In Chinese).
- 19 Deng, J., Zhu, B., Zhou, Z. X., Zheng, X. H., Li, C. S., Wang, T., and Tang, J. L.:
20 Modeling nitrogen loadings from agricultural soils in southwest China with
21 modified DNDC. *J. Geophys. Res.: Biogeosci.* (2005–2012), 116(G2), 2011.
- 22 Di Toro, D. M., Fitzpatrick, J. J., and Thomann, R. V.: Water quality analysis simulation
23 program (WASP) and model verification program (MVP)-Documentation.
24 Hydroscience, Inc., Westwood, NY, for U.S. EPA, Duluth, MN, Contract No.
25 68-01-3872, 1983.
- 26 Duan, Q., Sorooshian, S., and Gupta, V. K.: Optimal use of the SCE-UA global
27 optimization method for calibrating watershed models. *J. Hydrol.*, 158(3): 265-284,
28 1994.
- 29 Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration
30 approaches in hydrological modelling: a review. *Hydrol. Sci. J.*, 55:58-78, 2010.

- 1 Gassman, P.W., Reyes, M.R., Green,C.H., and Arnold, A.G.: The soil and water
2 assessment tool: historical development, applications, and future research
3 directions. *T. ASABE*, 50: 1211-1250, 2007.
- 4 Gleick, P.H. Water in crisis: paths to sustainable water use. *Ecol. Appl.*, 8(3): 571-579,
5 1998.
- 6 Goldberg, D. E.: Genetic algorithms in search, optimization, and machine learning,
7 Reading Menlo Park: Addison-Wesley, Massachusetts, USA, 1989.
- 8 Hamrick, J. M.: A three-dimensional environmental fluid dynamics computer code:
9 theoretical and computational aspects, Special Report, The College of William
10 and Mary, Virginia Institute of Marine Science, Virginia, USA, 317, 1992.
- 11 Hargreaves, G. H., and Samani, Z. A.: Estimating potential evapotranspiration. *J.*
12 *Irrigat. Drain. Div.*, 108(3), 225-230, 1982.
- 13 Henan Statistical Yearbook in 2003, 2004 and 2005. China Statistics Press, Beijing.
- 14 Her, Y., and Chaubey, I. Impact of the numbers of observations and calibration
15 parameters on equifinality, model performance, and output and parameter
16 uncertainty. *Hydrol. Process.*, 29:4220-4237, 2015.
- 17 Horst, W.J., Kamh, M., Jibrin, J.M. and Chude, V.O.: Agronomic measures for
18 increasing P availability to crops, *Plant. Soil.* 237: 211-223, 2001.
- 19 Johnes, P.J.: Evaluation and management of the impact of land use change on the
20 nitrogen and phosphorus load delivered to surface waters: the export coefficient
21 modelling approach. *J. Hydrol.*, 183(3): 323-349, 1996.
- 22 Johnsson, H., Bergstrom, L., Jansson, P. E., and Paustian, K.: Simulated nitrogen
23 dynamics and losses in a layered agricultural soil. *Agr. Ecosyst. Environ.*, 18(4),
24 333-356,1987
- 25 Kirchner J.W.: Getting the right answers for the right reasons: Linking measurements,
26 analyses, and models to advance the science of hydrology. *Water Resour. Res.*,
27 42(3), 2006. W03S04 doi: 10.1029/2005WR004362.
- 28 Kindler, J.: Integrated water resources management: the meanders. *Water Int.*,
29 25:312-319, 2000.

- 1 King, K. W., Arnold, J. G., Bingner, R. L. Comparison of Green-Ampt and curve
2 number methods on Goodwin Creek watershed using SWAT. *Trans. ASAE*, 42(4),
3 919-925, 1999.
- 4 Kennedy, J.: Particle swarm optimization, Encyclopedia of Machine Learning.
5 Springer USA, 760–766, 2010.
- 6 Krysanova, V., Mueller-Wohlfeil, D.I., Becker, A.: Development and test of a spatially
7 distributed hydrological / water quality model for mesoscale watersheds. *Ecol.*
8 *Model.*, 106, 261-289,1998.
- 9 Li, C., Frolking, S., Frolking, T.A.: A model of nitrous oxide evolution from soil
10 driven by rainfall events: 1. Model structure and sensitivity. *J. Geophys. Res.*
11 (1984–2012), 97(D9): 9759-9776, 1992.
- 12 Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges: A Simple hydrologically
13 based model of land surface water and energy fluxes for GSMs, *J. Geophys. Res.*,
14 99(D7), 14,415-14,428, 1994.
- 15 Lindström, G., Pers, C.P., Rosberg, R., Strömquist, J., Arheimer, B.: Development and
16 test of the HYPE (Hydrological Predictions for the Environment) model - A water
17 quality model for different spatial scales. *Hydrol. Res.* 41.3-4:295-319,2010.
- 18 Ma, F., Ye, A., Gong, W., Mao, Y., Miao, C., and Di, Z. An estimate of human and
19 natural contributions to flood changes of the Huai River. *Global Planetary*
20 *Change*, 119, 39-50, 2014.
- 21 Mantovan, P., and Todini, E.: Hydrological forecasting uncertainty assessment:
22 Incoherence of the GLUE methodology, *J. Hydrol.*, 330, 368–381, 2006.
- 23 Mantovan, P., Todini, E. and Martina, M. L.V.: Reply to comment by Keith Beven,
24 Paul Smith, and Jim Freer on “Hydrological forecasting uncertainty assessment:
25 Incoherence of the GLUE methodology”, *J. Hydrol.*, 338, 319–324, 2007.
- 26 McDonnell, J. J., Sivapalan, M., Vache, K., Dunn, S., Grant, G., Haggerty, R., Hinz,
27 C., Hooper, R., Kirchner, J., Roderick, M.L., Selker, J., and Weiler, M.: Moving
28 beyond heterogeneity and process complexity: A new vision for watershed
29 hydrology, *Water Resour. Res.*, 43, W07301, doi:10.1029/2006WR005467,2007.

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

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

6 Onstad, C. A. and Foster, G. R.: Erosion modeling on a watershed. *T.ASAE*
7 18(2):288-292, 1975.

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

13 Pushpalatha, R., Perrin, C., Le Moine, N., and Andréassian, V.: A review of efficiency
14 criteria suitable for evaluating low-flow simulations. *J.Hydrol.*, 420-421, 171-182,
15 2012.

16 Rallison, R.E. and Miller, N.: Past, present and future SCS runoff procedure. 353-364,
17 1981. In V.P. Singh (ed.). Rainfall runoff relationship. Water Resources
18 Publication, Littleton, CO.

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

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

24 Pohlert, T., L. Breuer, J.A. Huisman, and H.-G. Frede.: Integration of a detailed
25 biogeochemical model into SWAT for improved nitrogen predictions-model
26 development, sensitivity and uncertainty analysis. *Ecol. Model.* 203:215-228,
27 2006.

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

- 1 Sharpley, A.N. and Williams, J.R.: EPIC-erosion/productivity impact calculator: 1.
2 Model documentation. Technical Bulletin-United States Department of
3 Agriculture, Agric. Res. Service, Washington DC, USA, 1990.
- 4 Shi, P., Chen, C., Srinivasan, R., Zhang, X., Cai, T., Fang, X., Qu, S., Chen, X., and
5 Li, Q.. Evaluating the SWAT model for hydrological modeling in the Xixian
6 watershed and a comparison with the XAJ model. *Water Resour. Manag.*, 25(10),
7 2595-2612, 2011.
- 8 Singh, V.P. and Woolhiser, D.A.: Mathematical modeling of watershed hydrology. *J.*
9 *Hydrol. Eng.*, 7(4): 270-292, 2002.
- 10 Sivapalan, M. and Kalma, J. D.: Scale problems in hydrology: contributions of the
11 Robertson Workshop. *Hydrol. Process.*, 9(3-4), 243-250, 1995.
- 12 Tattari, S., Bärlund, I., Rekolainen, S., Posch, M., Siimes, K., Tuhkanen, H. R., and
13 Yli-Halla, M.: Modeling sediment yield and phosphorus transport in Finnish
14 clayey soils. *Transactions of the ASAE*, 44(2), 297-307, 2001.
- 15 van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasan,
16 R.: A global sensitivity analysis tool for the parameters of multi-variable
17 catchment models. *J. Hydrol.*, 324(1), 10-23, 2006.
- 18 Vinogradov, Y. B., Semenova, O. M., and Vinogradova, T. A.: An approach to the
19 scaling problem in hydrological modelling: the deterministic modelling
20 hydrological system. *Hydrol. Process.*, 25(7), 1055-1073, 2011.
- 21 Wang, G. S., Xia J., Tan G., and Lu A.F.: A research on distributed time variant gain
22 model: A case study on Chao River basin (in Chinese), *Prog. Geogr.*, 21(6), 573–
23 582, 2002
- 24 Wang, G., Xia, J., and Chen, J.: Quantification of effects of climate variations and
25 human activities on runoff by a monthly water balance model: A case study of the
26 Chaobai River basin in northern China. *Water Resour. Res.*, 45, W00A11,
27 doi:10.1029/2007WR006768, 2009.
- 28 Wang, J.Q., Ma, W.Q., Jiang, R.F. and Zhang, F.S.: Analysis about amount and ratio
29 of basal fertilizer and topdressing fertilizer on rice, wheat, maize in China. *Chin.*
30 *J. Soil Sci.*, 39(2):329-333, 2008. (In Chinese)

- 1 Wang, X.: Summary of Huaihe River Basin and Shandong Peninsula Integrated Water
2 Resources Plan, *China water resources*, 23,112-114,2011.
- 3 Wigmosta, M.S., Vail, L.W., and Lettenmaier,D.P.: A distributed hydrology-vegetation
4 model for complex terrain. *Water Resour. Res.*, 30(6): 1665-1679, 1994.
- 5 Williams, J.R., Jones, C.A., and Dyke, P.T. Modeling approach to determining the
6 relationship between erosion and soil productivity. *Trans. ASAE*, 27(1): 129-144,
7 1984.
- 8 Williams, J. R., Jones, C. A., Kiniry, J. R., and Spanel, D. A. The EPIC crop growth
9 model. . *Trans. ASAE*, 32(2):497-511, 1989.
- 10 Xia, J., Wang, G.S., Tan, G, Ye, A.Z., and Huang, G. H.: Development of distributed
11 time-variant gain model for nonlinear hydrological systems. *Sci. China: Earth Sci.*,
12 48(6), 713-723, 2005.
- 13 Xia, J.: Identification of a constrained nonlinear hydrological system described by
14 Volterra Functional Series, *Water Resour. Res.*, 27(9): 2415–2420, 1991.
- 15 Xing, G. X., and Zhu, Z. L.: An assessment of N loss from agricultural fields to the
16 environment in China. *Nutr. Cycl. Agroecosys.*, 57(1): 67-73, 2000.
- 17 Zhai, X.Y., Zhang, Y.Y., Wang X.L., Xia J. and Liang, T.: Non-point source pollution
18 modeling using Soil and Water Assessment Tool and its parameter sensitivity
19 analysis in Xin'anjiang Catchment, China. *Hydrol. Process.* 28, 1627-1640, 2014.
- 20 Zhang, Y.Y., Xia, J., Shao, Q.X., and Zhai, X.Y.: Water quantity and quality
21 simulation by improved SWAT in highly regulated Huai River Basin of China.
22 *Stoch. Env Res. Risk A.*, 27(1), 11-27, 2013.
- 23 Zhu, Z. L.: Loss of fertilizer N from plants-soil system and the strategies and
24 techniques for its reduction. *Soil Environ. Sci.*, 9(1):1-6, 2000. (in Chinese)

1 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, texture, 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	Runoff or other hydrological component observations, etc.	Hydrological parameter calibration	Hydrology
Water quality	Urban wastewater discharge outlets and the 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 indicators	Populations, breeding stock of large animals and livestock, water withdrawal in each sub-basin	Hydrology and water quality
Water projects	Reservoir design data attribute parameters	Regulation rules of reservoirs 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

1 Table 2 Sensitive parameters, their value ranges and relative importance for runoff
 2 and NH₄-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	-
$R_d(\text{COD})$	0.02 to 3.4	COD deoxygenation rate at 20 °C	-	6.62
$R_{set}(\text{COD})$	-0.36 to 0.36	COD settling rate at 20 °C	-	3.60
$R_d(\text{NH}_4)$	0.1 to 1	Bio-oxidation rate of NH ₄ -N at 20 °C	-	1.97
$K_{set}(\text{NH}_4)$	0 to 100	Settling rate of NH ₄ -N in the reservoirs	-	14.17
$K_d(\text{COD})$	0.02 to 3.4	COD deoxygenation rate in the reservoirs at 20°C	-	2.12
$K_d(\text{NH}_4)$	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

3

4

1 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

2

3

1 Table 4. The runoff simulation results at regulated stations with and without the dam
 2 regulation considered. Range means the difference of objective function value
 3 between regulations considered and not considered. If the range value is less than 0.0,
 4 then the simulation with regulation is better than that without regulation. Otherwise,
 5 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

6

1 Table 5. The comparison of NH₄-N simulation results between with and without dam
 2 regulation considered.

Stations	Periods	Regulated			Unregulated			Range	Ratio of nonpoint 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		
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		
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		
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		
Less-regulated stations									
Shenqiu	Calibration	0.13	0.62	0.26	-	-	-	-	47.13
	Validation	0.16	0.41	0.37	-	-	-		

3

4

1 **List of Figure Captions**

2

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

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

7 **Figure 3.** The flowchart of SBM (a) and CGM (b) in [the](#) ecological part and the
8 interactions with other modules.

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

11 **Figure 5.** The flowchart of PAT and its interactions with [other modules](#).

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

14 **Figure 7.** The daily runoff simulation at all stations.

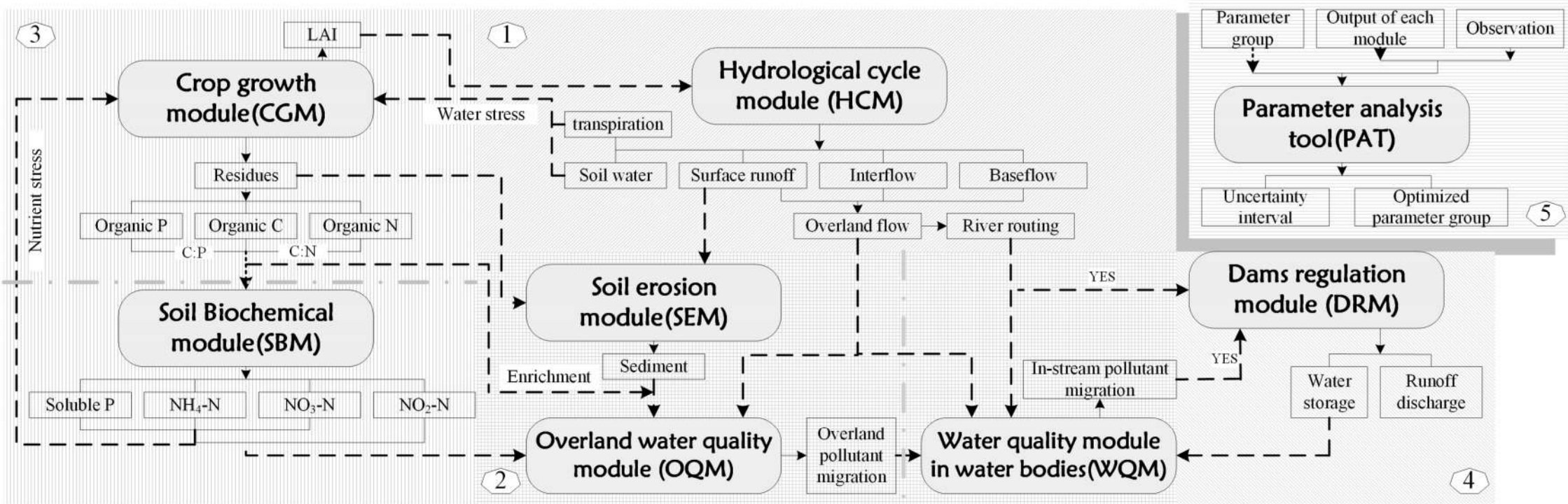
15 **Figure 8.** The cumulative distributions of simulated and observed daily runoff at all
16 stations

17 **Figure 9.** The simulated NH₄-N concentration variation at all [stations](#).

18 **Figure 10.** The spatial pattern of nonpoint source NH₄-N load (a) and its relationship
19 with paddy area (b) and rice yield (c) at the sub-basin and regional scale in the
20 Shaying River Catchment.

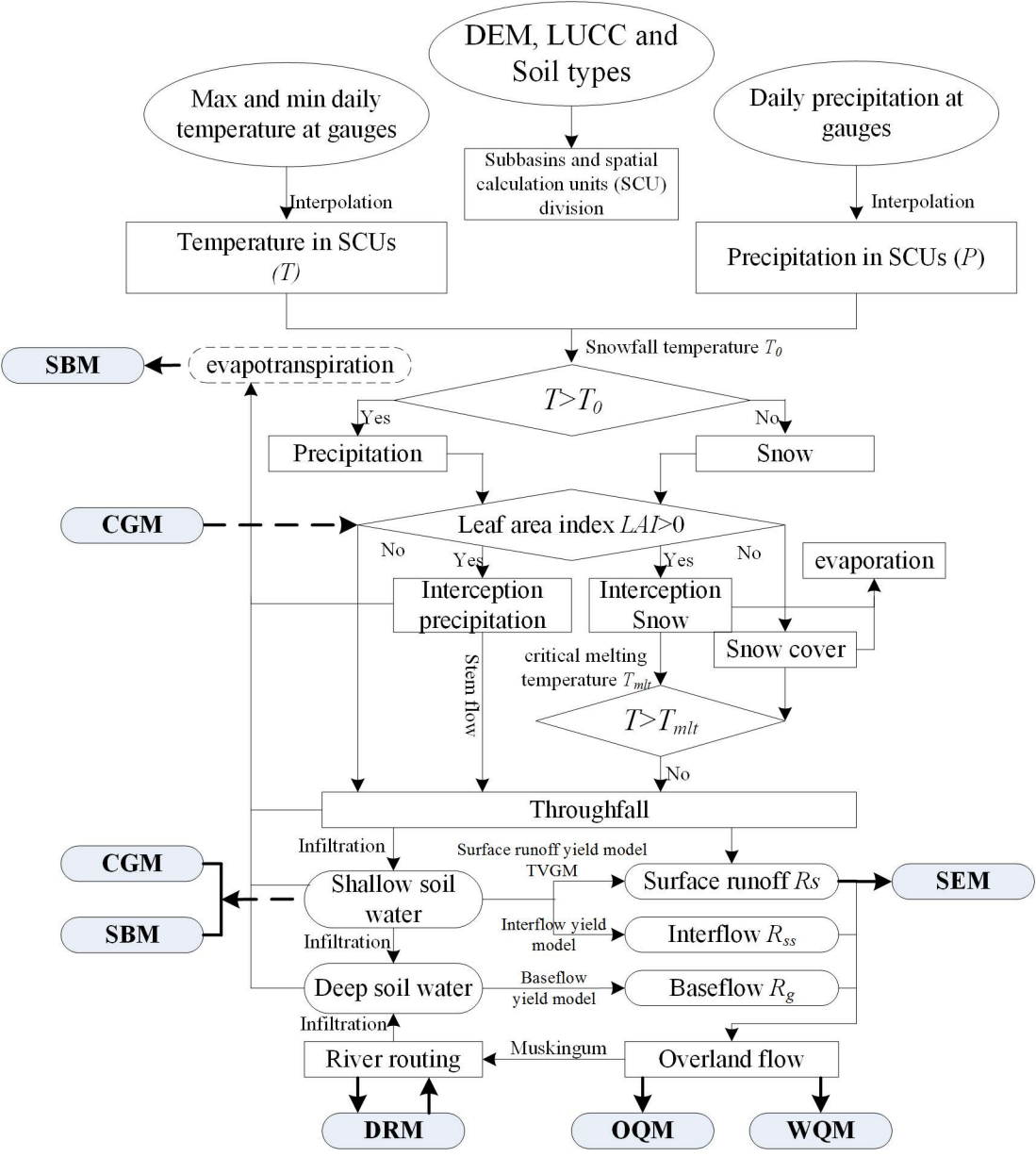
21 **Figure 11.** The spatial pattern of corn yield at the sub-basin and regional scale in the
22 Shaying River Catchment.

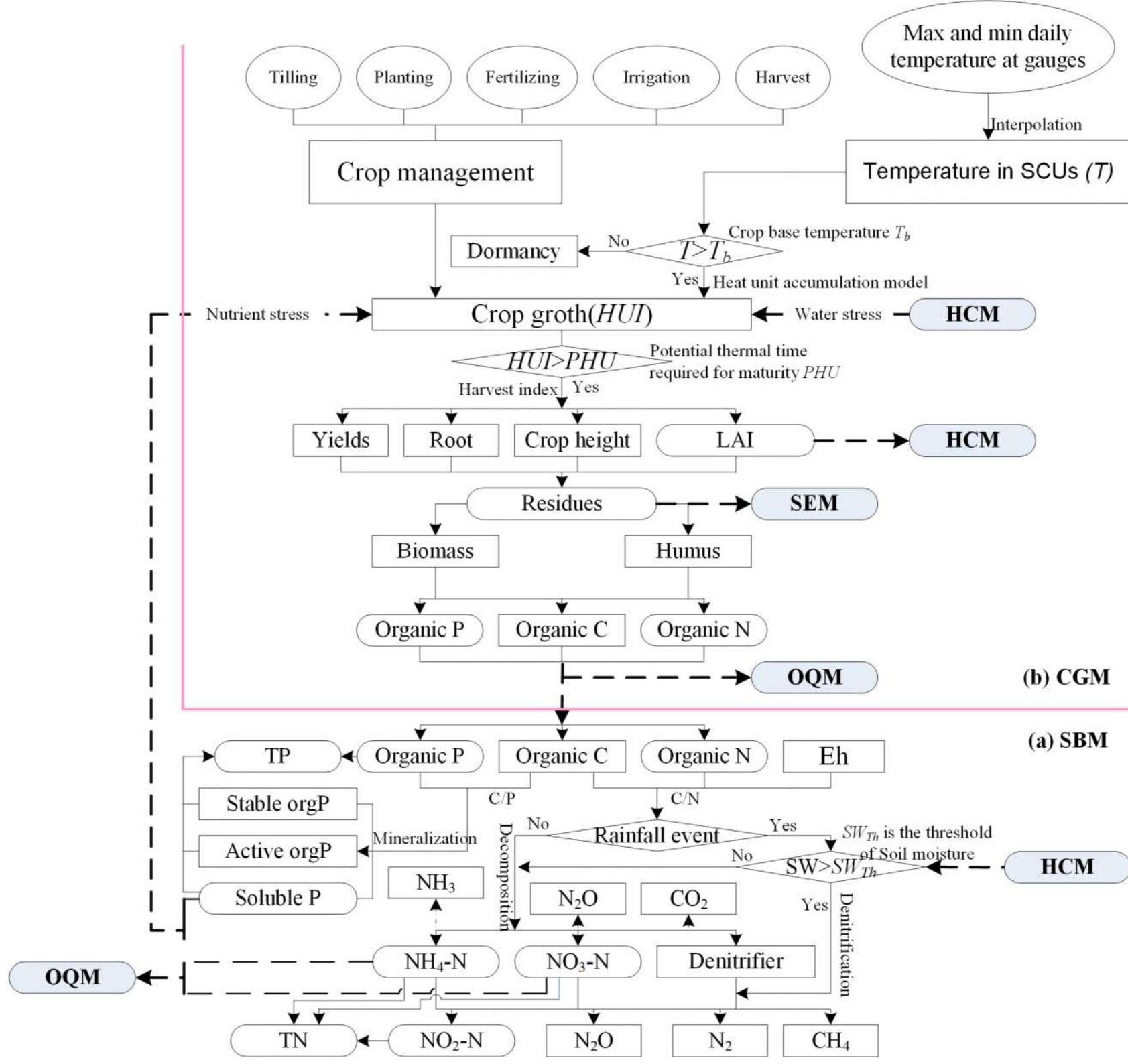
23

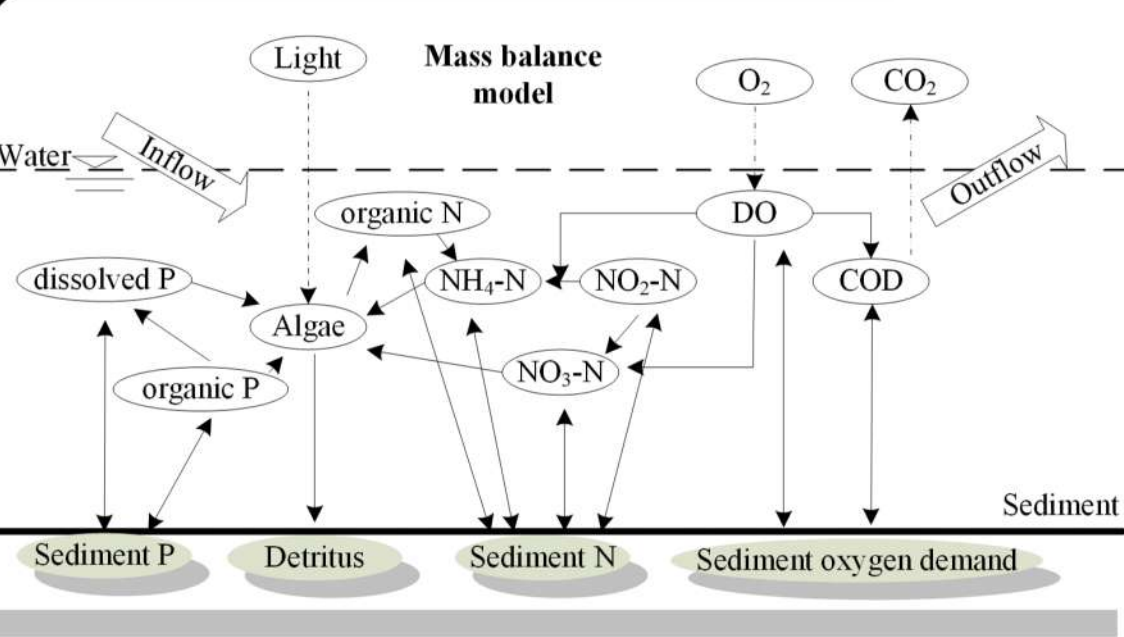
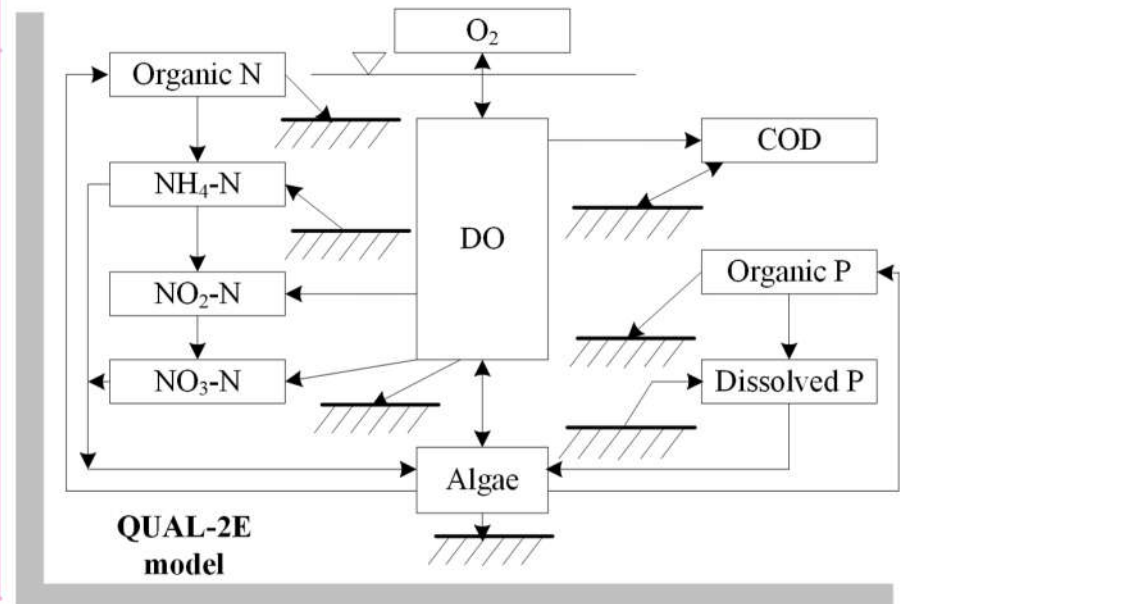
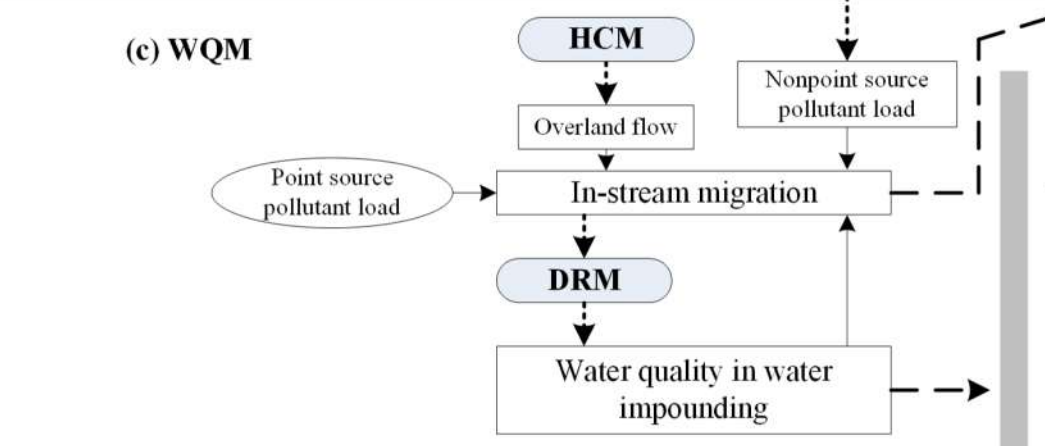
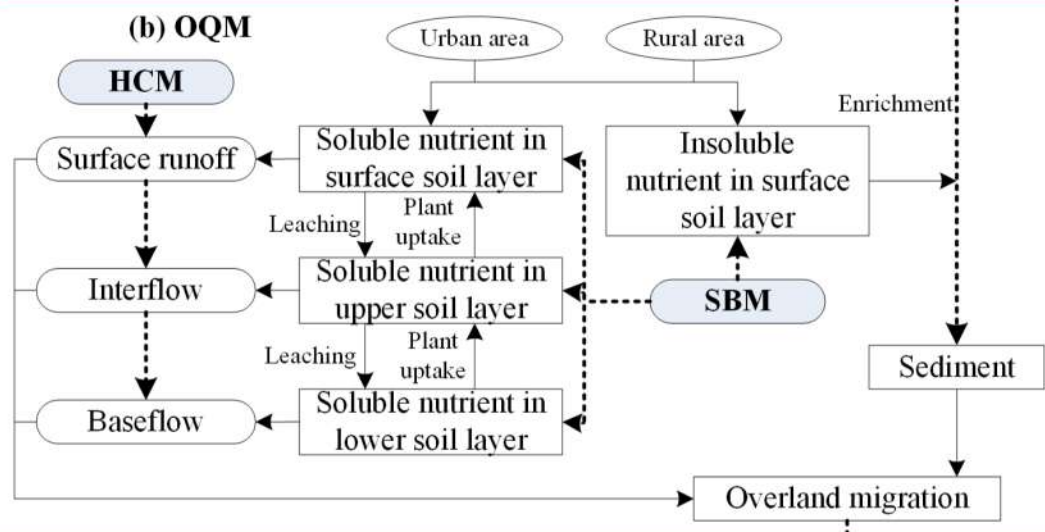
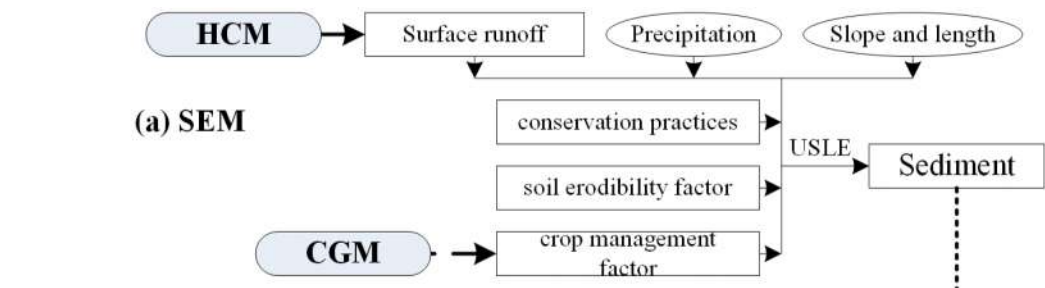


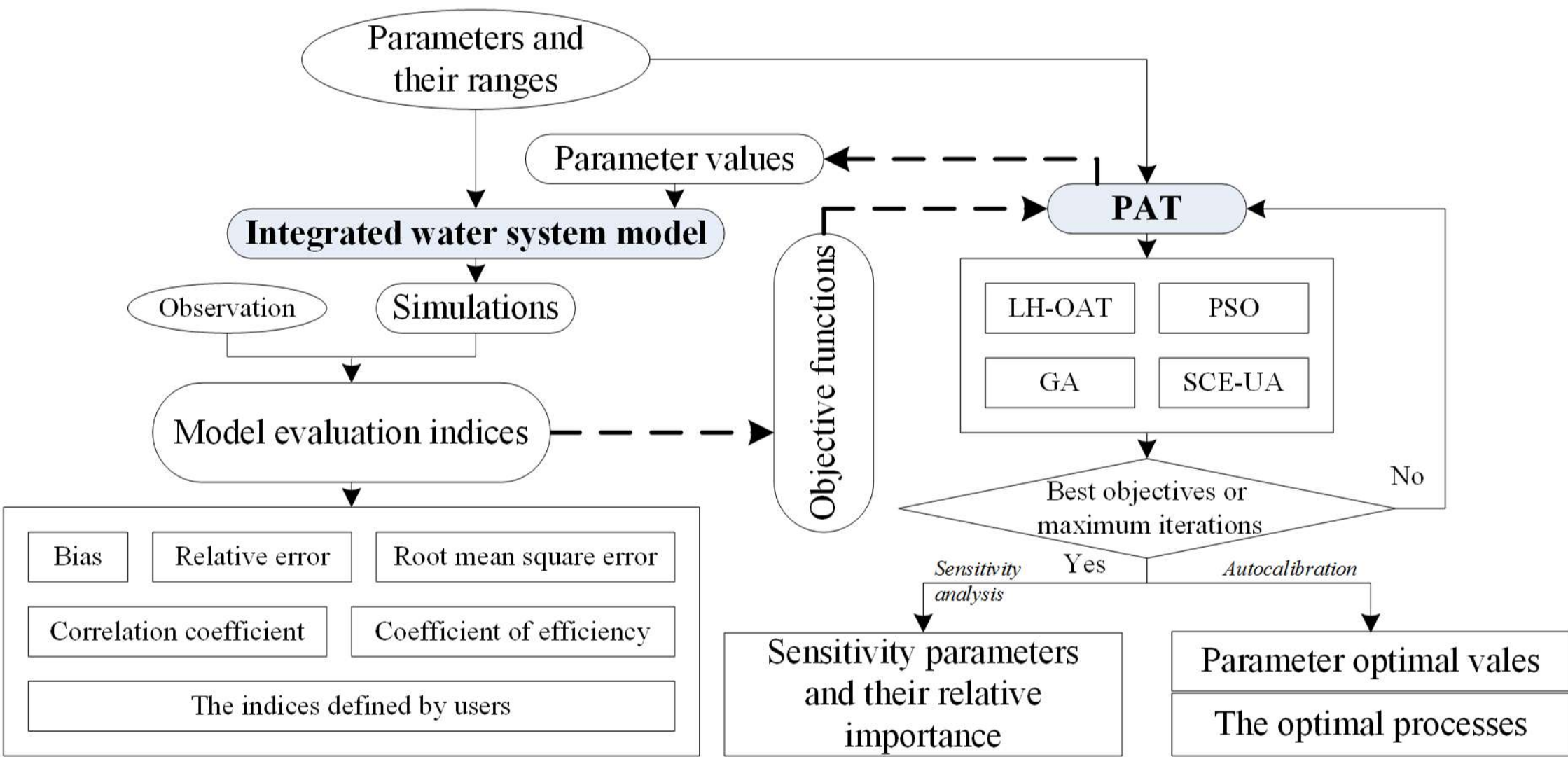
Legend

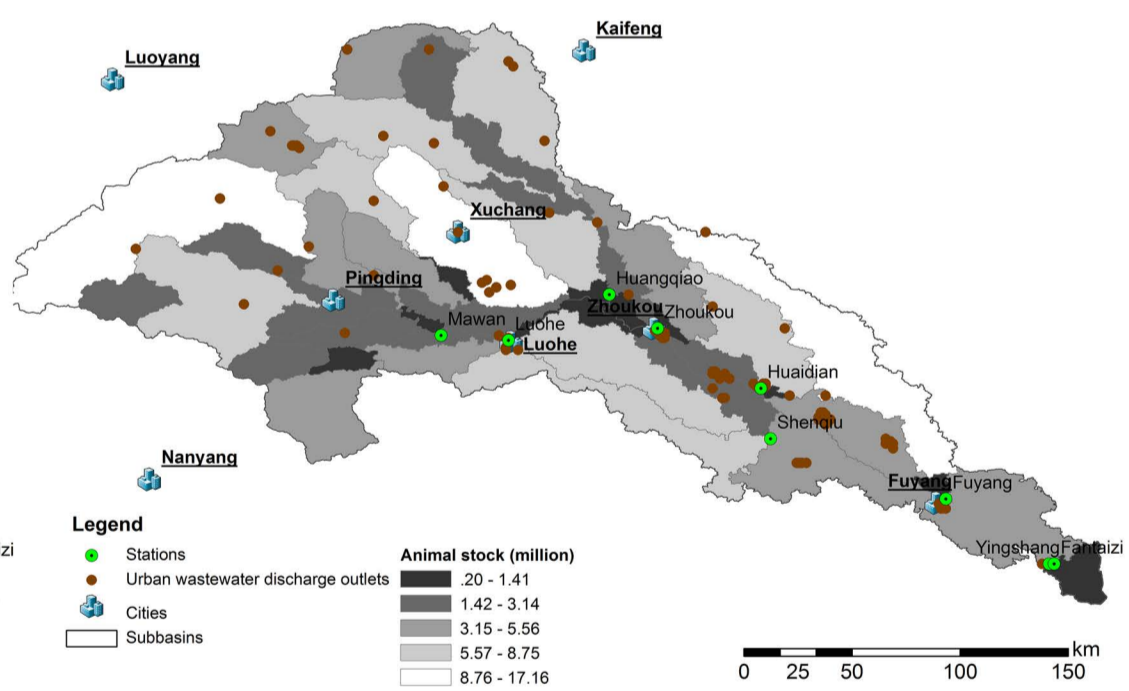
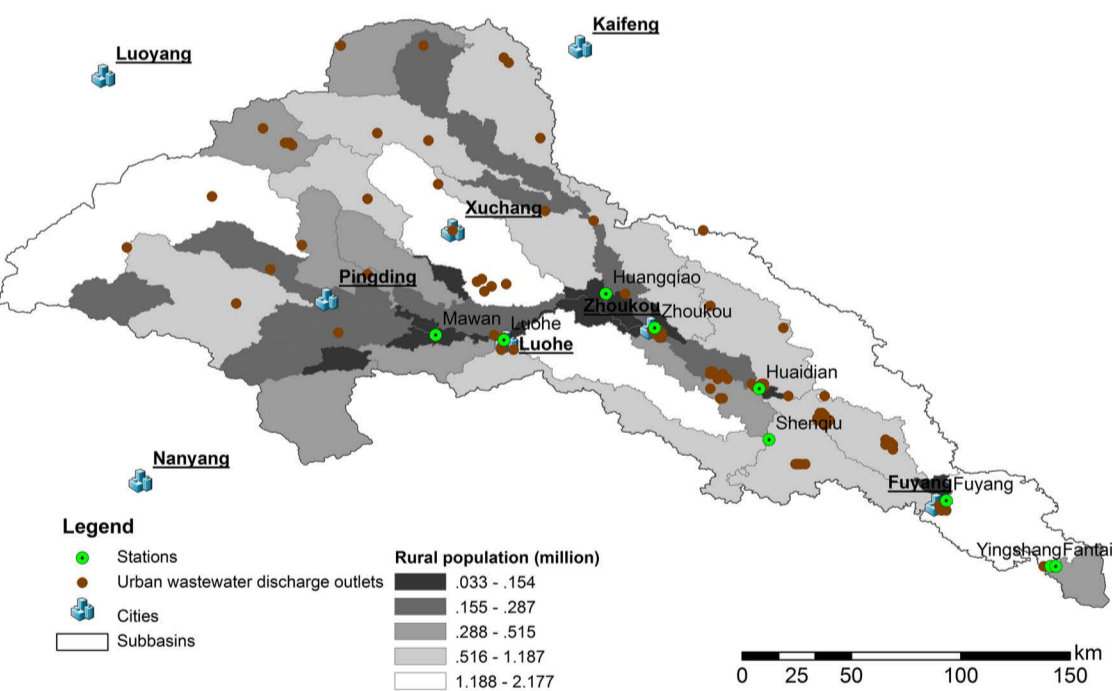
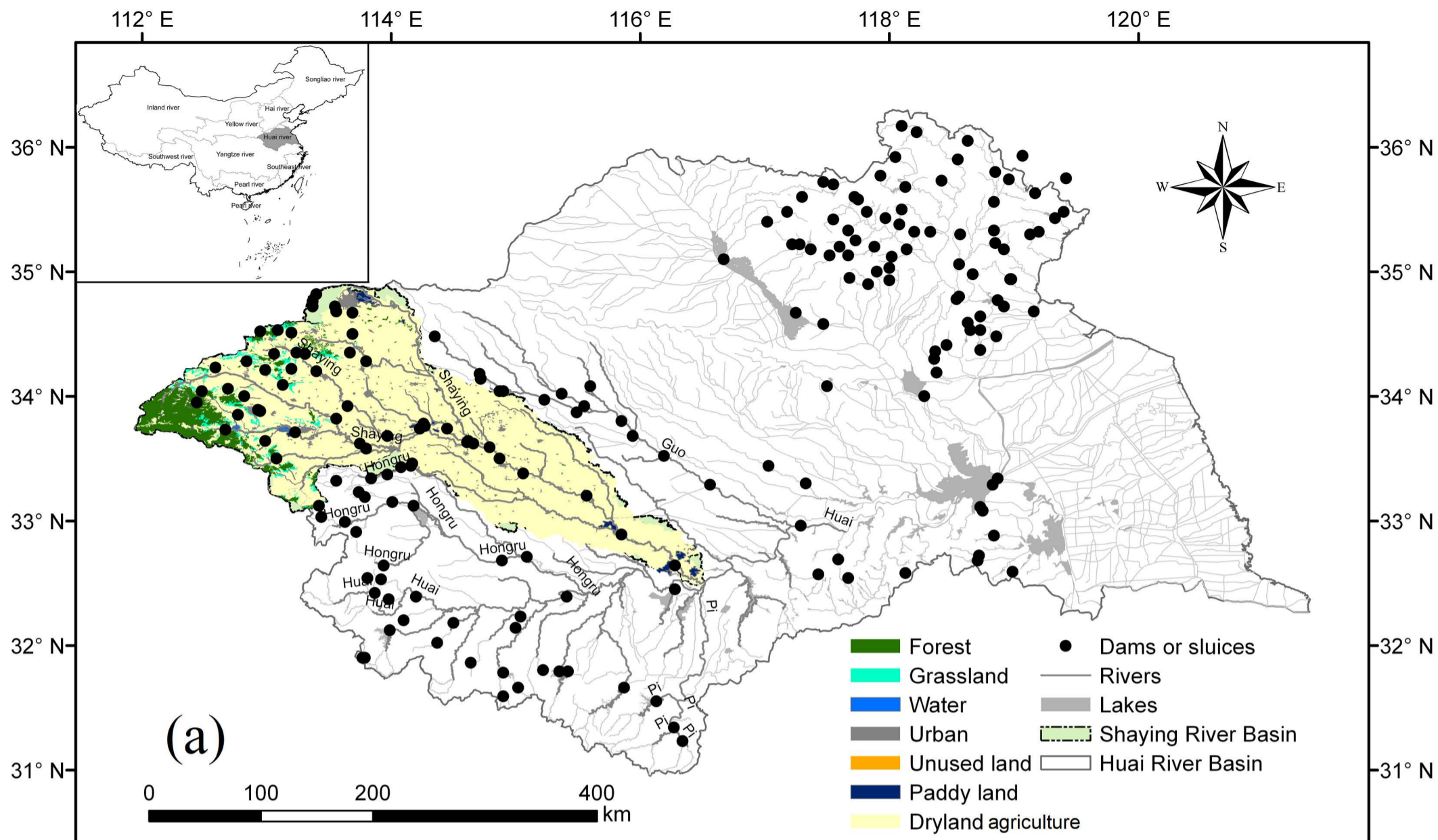
- Interactions between modules
- Module Name
- The major elements or processes which connect the different modules
- The major parts





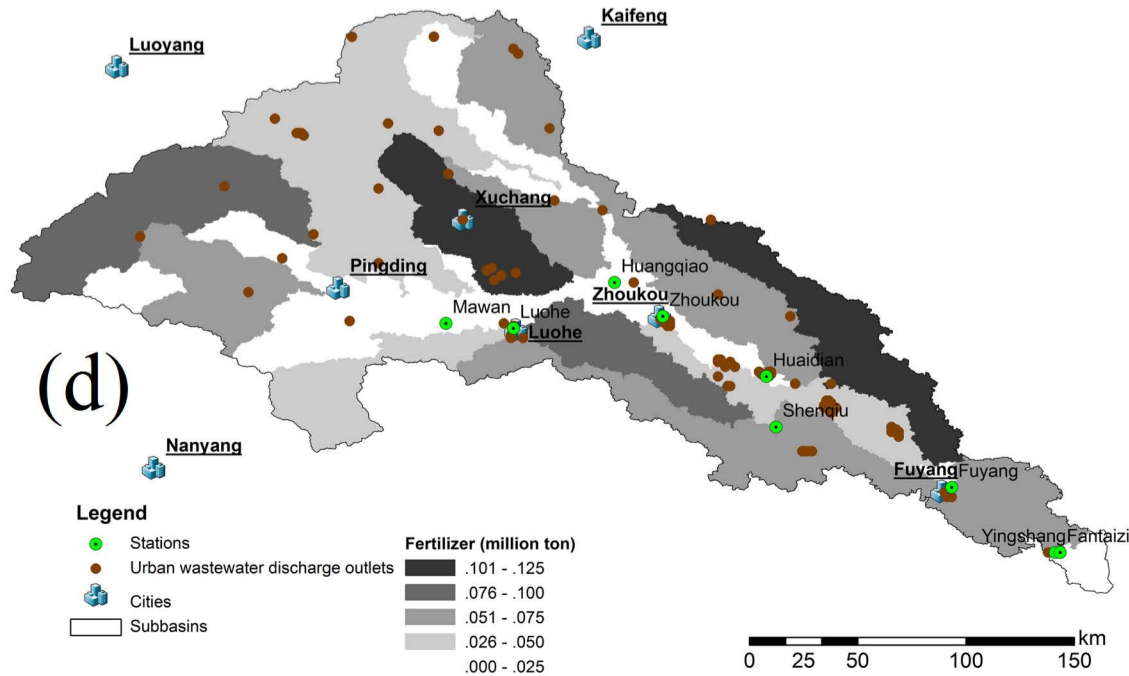




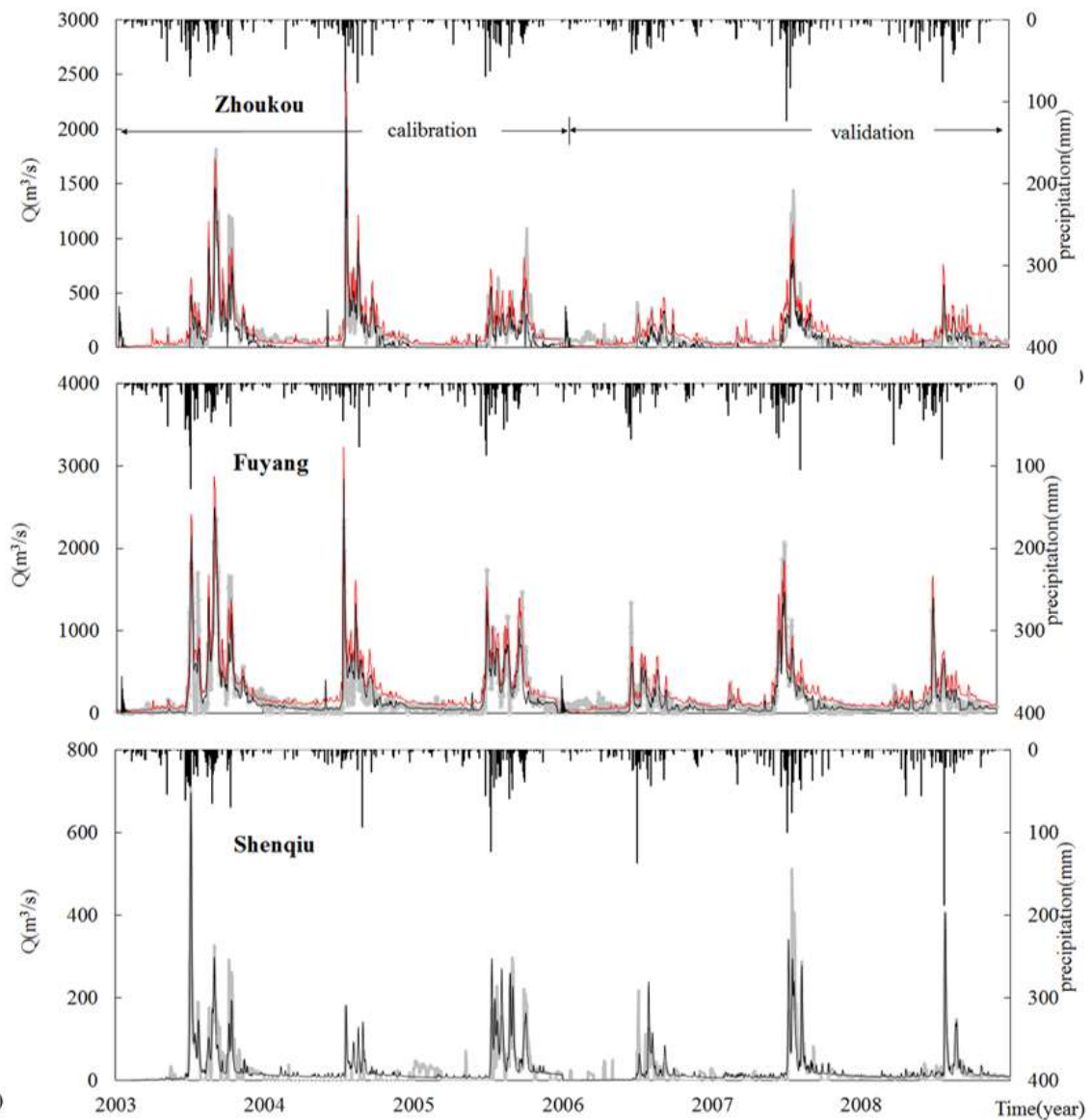
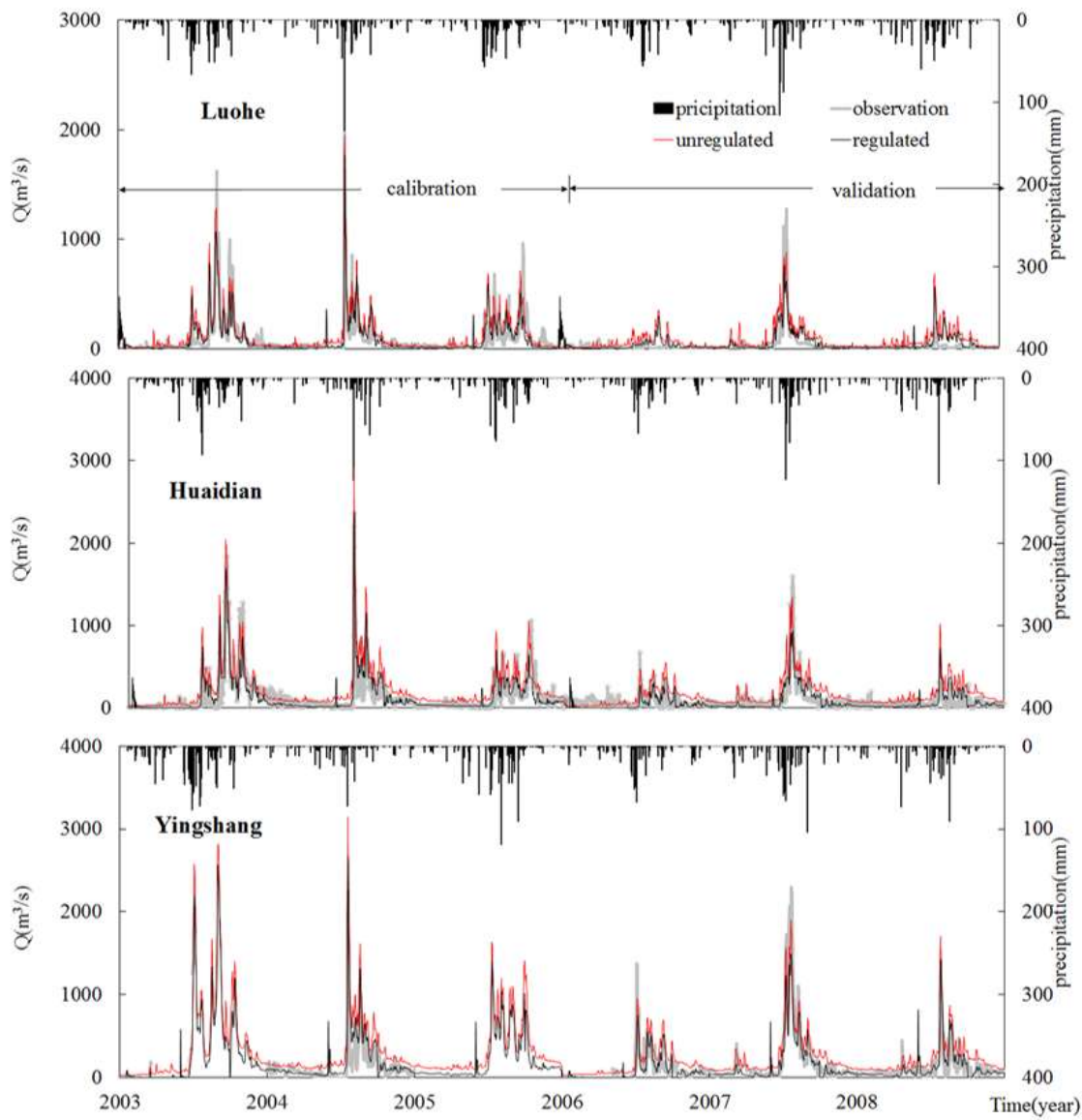


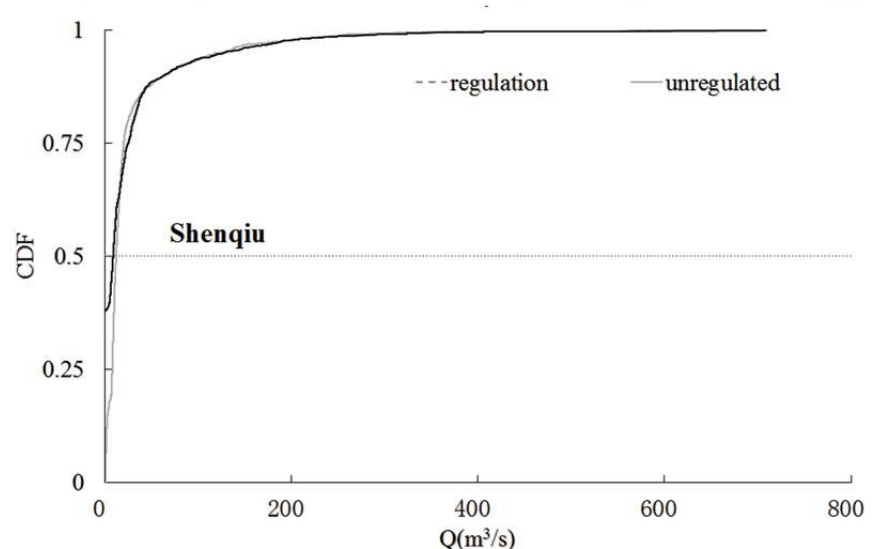
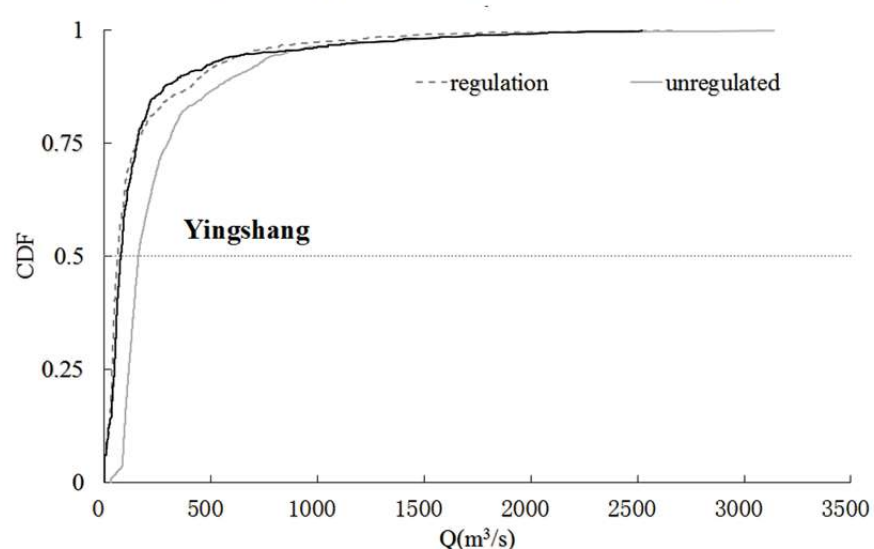
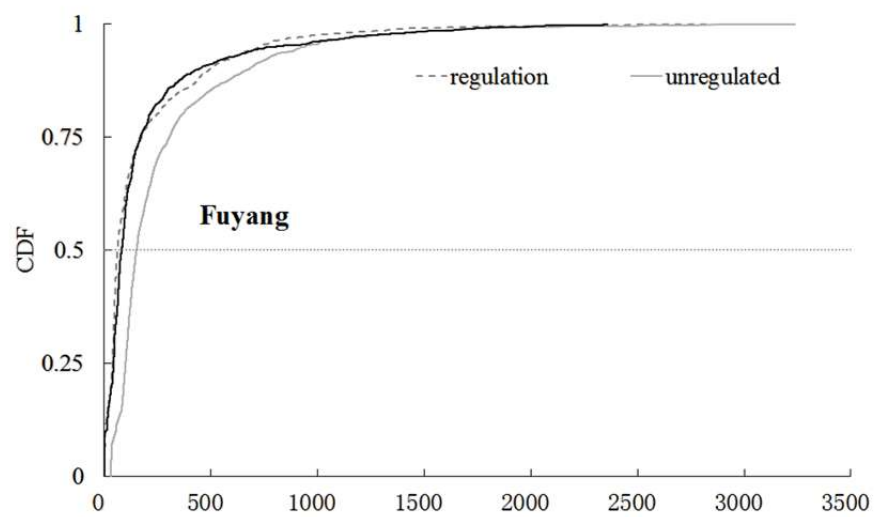
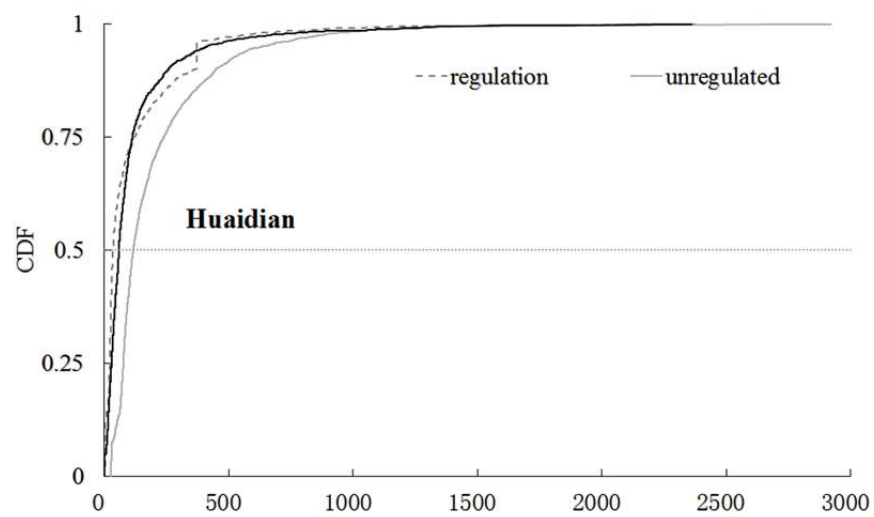
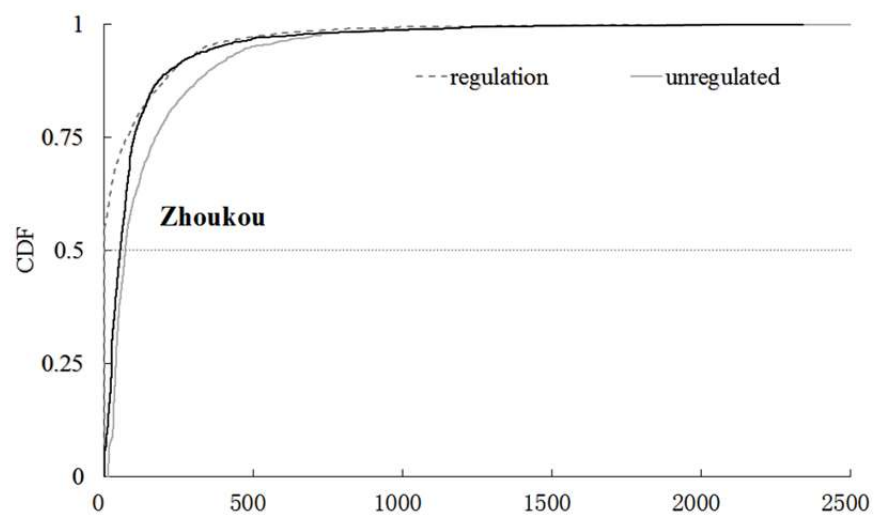
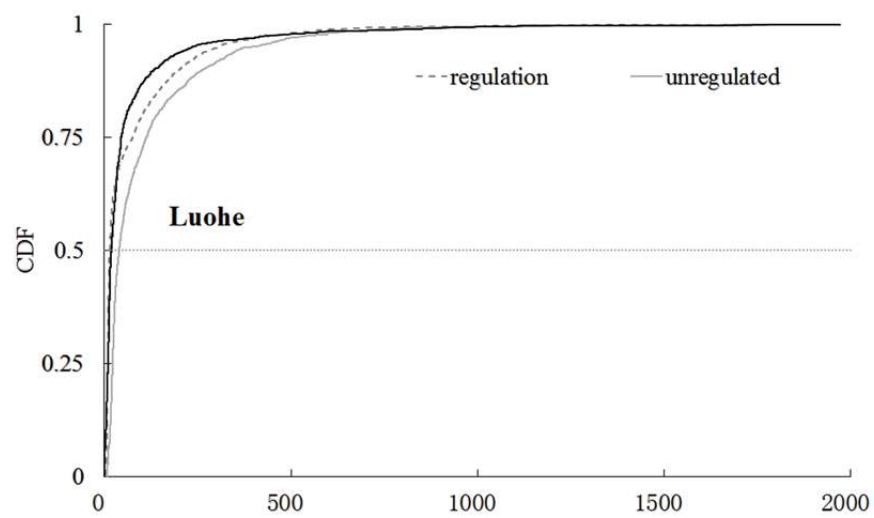
(b)

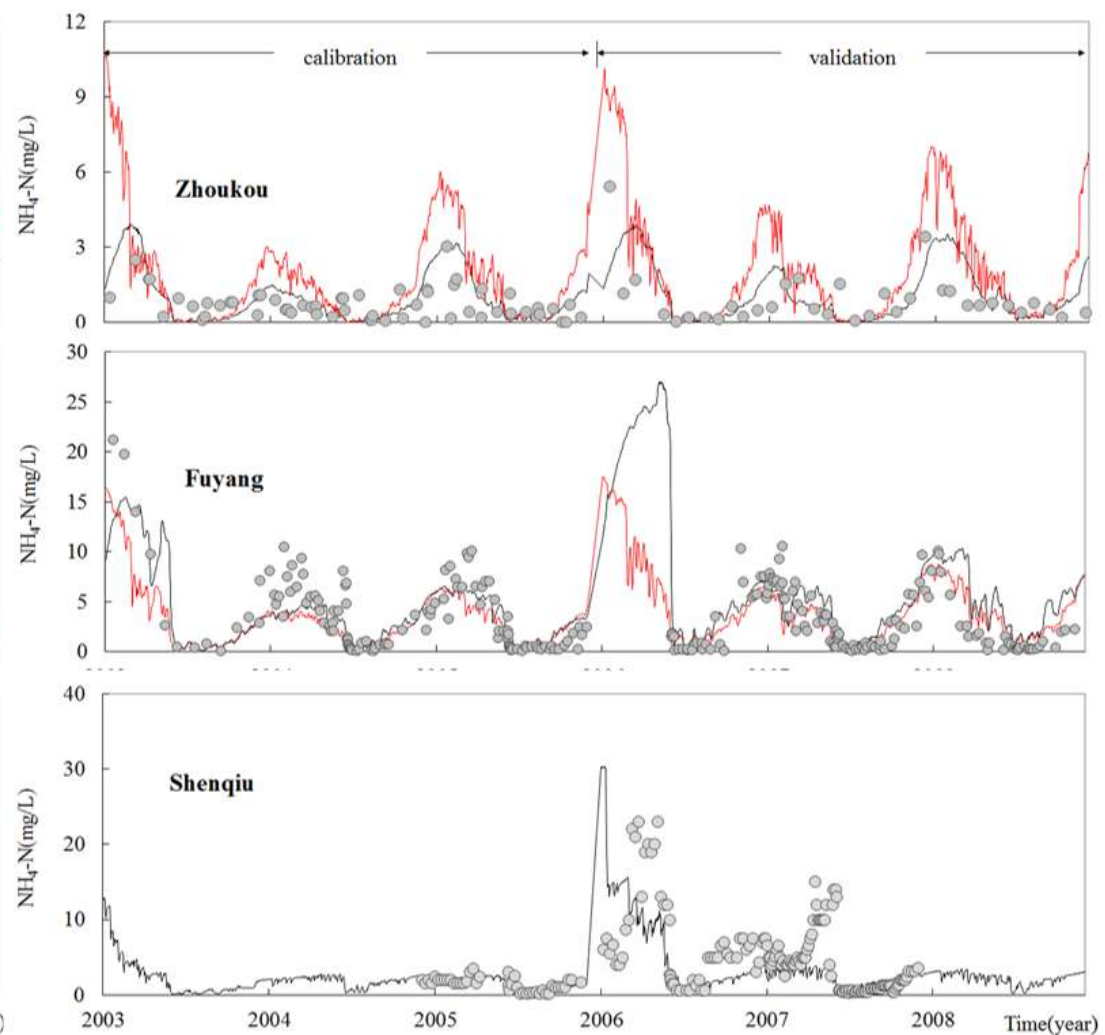
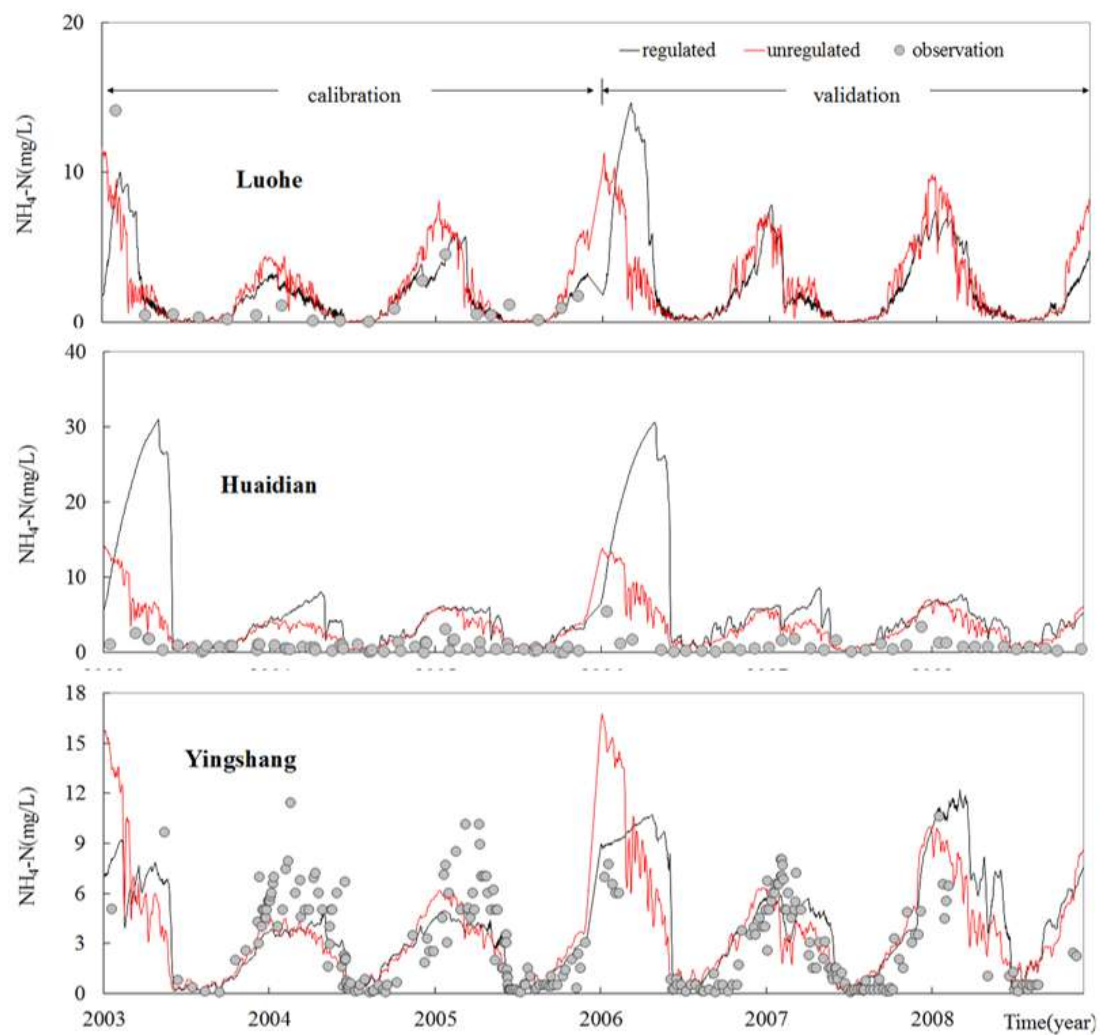
(c)

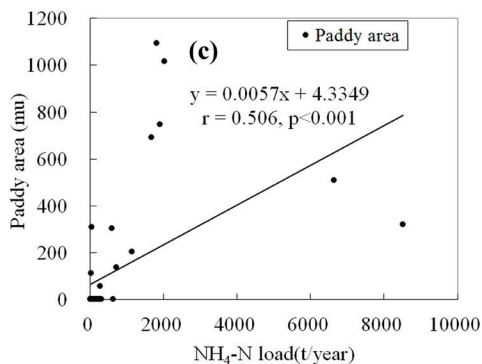
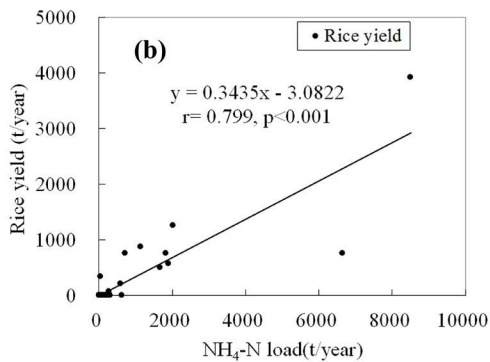
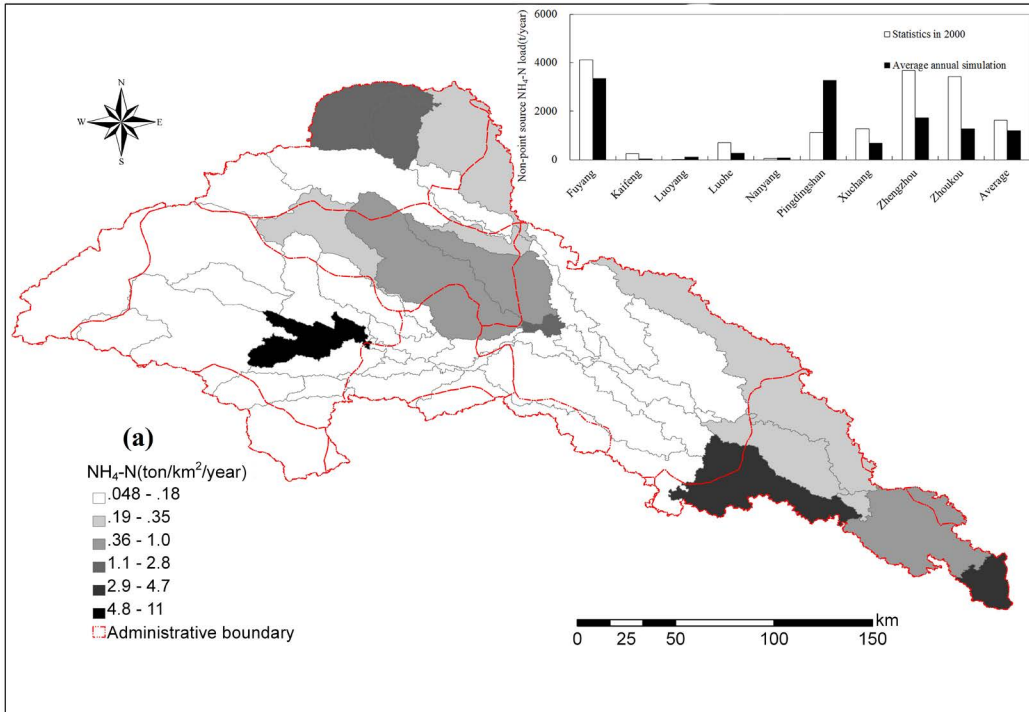


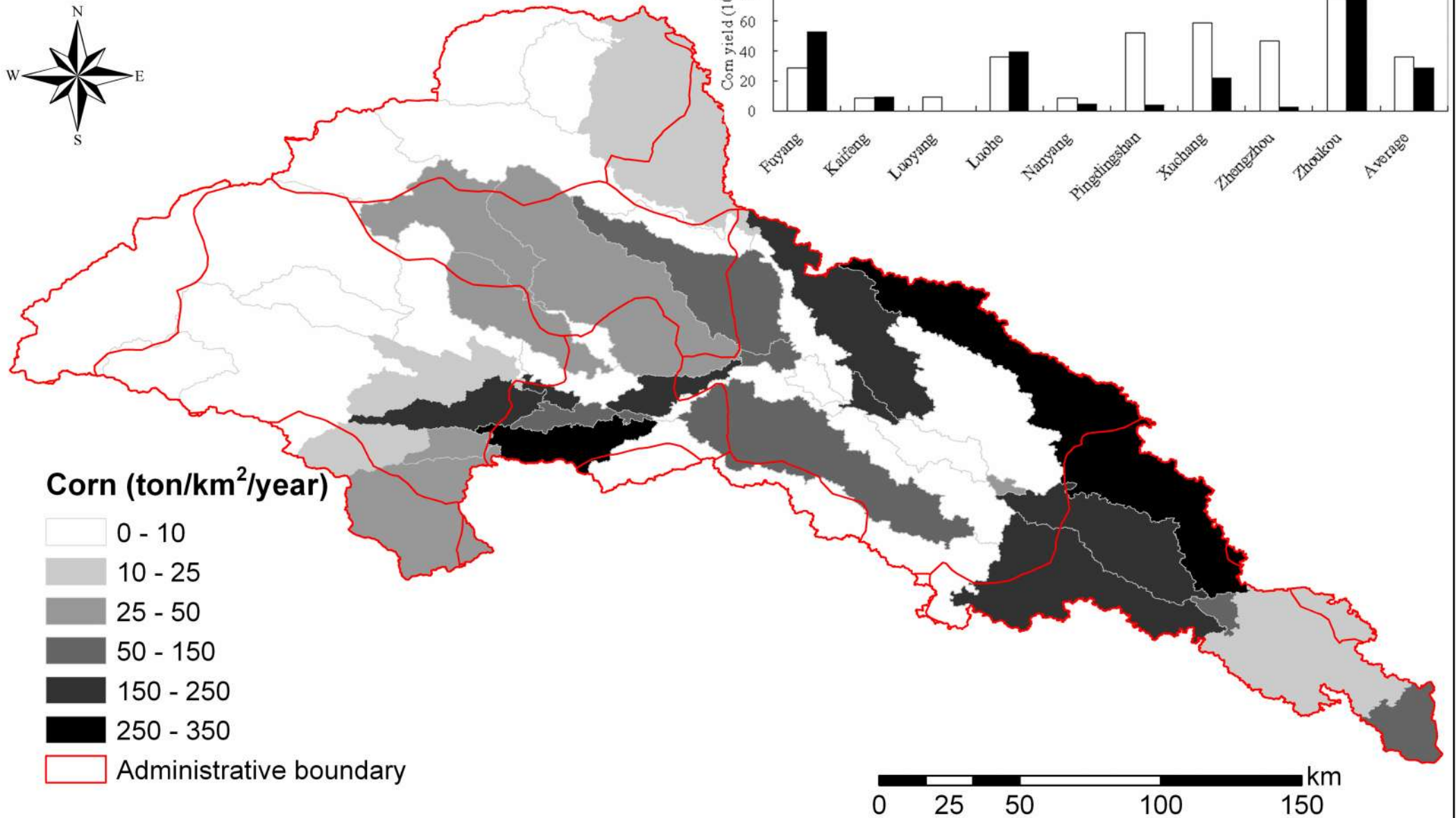
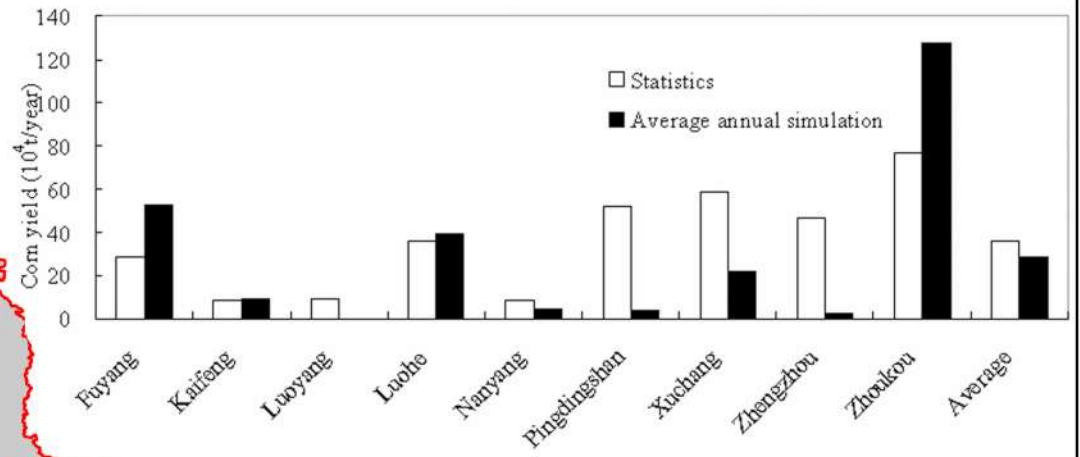
(d)











1 Supplementary material

2 1. Soil P cycle simulation (Neitsch *et al.*, 2011)

3 *Mineralization*: The mineralized P is added to the solution P pool. The amounts of
4 active and stable organic P pools ($orgP_{act}$ and $orgP_{sta}$, kg/ha) are calculated as

$$5 \begin{cases} orgP_{act} = orgP_{hum} \cdot orgN_{act} / (orgN_{act} + orgN_{sta}) \\ orgP_{sta} = orgP_{hum} \cdot orgN_{sta} / (orgN_{act} + orgN_{sta}) \end{cases} \quad (S1)$$

6 where $orgP_{hum}$ is the humic organic P amount (kg/ha); $orgN_{act}$ and $orgN_{sta}$ are the
7 amounts of N in the active organic pool and stable organic pool (kg/ha), respectively,
8 which are simulated by DNDC.

9 The **mineralization** rate of the humus active organic P pool (*RHP*) is calculated by

$$10 RHP = 1.4 \cdot \beta_{min} \cdot (\gamma_{tmp} \cdot \gamma_{SW})^{1/2} \quad (S2)$$

11 where β_{min} is the mineralization rate of the humus active organic P; γ_{tmp} and γ_{SW} are
12 **the** reduction factors of soil temperature and moisture.

13 The **mineralization** rate of the residue fresh organic P pool (*RRP*) is calculated as

$$14 RRP = 0.8 \cdot \delta_P = 0.8 \cdot \beta_{rsd} \cdot \gamma_P \cdot (\gamma_{tmp} \cdot \gamma_{SW})^{1/2} \quad (S3)$$

15 where δ_P and β_{rsd} are the residue decay rate and the mineralization rate of **the** residue
16 fresh organic P. γ_P is the P cycling residue composition factor.

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

$$18 DRP = 0.2 \cdot \delta_P \quad (S4)$$

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

20 and between active and stable mineral pools ($P_{act|sta}$, kg/ha) are

$$21 P_{sol|act} = \begin{cases} P_{sol} - \min P_{act} \cdot pai / (1 - pai) & \text{if } P_{sol} > \min P_{act} \cdot pai / (1 - pai) \\ 0.1 \cdot [P_{sol} - \min P_{act} \cdot pai / (1 - pai)] & \text{if } P_{sol} < \min P_{act} \cdot pai / (1 - pai) \end{cases} \quad (S5)$$

22 and

$$23 P_{act|sta} = \begin{cases} 0.0006 \cdot (4 \cdot \min P_{act} - \min P_{sta}) & \text{if } \min P_{sta} < 4 \cdot \min P_{act} \\ 0.00006 \cdot \beta_{eqP} \cdot (4 \cdot \min P_{act} - \min P_{sta}) & \text{if } \min P_{sta} > 4 \cdot \min P_{act} \end{cases} \quad (S6)$$

1 respectively. Where P_{sol} , $minP_{act}$ and $minP_{sta}$ are the amounts of soluble, mineral active
 2 and stable P (kg/ha), respectively; pai is P availability index.

3

4 **2. Crop growth module**

5 **2.1 Crop yield (Williams *et al.*, 1989)**

6 The crop growth depends on the accumulation of thermal time (Sharpley and Williams,
 7 1990). The daily thermal time (HU , °C) and the thermal time index for j^{th} crop (HUI)
 8 are calculated as:

$$9 \quad \begin{cases} HU_K = (T_{mx,K} + T_{mn,K})/2 - T_{b,j} \\ HUI_i = \sum_{K=1}^i HU_K / PHU_j \end{cases} \quad (S7)$$

10 where $T_{mx,K}$ and $T_{mn,K}$ are the maximum and minimum temperatures (°C) on the K^{th} day,
 11 respectively; $T_{b,j}$ is the base temperature of the j^{th} crop (°C). Crop growth will stop when
 12 HU_K is below 0.0. PHU_j is the required cumulative thermal time for the j^{th} crop from
 13 sowing to physical maturity (°C). The range of HUI is from 0.0 at sowing to 1.0 at
 14 maturity. i is the total days of crop growth.

15 The daily potential biomass accumulation (ΔB_p , t/ha/d) is estimated as follow:

$$16 \quad \begin{aligned} \Delta B_{p,i} &= 0.001 \cdot BE_i \cdot PAR_i \cdot [1 + \Delta HRLT_i]^3 \\ &= 0.0005 \cdot BE_i \cdot RA_i \cdot [1 - \exp(-0.65 \cdot LAI)] \cdot [1 + \Delta HRLT_i]^3 \end{aligned} \quad (S8)$$

17 where BE is the crop parameter for converting energy to biomass (kg·ha·m²/MJ);
 18 $HRLT$ and $\Delta HRLT$ are length of a day (hr) and its variation (hr/d), respectively; PAR
 19 is intercepted photosynthetic active radiation (MJ/m²). RA is solar radiation (MJ/m²)
 20 and LAI is leaf area index (m²/m²), which is a function of heat units, crop stress and
 21 crop development stages.

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

$$23 \quad \begin{aligned} LAI_i &= LAI_{i-1} + \Delta LAI \\ &= LAI_{i-1} + (\Delta HUF)(LAI_{mx})(1 - \exp(5 \cdot (LAI_{i-1} - LAI_{mx}))) \cdot \sqrt{REG_i} \end{aligned} \quad (S9)$$

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

$$25 \quad LAI_i = LAI_0 \cdot (1 - HUI_i / 1 - HUI_0)^{ad_j} \quad (S10)$$

1 where HUF is a thermal time factor; REG is minimum crop stress factor; Ad is a
 2 parameter controlled LAI decline rate for crop j ; HUI_0 is the HUI value when LAI begins
 3 to decline.

4 The biomass accumulation is constrained by the stresses of soil water, temperature,
 5 nutrient (N and P) and aeration conditions.

$$6 \quad \Delta B = \Delta B_p \cdot REG = \Delta B_p \cdot \min(WS, TS, SN, SP, AS) \quad (S11)$$

7 where REG is the crop growth regulating factor.

$$8 \quad \text{The water stress: } WS_i = \sum_{l=1}^M U_{l,i} / E_{P,i} \quad (S12)$$

$$9 \quad \text{The temperature stress: } TS_i = \sin[\pi \cdot (T_{g,i} - T_{b,j}) / 2(T_{o,j} - T_{b,j})] \quad 0 \leq TS_i \leq 1 \quad (S13)$$

$$10 \quad \text{The N stress: } \begin{cases} SN_{S,i} = 2[1 - \sum_{K=1}^i UN_K / (c_{NB,i} \cdot B_i)] \\ SN_i = 1 - SN_{S,i} / [SN_{S,i} + \exp(3.39 - 10.93SN_{S,i})] \end{cases} \quad (S14)$$

$$11 \quad \text{The P stress: } \begin{cases} SP_{S,i} = 2[1 - \sum_{K=1}^i UP_K / (c_{NP,i} \cdot B_i)] \\ SP_i = 1 - SP_{S,i} / [SP_{S,i} + \exp(3.39 - 10.93SP_{S,i})] \end{cases} \quad (S15)$$

$$12 \quad \text{The aeration stress: } \begin{cases} SAT = SW1/PO1 - CAF_j \\ AS_{S,i} = 1 - SAT / [SAT + \exp(-1.291 - 56.1 \cdot SAT)] \quad SAT > 0.0 \end{cases}$$

13 (S16)

14 where T_g and T_0 are the average daily soil surface temperature and the optimal
 15 temperature ($^{\circ}\text{C}$) for crop j , respectively; SAT is the saturated soil moisture (mm/mm);
 16 $SW1$ and $PO1$ are the soil moisture (mm/mm) and porosity of the top 1m of soil,
 17 respectively; CAF is critical aeration factor for crop j ; AS is aeration stress factor.

18 The crop yield (YLD , t/ha) is estimated using the harvest index, viz.:

$$19 \quad YLD_j = HI_j \cdot B_{AG} \quad (S17)$$

20 where HI is harvest index for crop j ; B_{AG} is the above-ground biomass (t/ha)..

21 2.2 Water use

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

$$U_{p,i} = E_{p,i} \cdot [1 - \exp(-\Lambda \cdot Z/RZ)] / [1 - \exp(-\Lambda)] \quad (\text{S18})$$

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

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

where $U_{p,i}$ is the total water used to depth Z m on day i (mm); RZ is the root zone depth (m); Λ is a water use distribution parameter.

Restricted by soil moisture, the water use ($U_{l,i}$, mm/day) in layer l on day i is calculated with the following equations (Jones and Kiniry, 1986).

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

(S20)

2.3 Nutrient uptake

The daily crop nutrient (N and P) demands are the differences between crop nutrient demands and potential nutrient contents.

$$\begin{cases} UND_i = c_{NB,i} \cdot B_i - \sum_{K=1}^i UN_K \\ UPD_i = c_{PB,i} \cdot B_i - \sum_{K=1}^i UP_K \end{cases} \quad (\text{S21})$$

where UND and UNP are the potential daily N and P demands, respectively (kg/ha); UN and UP are the actual uptakes of N and P , respectively (kg/ha); c_{NB} and c_{NP} are the optimal N and P concentrations of the crop, respectively (kg/t); B is the daily biomass accumulation (t/ha).

The actual soluble N ($\text{NO}_3\text{-N}$ and $\text{NH}_4\text{-N}$) mass uptaken by crops is calculated as.

$$\begin{cases} UN_{l,i} = U_{l,i} \cdot (WN_l / SW_l)_i \\ UNS_i = \sum_{K=1}^M UN_{l,i} \end{cases} \quad (\text{S22})$$

where $UN_{l,i}$ is the actual uptakes of N in the layer l on day i . WN is $\text{NO}_3\text{-N}$ or $\text{NH}_4\text{-N}$ amount in the soil (kg/ha). The amount of N supplied by soil (UNS) is estimated by summing UN of all layers (kg/ha).

1 The soil P availability is calculated as.

$$2 \begin{cases} UPS_i = 1.50 \cdot UPD_i \cdot \sum_{l=1}^M LF_{u,l} \cdot (RW_l / RWT_i) \\ LF_{u,l} = 0.1 + 0.9 \cdot c_{LP,l} / [c_{LP,l} + 117 \cdot \exp(-0.283 \cdot c_{LP,l})] \end{cases} \quad (S23)$$

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

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

$$7 \begin{cases} WFX_i = FXR_i \cdot UND_i & WFX \leq 6.0 \\ FXR = \min(1.0, FXW, FXN) \cdot FXG \end{cases} \quad (S24)$$

8 where FXG is the growth stage factor; FXW and FXN are the factors of soil water and
9 NO_3-N , respectively. All of these factors are calculated using the follow equations.

$$10 \begin{cases} 0.0 & HUI_i \leq 0.15, HUI_i \geq 0.75 \\ 6.67HUI_i - 1.0 & 0.15 < HUI_i \leq 0.3 \\ 1.0 & 0.3 < HUI_i \leq 0.55 \\ 3.75 - 5.0HUI_i & 0.55 < HUI_i < 0.75 \end{cases} \quad (S25)$$

$$11 \quad FXW_i = (SW_{0.3,i} - WP_{0.3}) / 0.85 \cdot (FC_{0.3} - WP_{0.3}) \quad SW_{0.3} < 0.85(FC_{0.3} - WP_{0.3}) + WP_{0.3} \quad (S26)$$

$$12 \quad FXN_i = \begin{cases} 0.0 & WNO_3 > 300 \text{ kg} \cdot \text{ha}^{-1} \cdot \text{m}^{-1} \\ 1.5 - 0.005 \cdot WNO_3 / RD & 100 < WNO_3 \leq 300 \\ 1.0 & WNO_3 \leq 100 \end{cases} \quad (S27)$$

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

15

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

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

$$19 \quad Y = \begin{cases} (0.646EI + 0.45Q \cdot q_p^{0.333}) \cdot K \cdot CE \cdot PE \cdot LS & Q > 0. \\ 0 & Q = 0. \end{cases} \quad (S28)$$

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

4 LS is the factor of slope length and steepness:

$$5 \quad \begin{cases} LS = (\lambda/22.1)^\xi (65.41S^2 + 4.56S + 0.065) \\ \xi = 0.6 \cdot [1 - \exp(-35.835S)] \end{cases} \quad (S29)$$

6 CE is the crop management factor:

$$7 \quad CE = (0.8 - CE_{mn,j}) \exp(-0.00115CV) + CE_{mn,j} \quad (S30)$$

8 EI is the precipitation energy factor:

$$9 \quad EI = P \cdot [12.1 + 8.9 \cdot (\log r_p - 0.434) \cdot r_{0.5}] / 1000 \quad (S31)$$

10 where S and λ are the land surface slope (m/m) and slope length (m), both of which
 11 are obtained during the procedure of preparing the spatial simulation units; ξ is a
 12 parameter dependent upon slope; $CE_{mn,j}$ is the minimum crop management factor of
 13 crop j ; CV is soil cover (above ground biomass and residue) (kg/ha); P is the daily
 14 precipitation (mm); r_p , $r_{0.5}$ is the peak precipitation rate and maximum 0.5 h
 15 precipitation intensity (mm/hr). The value of r_p is obtained according to the exponential
 16 precipitation distribution.

17

18 **4. Overland water quality module**

19 **4.1 Nutrient loss in urban and rural area**

20 Generally, the inhabitant and industrial sewage in the urban area are collected, treated
 21 and discharged into river system from the wastewater discharge outlets. This amount
 22 of nutrient flux is the model input as the point source pollutant load. The nonpoint
 23 source nutrient loss in urban area takes place along the overland flow and is estimated
 24 using the export coefficient model (Johnes, 1996).

$$25 \quad V_{ur_N} = 100 \cdot c_{ur_N} \cdot Area_{urban} \quad (S32)$$

1 where V_{ur_N} , c_{ur_N} and $Area_{urban}$ are the amount of nutrient loss in the urban area (kg);
 2 the export coefficient (kg/ha/year) and urban area (km²), respectively.
 3 The farm manure of rural living and livestock farming is also considered as one of
 4 important nonpoint source of nutrient due to the deficiency of sewage treatment
 5 facilities in the rural area. The total loss is estimated as.

$$6 \quad \begin{cases} V_{liv_N} = c_{liv_N} \cdot Pop_{rural} \\ V_{lst_N} = c_{lst_N} \cdot Pop_{stock} \end{cases} \quad (S33)$$

7 where V_{liv_N} and V_{lst_N} are the amounts of nutrient loss from living and livestock
 8 farming in the rural area, respectively (kg/year). C_{liv_N} and C_{lst_N} are the export
 9 coefficients of living (kg/day/person) and livestock (kg/day/animal), respectively;
 10 Pop_{rural} and Pop_{stock} are the population and the animal stock, respectively.

11 4.2 Nutrient loss of soil layer

12 The loss of soluble nutrient is considered to happen in both upper and lower soil
 13 layers. The loss weights of NO₃-N, NH₄-N and soluble P are calculated using the
 14 equations (Williams *et al.*, 1989), respectively.

$$15 \quad \begin{cases} V_{N_up} = W_{N_up} \cdot [1 - \exp(-\frac{R_s + R_{ss}}{UL})] \\ V_{N_low} = W_{N_low} \cdot [1 - \exp(-\frac{R_{bs}}{UL})] \end{cases} \quad (S34)$$

16 where W_{N_up} and W_{N_low} are the soluble nutrient weights in the upper and lower soil
 17 layers, respectively (kg/ha); UL is the maximum soil moisture (mm); V_{N_up} and V_{N_low}
 18 is the soluble nutrient loss in the upper and lower soil layers, respectively (kg/ha); R_s ,
 19 R_{ss} and R_{bs} are the surface runoff, interflow and baseflow (mm), respectively, which
 20 are obtained from the hydrological cycle module.

21 The amount of insoluble nutrients migrated with the sediment is estimated using the
 22 equation (Neitsch *et al.*, 2011)

$$23 \quad Y_{ON} = 0.001 \cdot Y \cdot c_{ON} \cdot ER \quad (S35)$$

24 where Y_{ON} is loss of organic N or P (kg/ha); c_{ON} is the insoluble nutrient concentration
 25 in the soil layer (g/m³); ER is enrich ratio.

4.3 Overland migration (Neitsch *et al.*, 2011)

$$N_{overl,i} = (N'_{overl,i} + N_{stor,i-1}) \cdot [1 - \exp(-T_{retain}/T_{route})] \quad (S36)$$

where $N_{overl,i}$ is the amount of overland pollutant discharged into river system on day i including sediment (tons/day), soluble and insoluble nutrient (kg/day); $N'_{overl,i}$ and $N_{stor,i}$ are pollutant loads generated in the sub-basin on day i , retained from the previous day (tons for sediment, kg for nutrient), respectively. T_{retain} and T_{route} are the retain time and routing time of flow (days), respectively.

5. Water quality module of water bodies

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

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

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

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

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

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

Table S1. All the parameters in the extended model

ID	Variables	Definition	Unit	Affected components
Sub-basin parameters				
1	W_m	Minimum soil moisture	none	flow
2	W_w	Soil moisture at permanent wilting point	none	flow
3	W_{fc}	Field capacity of soil	none	flow
4	$W_{sat,u}$	Saturated moisture capacity of upper soil layer	none	flow
5	$W_{sat,l}$	Saturated moisture capacity of lower soil layer	none	flow
6	g_1	Basic surface runoff coefficient	none	flow
7	g_2	Influence coefficient of soil moisture	none	flow
8	K_{ET}	Adjustment factor of evapotranspiration	none	flow
9	K_{ss}	Interflow yield coefficient	none	flow
10	T_g	Delay time for aquifer recharge	day	flow
11	K_{bs}	Baseflow yield coefficient	none	flow
12	K_{sat}	Steady state infiltration rate of soil	mm/hr	flow
13	k_{fmx}	Ratio of state infiltration rate to maximum rate in soil	none	flow
14	$DtoW$	Ratio of width to depth of channel	none	flow
15	rch_k	Infiltration rate of channel	mm/hr	sediment
16	ch_cov	Channel cover factor	none	sediment
17	ch_erod	Channel erodibility factor	cm/hr/Pa	sediment
18	$R_{set}(\text{algae})$	Settling rate of Algae at 20 °C	mg/day	algae
19	$R_{set}(\text{solP})$	Settling rate of soluble P at 20 °C	mg/m ² /day	P
20	$R_{set}(\text{NH}_4)$	Settling rate of NH ₄ -N at 20 °C in channel	mg/m ² /day	N
21	$R_{set}(\text{orgN})$	Settling rate of organic N at 20 °C in channel	day ⁻¹	N
22	$R_{set}(\text{orgP})$	Settling rate of organic P at 20 °C in channel	day ⁻¹	P
23	$R_d(\text{COD})$	COD deoxygenation rate at 20 °C in channel	day ⁻¹	COD
24	Rch_k_1	Reaeration coefficients at 20 °C in channel	day ⁻¹	DO
25	$R_{set}(\text{COD})$	Settling rate of COD at 20 °C in channel	day ⁻¹	COD
26	Rch_k_2	DO adsorption rate of sediment at 20 °C in channel	day ⁻¹	DO
27	$R_d(\text{NH}_4)$	Bio-oxidation rate of NH ₄ -N at 20 °C in channel	day ⁻¹	N
28	$R_d(\text{NO}_2)$	Oxidation rate of NO ₂ -N to NO ₃ -N at 20 °C in channel	day ⁻¹	N
29	$R_d(\text{orgN})$	Hydrolysis rate of organic N to NH ₄ -N at 20 °C in channel	day ⁻¹	N
30	$R_d(\text{orgP})$	Hydrolysis rate of organic P to soluble P at 20 °C in channel	day ⁻¹	N
31	$CtoB$	Relationship between COD and BOD	none	COD
32	res_k	Infiltration rate in reservoir or sluice	mm/hr	flow
33	$K_{set}(\text{COD})$	Settling rate of COD at 20 °C in reservoir or sluice	m/year	COD
34	$K_{set}(\text{NH}_4)$	Settling rate of NH ₄ -N at 20 °C in reservoir or sluice	m/year	N
35	$K_{set}(\text{NO}_2)$	Settling rate of NO ₂ -N at 20 °C in reservoir or sluice	m/year	N
36	$K_{set}(\text{NO}_3)$	Settling rate of NO ₃ -N at 20 °C in reservoir or sluice	m/year	N
37	$K_{set}(\text{orgN})$	Settling rate of organic N at 20 °C in reservoir or sluice	m/year	N
38	$K_{set}(\text{orgP})$	Settling rate of organic P at 20 °C in reservoir or sluice	m/year	P

39	$K_{set}(\text{solP})$	Settling rate of soluble P at 20 °C in reservoir or sluice	m/year	P
40	$K_{set}(\text{DO})$	Settling rate of DO at 20 °C in reservoir or sluice	m/year	DO
41	$K_{set}(\text{algae})$	Settling rate of algae at 20 °C in reservoir or sluice	m/year	algae
42	$K_{set}(\text{TN})$	Settling rate of TN at 20 °C in reservoir or sluice	m/year	N
43	$K_{set}(\text{TP})$	Settling rate of TP at 20 °C in reservoir or sluice	m/year	P
44	$K_d(\text{COD})$	COD deoxygenation rate in reservoirs at 20 °C	day ⁻¹	COD
45	res_kl	Reaeration coefficients at 20 °C in reservoir or sluice	day ⁻¹	DO
46	$K_d(\text{NH}_4)$	Bio-oxidation rate of NH ₄ -N in reservoir at 20 °C	day ⁻¹	N
47	$K_d(\text{NO}_2)$	Oxidation rate of NO ₂ -N to NO ₃ -N at 20 °C in reservoir or sluice	day ⁻¹	N
48	$K_d(\text{orgN})$	Hydrolysis rate of organic N to NH ₄ -N at 20 °C in reservoir or sluice	day ⁻¹	N
49	$K_d(\text{orgP})$	Hydrolysis rate of organic P to soluble P at 20 °C in reservoir or sluice	day ⁻¹	P
50	$K_{scu}(\text{COD})$	Resuspension rate of COD at 20 °C in reservoir or sluice	m/year	COD
51	$K_{scu}(\text{NH}_4)$	Resuspension rate of NH ₄ -N at 20 °C in reservoir or sluice	m/year	N
52	$K_{scu}(\text{NO}_2)$	Resuspension rate of NO ₂ -N at 20 °C in reservoir or sluice	m/year	N
53	$K_{scu}(\text{NO}_3)$	Resuspension rate of NO ₃ -N at 20 °C in reservoir or sluice	m/year	N
54	$K_{scu}(\text{orgN})$	Resuspension rate of organic N at 20 °C in reservoir or sluice	m/year	N
55	$K_{scu}(\text{orgP})$	Resuspension rate of organic P at 20 °C in reservoir or sluice	m/year	P
56	$K_{scu}(\text{solP})$	Resuspension rate of soluble P at 20 °C in reservoir or sluice	m/year	P
57	$K_{scu}(\text{DO})$	Resuspension rate of DO at 20 °C in reservoir or sluice	m/year	DO
58	$K_{scu}(\text{algae})$	Resuspension rate of algae at 20 °C in reservoir or sluice	m/year	algae
59	$K_{scu}(\text{TN})$	Resuspension rate of TN at 20 °C in reservoir or sluice	m/year	N
60	$K_{scu}(\text{TP})$	Resuspension rate of TP at 20 °C in reservoir or sluice	m/year	P
61	$K_{bur}(\text{COD})$	Decay rate of COD at 20 °C in reservoir or sluice	m/year	COD
62	$K_{bur}(\text{NH}_4)$	Decay rate of NH ₄ -N at 20 °C in reservoir or sluice	m/year	N
63	$K_{bur}(\text{NO}_2)$	Decay rate of NO ₂ -N at 20 °C in reservoir or sluice	m/year	N
64	$K_{bur}(\text{NO}_3)$	Decay rate of NO ₃ -N at 20 °C in reservoir or sluice	m/year	N
65	$K_{bur}(\text{orgN})$	Decay rate of organic N at 20 °C in reservoir or sluice	m/year	N
66	$K_{bur}(\text{orgP})$	Decay rate of organic P at 20 °C in reservoir or sluice	m/year	P
67	$K_{bur}(\text{solP})$	Decay rate of soluble P at 20 °C in reservoir or sluice	m/year	P
68	$K_{bur}(\text{DO})$	Decay rate of DO at 20 °C in reservoir or sluice	m/year	DO
69	$K_{bur}(\text{algae})$	Decay rate of algae at 20 °C in reservoir or sluice	m/year	algae
70	$K_{bur}(\text{TN})$	Decay rate of TN at 20 °C in reservoir or sluice	m/year	N
71	$K_{bur}(\text{TP})$	Decay rate of TP at 20 °C in reservoir or sluice	m/year	P
72	$usle_k$	Soil erodibility factor of USLE equation	none	sediment
73	$usle_p$	Erosion control practice factor of USLE equation	none	sediment
74	$MicrIn$	Microbe index	none	C, N
75	K_l	Decomposition rate of labile organic C	day ⁻¹	C
76	μ_{CLAY}	Reduction factor of clay content on organic matter decomposition	none	C
77	μ_t	Reduction factor of soil temperature on growth of denitrifier or nitrifier	none	N
78	S	Labile fraction of organic C compounds	none	C
79	kr_{cvl}	Decomposition rate of very labile organic C in residue pool	day ⁻¹	C
80	kr_{cl}	Decomposition rate of labile organic C in residue pool	day ⁻¹	C
81	kr_{cr}	Decomposition rate of stable organic C in residue pool	day ⁻¹	C

82	km_{sc}	Decomposition rate of stable organic C in microbial biomass pool	day ⁻¹	C
83	km_{cl}	Decomposition rate of labile organic C in microbial biomass pool	day ⁻¹	C
84	km_h	Decomposition rate of microbial biomass to humands	day ⁻¹	C
85	K_C	Half velocity constant of organic C on denitrifier biomass growth	none	N
86	$K_{N_xO_y}$	Half velocity constant of NO ₃ -N, NO ₂ -N, NO and N ₂ O on denitrifier biomass growth	none	N
87	u_{NO_3}	Maximum growth rate of NO ₃ -N denitrifier	day ⁻¹	N
88	u_{NO_2}	Maximum growth rate of NO ₂ -N denitrifier	day ⁻¹	N
89	u_{NO}	Maximum growth rate of NO denitrifier	day ⁻¹	N
90	u_{N_2O}	Maximum growth rate of N ₂ O denitrifier	day ⁻¹	N
91	M_C	Maintenance coefficient of C	hr ⁻¹	C
92	Y_C	Maximum growth yield of soluble C	kg/ha/hr	C
93	M_{NO_3}	Maintenance coefficient of NO ₃ -N	hr ⁻¹	N
94	Y_{NO_3}	Maximum growth yield of NO ₃ -N	kg/ha/hr	N
95	$CDR_{D:N}$	C:N ratio in bacteria	none	N
96	M_{NO_2}	Maintenance coefficient of NO ₂ -N	hr ⁻¹	N
97	Y_{NO_2}	Maximum growth yield of NO ₂ -N	kg/ha/hr	N
98	M_{NO}	Maintenance coefficient of NO	hr ⁻¹	N
99	Y_{NO}	Maximum growth yield of NO	kg/ha/hr	N
100	M_{N_2O}	Maintenance coefficient of N ₂ O	hr ⁻¹	N
101	Y_{N_2O}	Maximum growth yield of N ₂ O	kg/ha/hr	N
102	$\mu_{SW,n}$	Soil water content adjusted factor for denitrification	none	C, N
103	β_{min}	Mineralization rate of humus active organic P	day ⁻¹	P
104	β_{rsd}	Mineralization rate of residue fresh organic P	day ⁻¹	P
Watershed parameters				
105	$C_{ur}(COD)$	Export coefficient of COD load in urban area	kg/ha/year	COD
106	$C_{ur}(NH_4)$	Export coefficient of NH ₄ -N load in urban area	kg/ha/year	N
107	$C_{ur}(TN)$	Export coefficient of TN load in urban area	kg/ha/year	N
108	$C_{ur}(TP)$	Export coefficient of TP load in urban area	kg/ha/year	P
109	$C_{ur}(COD)$	Export coefficient of COD load in unused area	kg/ha/year	COD
110	$C_{ur}(NH_4)$	Export coefficient of NH ₄ -N load in unused area	kg/ha/year	N
111	$C_{ur}(TN)$	Export coefficient of TN load in unused area	kg/ha/year	N
112	$C_{ur}(TP)$	Export coefficient of TP load in unused area	kg/ha/year	P
113	R_{ur}	Loss rate of non-point source load from soil layer	none	pollutant load
114	$C_{liv}(COD)$	Export coefficient of COD load from living in rural area	kg/year	COD
115	$C_{liv}(NH_4)$	Export coefficient of NH ₄ -N load from living in rural area	kg/year	N
116	$C_{liv}(TN)$	Export coefficient of TN load from living in rural area	kg/year	N
117	$C_{liv}(TP)$	Export coefficient of TP load from living in rural area	kg/year	P
118	$C_{lst}(COD)$	Export coefficient of COD load from livestock in rural area	kg/year	COD
119	$C_{lst}(NH_4)$	Export coefficient of NH ₄ -N load from livestock in rural area	kg/year	N
120	$C_{lst}(TN)$	Export coefficient of TN load from livestock in rural area	kg/year	N
121	$C_{lst}(TP)$	Export coefficient of TP load from livestock in rural area	kg/year	P
122	R_{liv}	Loss rate of non-point source load from living	none	pollutant load

123	R_{lst}	Loss rate of non-point source load from livestock	none	pollutant load
124	$C_{pcp}(\text{COD})$	COD concentration in precipitation	mg/L	COD
125	$C_{pcp}(\text{NH}_4)$	NH ₄ -N concentration in precipitation	mg/L	N
126	$C_{pcp}(\text{TN})$	TN concentration in precipitation	mg/L	N
127	$C_{pcp}(\text{TP})$	TP concentration in precipitation	mg/L	P
128	SF_{tmp}	Snowfall temperature	°C	flow
129	SM_{tmp}	Snow melt base temperature	°C	flow
130	SMF_{mx}	Melt factor for snow on June 21	mm/day	flow
131	SMF_{mn}	Melt factor for snow on December 21	mm/day	flow
132	$TIMP$	Snow pack temperature lag factor	none	flow
133	$Coefrad$	Factor of maximum possible radiation to net radiation	none	flow
134	SC_{max}	Minimum snow water content that corresponds to 100% snow cover	mm	flow
135	SC_{50}	Fraction of snow volume represented by SCMX that corresponds to 50% snow cover	none	flow
136	SC_1	Coefficients that define shape of snow curve 95% coverage at 100% snow cover	none	flow
137	SC_2	Coefficients that define shape of snow curve 50% coverage at 100% snow cover	none	flow
138	$Surlag$	Surface runoff lag time	day	flow
139	n_{ch}	Roughness of Channel	none	flow
140	msk_x	Weighting factor in Muskingum equation	none	flow
141	msk_k	Storage time constant of channel in Muskingum equation	day	flow
142	AI_1	Fraction of algal biomass that is N	none	N
143	AI_2	Fraction of algal biomass that is P	none	P
144	AI_3	Adjusted rate of oxygen production per unit of algal photolysis	none	DO
145	AI_4	Adjusted rate of oxygen uptake per unit of algal respiration	none	DO
146	AI_5	Adjusted rate of oxygen uptake per unit of NH ₄ -N oxidation	none	N
147	AI_6	Adjusted rate of oxygen uptake per unit of NO ₂ -N oxidation	none	N
148	AI_7	Adjusted rate of NH ₄ -N oxidation to NO ₂ -N	none	N
149	g_{max}	Maximum specific algal growth rate at 20°C	day ⁻¹	algae
150	$RHOQ$	Algal respiration rate at 20°C	day ⁻¹	algae
151	$TFACT$	Fraction of solar radiation computed in temperature heat balance	none	algae
152	K_l	Half-saturation coefficient for light	kJ/m ²	algae
153	K_N	Michaelis-Menton half-saturation constant for N	mg/L	algae
154	K_P	Michaelis-Menton half-saturation constant for P	mg/L	algae
155	Lec	Non-algal portion of light extinction coefficient	m ⁻¹	algae
156	Lec_1	Linear algal self-shading coefficient	m ⁻¹ ·(μg/L) ⁻¹	algae
157	Lec_2	Nonlinear algal self-shading coefficient	m ⁻¹ ·(μg/L) ^{-2/3}	algae
158	P_N	Algal preference factor for ammonia	none	N
159	PRF	Peak rate adjustment factor for sediment routing in channel	none	sediment
160	SP_{con}	Linear parameter for calculating maximum transport capacity of sediment in channel	none	sediment

161	<i>SPexp</i>	Exponent parameter for calculating maximum transport capacity of sediment in channel	none	sediment
162	<i>f_Ph</i>	Flood PH value	none	C, N
163	<i>rcn_{rvl}</i>	Ratio of C/N of very labile litter	none	C, N
164	<i>rcn_{rl}</i>	Ratio of C/N of labile litter	none	C, N
165	<i>rcn_{rr}</i>	Ratio of C/N of resistant litter	none	C, N
166	<i>rcn_b</i>	Ratio of C/N of labile biomass	none	C, N
167	<i>rcn_h</i>	Ratio of C/N of labile humus	none	C, N
168	<i>rcn_m</i>	Ratio of C/N of humads	none	C, N
169	<i>pavi</i>	P availability index	none	C, N
170	<i>TtoC</i>	Relationship between TOC and COD	none	COD
171-182	<i>rpnt_{01~12}</i>	Ratio of point pollutant source from Jan. to Dec.	none	pollutant load

1

2

1 Table S2. The detailed information of data sets for the case study

Category	Data	Spatial scale	Temporal scale	Source
GIS	DEM	Grid: 90m*90 m	none	Institute of Geographic Science and Natural Resources
	Land use	1:1,000,000	none	Research, Chinese Academy of Sciences
	Soil	1:4,000,000	none	
Weather	Precipitation	65 stations	daily (from 2003 to 2008)	Hydrological Yearbooks of Henan Province, China
	Maximum and minimum temperature	6 stations	daily (from 2003 to 2008)	National Meteorological Infomation Center of China
Hydrology	Total runoff, high and low flows	6 stations	daily (from 2003 to 2008)	Hydrological Yearbooks of Henan Province, China
Water quality	Wastewater discharge outlets and the discharge load (wastewater, NH ₄ -N, etc.)	over 200 outlets	annual (from 2003 to 2008)	Water Resources Protection Bureau of Huai River Basin, China
	Water quality variable concentrations (NH ₄ -N)	6 stations	daily (from 2003 to 2008)	Water Resources Protection Bureau of Huai River Basin, China
	Nonpoint source load (NH ₄ -N)	9 administrative regions	average annual (from 2003 to 2005)	Huai River Commission, China
Ecology	Corn yield	9 administrative regions	average annual (from 2003 to 2005)	Henan Statistical Yearbook, China
Economy	Populations in rural area, breeding stock of large animals and livestock, water withdrawal	9 administrative regions	annual (from 2003 to 2008)	Henan Statistical Yearbook, China
Water projects	Water storage capacities of dead, usable, flood control and maximum flood levels and the corresponding water surface areas; the relationship among water level, storage volume and outflow	5 reservoirs and 12 sluices	none	Water Resources Protection Bureau of Huai River Basin, China
Agricultural management	Fertilization and irrigation types, timing and amount, the time of seeding and harvest, crop types	9 administrative regions	average annual (from 2003 to 2008)	Henan Statistical Yearbooks, China, Wang et al., (2008) and Zhai et al. (2014)

2

3

4

5

1

2 Table S3. The agricultural management scheme in the Shaying River Catchment

Crop	Management	Time		Ratio distribution of annual TN fertilizer	Ratio distribution of annual TP fertilizer	Fertilizer intensity (kg/ha)	
		Start (month- day)	Duration (day)			TN	TP
Early rice	Base fertilization	4-1	1	0.60	0.86	40.60-86.17	25.46-59.47
	Plant	4-15	1	-	-		
	Additional Fertilization	5-1	1	0.40	0.14	27.06-57.45	4.14-9.68
Late rice	Harvest & Kill	7-31	1	-	-		
	Base fertilization	8-1	1	0.50	0.86	33.83-71.81	25.46-59.47
	Plant	8-15	1	-	-		
	Additional Fertilization	9-1	1	0.50	0.14	33.83-71.81	4.14-9.68
Winter wheat	Harvest & Kill	10-31	1	-	-		
	Base fertilization	10-1	1	0.64	0.02	43.30-271.04	0.59-4.10
	Plant	10-15	1	-	-		
Cron	Additional Fertilization	1-1	1	0.36	0.98	24.36-152.46	29.00-201.11
	Harvest & Kill	6-1	1	-	-		
	Base fertilization	6-1	1	0.41	0.88	27.74-173.63	26.05-180.59
	Plant	6-15	1	-	-		
	Additional Fertilization	7-15	1	0.59	0.12	39.92-249.86	3.55-24.62
	Harvest & Kill	9-30	1	-	-		

3

4