


Provided by the author(s) and University College Dublin Library in accordance with publisher policies. Please cite the published version when available.

<b>Title</b>	Coupling system model with fuzzy logic rules for use in runoff and total phosphorus load prediction in a catchment
<b>Author(s)</b>	Nasr, Ahmed Elssidig; Bruen, Michael
<b>Publication date</b>	2007-05
<b>Conference details</b>	Presented at the 7th International Water Association Symposium on Systems Analysis and Integrated Assessment in Water Management, Washington, D.C., 7-9 May 2007
<b>Item record/more information</b>	<a href="http://hdl.handle.net/10197/2279">http://hdl.handle.net/10197/2279</a>

Downloaded 2018-07-20T05:04:58Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd\_oa) 

Some rights reserved. For more information, please see the item record link above.



# **Coupling System Model with Fuzzy Logic Rules for Use in Runoff and Total phosphorus Load Prediction in a Catchment**

A. Nasr\*, and M. Bruen\*\*

Centre for Water Resources Research, University College Dublin, Earlsfort Terrace, Dublin 2  
(E-mails: [\*]ahmed.nasr@ucd.ie; [\*\*] michael.bruen@ucd.ie)

## **Abstract**

Tackling the problem of eutrophication in fresh waters is at the top of the agenda for the implementation of the Water Framework Directive (WFD) in Europe. The problem is caused primarily by an increase in phosphorus loading from diffuse sources. Therefore there is a need to apply appropriate measures, which are able to reduce the phosphorus diffuse pollution, at a catchment scale in each River Basin District (RBD). As the implementation of such measures disturbs the existing system in the catchment it is important to be able to predict their impact and this requires a reliable mathematical model to represent the system. In this study, a new, lumped catchment, methodology to improve on an existing diffuse phosphorus pollution model, the Grid Oriented Phosphorus Component (GOPC) model, is proposed. This methodology consists of two elements; (i) the Soil Moisture Accounting and Routing (SMAR) hydrological model was used to provide the required hydrological variables to the GOPC model; and (ii) fuzzy logic rules were formulated with the notion that each rule corresponds to a sub-model representing a particular hydrological behaviour in the catchment and the combined results of all rules give the total response. Sixteen modelling cases, each of which uses different numbers of fuzzy sub-sets for the rainfall and the evaporation, were compared for their discharge and total phosphorus (TP) simulations in a catchment in Northern Ireland. The comparison was based on the validation results as they allow testing the applicability of the models for conditions different from those used in the calibration period. Using 2 fuzzy sub-sets for the rainfall and a single fuzzy sub-set for the evaporation produced the best simulation for the discharge whereas the best TP simulation was obtained from the case of 4 rainfall fuzzy sub-sets and 3 evaporation fuzzy sub-sets.

## **Keywords**

Runoff; Phosphorus; modelling; GOPC; SMAR; Fuzzy Rules

## **INTRODUCTION**

The European Water Framework Directive (WFD) (EEC, 2000) has set a stringent target of bringing all fresh water bodies in Europe to “good status” by 2015. Therefore immediate action is required to alleviate the existing pollution pressures and eutrophication of rivers and lakes, for which phosphorus (P) pollution from diffuse sources is the main culprit (Mainstone and Parr, 2002). Before choosing management measures it is essential to predict their impact with the aid of a catchment-scale mathematical model. Three hydrological variables, soil moisture, surface runoff volume and baseflow volume have a direct influence on some of the bio-chemical processes of the soil P cycle as well as the P transportation mechanisms. Therefore a P model must always be accompanied by a hydrological model to provide values for these variables. A wide range of mathematical models are readily available for modelling the catchment system of P diffuse sources pollution, e.g. SWAT (Arnold et al., 1998), HSPF (Bicknell et al., 1997), LASCAM (Viney et al., 2000), INCA-P (Wade et al., 2002), GOPC (Nasr et al., 2005). All these simulate the temporal and spatial variations of the soil P cycle variables using the principal of

mass balance. In addition, some sort of a hydrodynamic model is incorporated to account for the transport of P to the receiving water. Despite the successful application of these models to catchments of different climatic, hydrologic, and agro-chemical conditions none of them has exhibited a consistent performance for all conditions. For instance the performance of a model could vary from one catchment to another although they both lie in the same region. Therefore there is still scope for more experimental studies to identify the factors determining the success or failure of these models. One of these factors is the degree of complexity required in the model and, in particular, whether a fully distributed model is required or whether extensions of lumped models are adequate. This is explored in this paper, by developing a new methodology to implement the Grid Oriented Phosphorus Component (GOPC) model in a lumped manner. The methodology consists of two elements; (i) using the Soil Moisture Accounting and Routing (SMAR) model as a provider of the required hydrological variables to the GOPC; and (ii) describing different sub-models using fuzzy logic rules to account for the effects of the temporal variations in the processes. The following two sections describe the basis of the GOPC and the SMAR models. These are then followed by a description of the procedure by which the fuzzy logic rules can be used in the modelling. The last three sections are devoted for describing the study catchment, discussing the results, and presenting the conclusions

### **GOPC**

The GOPC model (Nasr et al., 2005) is a generic phosphorus module developed to simulate the processes in the soil P cycle and the transportation of different P components over the land and through the sub-surface. The soil P state variables in the GOPC consist of the soil soluble P (SSP), the fresh organic P (FROP), the fixed organic P (FXOP), the easily soluble inorganic P (ESIP), and the fixed or insoluble inorganic P (FIP). The FROP represents the organic matter that easily mineralised and it consists of the manures, and the decayed plant and microbial biomass. On the other hand the FXOP contains the humus material which mineralises slowly. The inorganic P in the soil is divided into two types, the ESIP and the FIP. With respect to the P export in the GOPC model, the overland flow transports two forms of P, dissolved P (DP) and particulate P (PP), whereas the DP is the only form delivered by baseflow. The dynamic changes in the soil storage of each state variable is described by a mass balance equation which relates the input fluxes, the output fluxes, and the rate of change of the storage. All mass balance equations are solved simultaneously to obtain the mass of the each state variable at each time step (Nasr, 2004).

### **SMAR**

The model is a quasi-physical rainfall-runoff model known as the layers model because a procedure of moisture balancing between the rainfall, evaporation, and runoff is applied to the soil storage which consists of a number of layers. In this model, a water balance component is connected with a flow routing component to create an adequate conceptualization of the hydrological processes involved in flow generation. Using a number of empirical and physically plausible relationships, the non-linear water balance component distributes the available moisture between evaporation, soil storage and overland runoff. The routing component of the model simulates the flow of water across the land and inside the stream channels until it reaches a controlling point where the discharge is measured. The simulation accounts for attenuation and wave diffusion of the runoff and baseflow volumes separately with different conservative linear time-invariant elements. A number of modifications to the original structure of the model have

been introduced (Khan, 1986; Liang, 1992) and the latest version by Tan and O'Connor (1996) is used here.

### **FUZZY MODELLING APPROACH**

Implementation of most of the existing catchment models is always preceded by a long process of building a GIS database of information required by such models. This might not be appropriate when immediate answers are sought from the models. Therefore there should be a type of models which requires minimum amount of input data and at the same time can provide quick answers with similar degree of reliability to those sophisticated models. In this study, a fuzzy modelling approach is proposed as a way of building a model which uses only the available time series of the input and output variables to provide reliable estimates of the discharge and total phosphorus (TP) load. The proposed fuzzy logic model has the structure of an artificial neural network with five layers and hence is called a neuro-fuzzy model. This model is an abstract of a rule-based or knowledge-based system consisting of three conceptual components; (i) fuzzy logic; (ii) fuzzy decision rules; and (iii) fuzzy reasoning (Jang, 1997).

The function of the fuzzy logic component is to represent the uncertainty of assigning a membership for any value of the system input and output variables to certain fuzzy sets of that variable. To achieve this, a continuous and multi-valued logic membership function between 0 and 1 is defined for each of the fuzzy sets. The value obtained from this function provides a qualitative representation of the uncertainty in such a way that the range between 0 and 0.5 can be divided to encompass the various degrees of uncertainty whereas the remaining range is made for the various degrees of certainty.

The fuzzy decision rules consist of a number of fuzzy IF-THEN rules. The antecedents or premises of the IF-THEN rules define a fuzzy region of the input space while the output or the consequent parameters specify the corresponding output. Each of the IF-THEN rules describes the local behaviour of the mapping between the inputs and the outputs of the system and in this sense it can be interpreted as a sub-model of the entire system. To perform the mapping an appropriate mathematical model must be used. The discharge and TP load are the two targeted variables in the modelling here. To obtain simulations of these variables two models in a series must be used. The first model generates the discharge as its output and moreover it also produces all the hydrological variables required by the second model which simulates the TP load.

In the fuzzy reasoning component, an inference procedure is implemented whereby outputs of all the IF-THEN rules are combined and then transformed into crisp values, if they were not already so, to obtain the final outputs from the fuzzy model. A weighted average combination of the individual outputs is applied in this case. The weight given to the output of certain IF-THEN rule is determined by multiplying the values of the fuzzy membership functions of the input variables which constitute that rule.

### **STUDY CATCHMENT**

All cases of the proposed model in this study were tested by applying them to a 96 km<sup>2</sup> catchment which is part of an International RBD managed collaboratively by the Republic of Ireland and Northern Ireland. The dominant land use in this catchment is grassland which ranges in quality from unimproved pasture to improved pasture and intensively used silage meadows. Significant areas of natural vegetation also exist. Carboniferous series sandstone and limestone characterise the catchment geology while the soil is part of an extensive drumlin belt and is clay rich and highly gleyed with low infiltration rates.

## IMPLEMENTATION OF THE GOPC MODEL

The way in which the GOPC module has been formulated allows its use in conjunction with any potential hydrological model that produces the required variables. Nasr (2004) used the SHETRAN model (Ewen et al. 2000) as a hydrological model for the GOPC and applied them along with other models to simulate the total P load in three catchments in Ireland. In his study, each catchment was divided into square cells of 100 m sides and hence the model can be considered as fully distributed. In contrast, the GOPC in this present study is used in conjunction with the SMAR model and the study catchment is considered as a single lumped unit, however, with the aid of IF-THEN fuzzy logic rules this lumped catchment configuration is transformed into a combination of sub-models. Each sub-model has a separate set of parameters representing different P transport behaviours that corresponds to different patterns of the two driving variables, rainfall and evaporation. Therefore the total number of parameters ( $NP_{total}$ ) which require identification can be determined as follows:

$$NP_{total} = (NP_{SMAR} + NP_{GOPC}) * NRL + NP_{MF} \quad (1)$$

where:

$NP_{SMAR}$  : number of parameters in the SMAR model;

$NP_{GOPC}$  : number of parameters in the GOPC model;

$NRL$  : number of the fuzzy rules;

$NP_{MF}$  : total number of parameters of all membership functions of the fuzzy sub-sets of the input variables.

**Table 1.** Details of the fuzzy rule cases tested in this study

Case	Rainfall fuzzy sub-sets	Evaporation fuzzy sub-sets	Number of rules
1	1	1	1
2	1	2	2
3	1	3	3
4	1	4	4
5	2	1	2
6	2	2	4
7	2	3	6
8	2	4	8
9	3	1	3
10	3	2	6
11	3	3	9
12	3	4	12
13	4	1	4
14	4	2	8
15	4	3	12
16	4	4	16

There are 21 parameters in the GOPC model to model the TP load whereas SMAR uses 11 parameters to calculate the required hydrological variables. Table 1 shows the number of fuzzy rules in each modelling case tested in this study. The table also shows the number of fuzzy sub-sets used for rainfall and evaporation. A two parameter Gaussian function was used for each membership function of the fuzzy sub-sets of the input variables and as a result the total number of parameters is twice the number of fuzzy sub-sets. The problem of calibrating all the above

parameters is a non-linear one and was done with the Genetic algorithm (GA) (Holland, 1975). A global heuristic search method, the operation of the GA is broadly based on the Darwinian theory of ‘survival of the fittest’, as potential solutions to a given optimisation problem contend and ‘mate’ with each other to evolve improved solutions. The GA codes the parameter values as genes in a chromosome and uses probabilistic rules to advance the search process. The starting point for the operation of the GA is the random generation of an initial population of parameter sets. From this initial population, pairs of parameters sets are randomly chosen depending on their fitness evaluated on the basis of the value of the selected objective function. The chosen pairs are subsequently used to generate a new population of parameters sets (i.e. the next generation) utilising the genetic operators of ‘crossover’ and ‘mutation’ to generate ‘offspring’. The newly generated population is anticipated to be better than the older one. The process of generating new populations continues until a pre-specified ‘stopping-criterion’ is fulfilled (e.g. when the specified number of function evaluations is reached).

For each modelling case two groups of data were required. The first group, which includes rainfall, evaporation, and discharge, was used to run and calibrate the SMAR model. Similarly, the second group, which includes the estimated amount of inorganic and organic fertiliser application and the total Phosphorus (TP) load at the catchment outlet, was also used in the GOPC model. The available data covers the period from 1/10/2001 up to 31/1/2003. The first 67% of the data was used in the model calibration while the remaining 33% was used to verify the calibrated models. The Nash-Sutcliffe criterion ( $R^2$ ) (Nash and Sutcliffe, 1970) was used to compare the results of all the modelling cases.

**Table 2.**  $R^2$  values of discharge (Q) and TP simulations by SMAR and GOPC for all tested cases of the fuzzy rules

Case	Q/SMAR			TP/GOPC		
	Calibration	Verification	Rank	Calibration	Verification	Rank
1	80.49	<b><u>73.21</u></b>	2	67.20	42.62	5
2	80.12	66.24	10	68.48	41.31	6
3	81.01	64.91	13	66.18	4.86	15
4	80.96	62.25	15	67.48	29.92	11
5	81.47	<b><u>73.11</u></b>	1	68.54	36.77	9
6	82.00	72.85	3	69.99	-22.25	16
7	<b><u>82.87</u></b>	65.35	12	<b><u>72.01</u></b>	13.21	13
8	81.28	66.44	9	69.38	37.38	8
9	81.68	65.64	11	69.55	34.67	10
10	81.92	70.01	4	69.96	<b><u>53.75</u></b>	2
11	81.11	69.88	5	69.59	38.17	7
12	81.63	67.06	8	71.80	6.84	14
13	74.82	62.12	16	49.98	48.15	4
14	81.93	67.10	7	69.46	17.44	12
15	81.94	63.89	14	71.55	<b><u>54.05</u></b>	1
16	82.64	68.87	6	69.40	48.79	3

## RESULTS

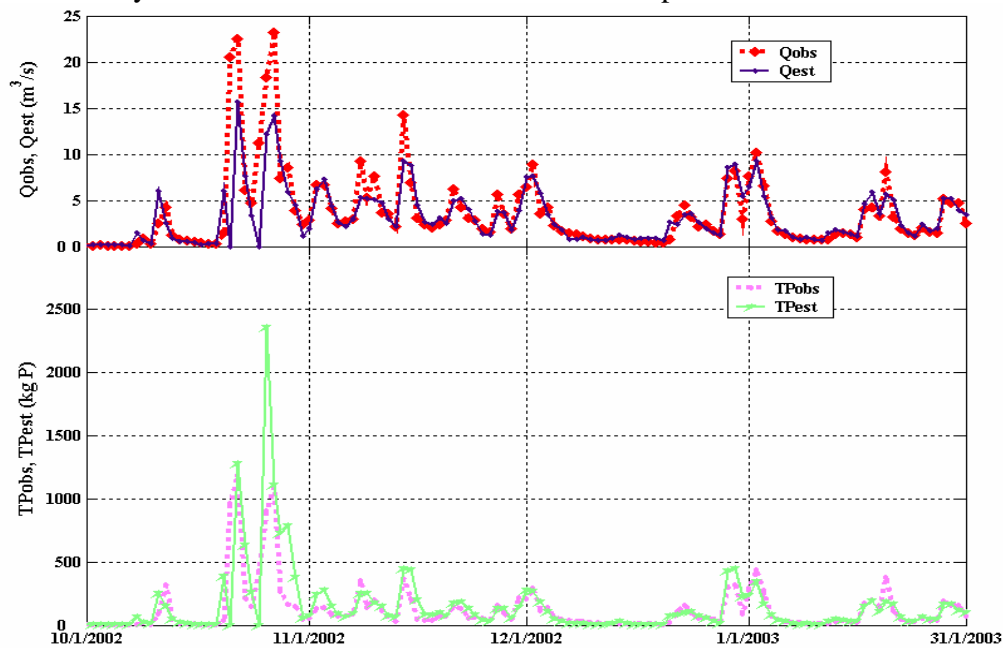
In Table 2, the  $R^2$  values for the SMAR and GOPC models during calibration and verification in each case of the fuzzy rules are given. The SMAR results indicate the performance of simulating

the discharge values at the catchment outlet whereas the GOPC values assess the TP load simulation. The performances of all cases have been ranked according to their verification results since model validation is always crucial in determining the applicability of the model. If two or more cases score similar  $R^2$  values in validation, then their rank is determined by their  $R^2$  values in calibration. Two columns showing the ranking results for SMAR and GOPC are added in Table 2.

### Flow results

Cases 5 and 1 have  $R^2$  values in validation of almost the same magnitude, however, the  $R^2$  value for case 5 in calibration outperforms the one for case 1. Therefore case 5 has been ranked the top of the list. The best  $R^2$  in calibration was obtained by case 7 but its  $R^2$  in validation was considerably poor and relegates it to 11<sup>th</sup> position in the ranking list. The worst  $R^2$  values in calibration and validation are from case 13 and both values are significantly lower than the ones for the best case.

For the best case (case 5) 2 fuzzy sub-sets were used for the rainfall while a single one was used for the evaporation and this resulted in 2 sub-models. On the other hand, the worst case (case 7) used 2 fuzzy sub-sets for the rainfall and 3 for the evaporation.

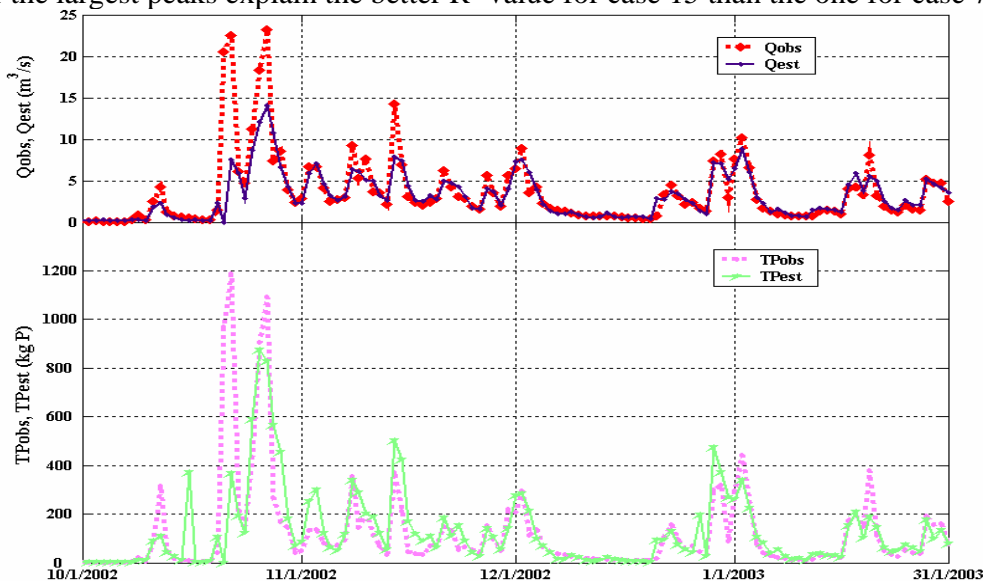


**Figure 1.** Plots of discharge (Q) and TP during validation for case 7

### TP results

The  $R^2$  value of case 7 for TP is the best and this case was also the best for the flow. This is not a surprising result because the flow simulation has a direct influence on the TP simulation and hence in calibration the best results for both variables are obtained from the same case. For case 7 Fig. 1 shows plots of the observed and estimated values of the discharge and the TP during the validation period. The graphs illustrate the good performance of SMAR and GOPC in reproducing the low values while there are some deficiencies in their capturing some of the peaks.

The ranking list for TP results is different from the flow one. At the top of this list is case 15 which used 4 and 3 fuzzy sub-sets for the rainfall and evaporation respectively. In contrast Case 6 with 2 rainfall fuzzy sub-sets and 2 evaporation fuzzy sub-sets is at the bottom. It is obvious that as many as 12 sub-models are required to adequately capture the variations in the processes involved in the TP simulation by case 15. The discharge and TP plots of case 15 are presented in Fig. 2. The hydrograph shapes do not much differ from those in Fig.1 for case 7. Likewise the TP graphs for cases 7 and 15 have similar shapes except for the largest two peaks which have been underestimated by the latter and overestimated by the former. The differences in simulation of the largest peaks explain the better  $R^2$  value for case 15 than the one for case 7.



**Figure 2.** Plots of discharge (Q) and TP during validation for case 15

## CONCLUSIONS

In this study, the SMAR model has been used with the GOPC to simulate the discharge and TP load at the outlet of a catchment in Northern Ireland. The rainfall and evaporation inputs have been represented in a fuzzy way. Different numbers of fuzzy sub-sets for both variables has been used in sixteen different combinations and for each combination different fuzzy logic rules have been formulated. The combination of the results of all rules in each case has been interpreted as a use of different sub-models representing various hydrological behaviours. The results showed that the best simulation for the discharge was achieved when using 2 fuzzy sub-sets for the rainfall and a single sub-set for the evaporation. For the TP, the use of 4 rainfall sub-sets along with 3 evaporation sub-sets was the best combination. The results are generally encouraging and they prove that the use of GOPC/SMAR along with fuzzy logic rules is worthwhile when good prediction of TP export has to be obtained and a fully distributed model, and the spatial database it requires, is not available.

## ACKNOWLEDGEMENTS

The work of this paper was done as part of a project funded by the Environmental Protection Agency in Ireland.



## REFERENCES

- Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R. (1998). Large area hydrological modelling and assessment Part I: Model development. *Journal of the American Water Resources Association (JAWRA)*, 34(1), 73-89.
- Bicknell, B.R., Imhoff, J.C., Donigan, A.S. and Johanson, R.C. (1997). Hydrological Simulation Program – FORTRAN (HSPF), User's Manual for release 11. EPA – 600/R-97/080. United States Environmental Protection Agency, Athens, GA.
- EEC (European Economic Community) (2000). The council of 23 October 2000 establishing a framework for Community action in the field of water policy (2000/60/EC). *Official Journal of the European Communities*, L327, 1-72.
- Ewen, J., Parkin, G. and O'Connell, E. (2000). SHETRAN: Distributed basin flow and transport modelling system. *Journal of Hydrologic Engineering*, 5(3), 250-258.
- Holland, J.H. (1975). *Adaptation in natural and artificial systems*. University Michigan Press, Ann Arbor.
- Jang, J.S.R. (1997). ANFIS architecture. In: *Neuro-fuzzy and Soft Computing*, J.S.R. Jang, C.-T. Sun and E. Mizutani (ed.), Prentice Hall, New Jersey.
- Khan, H. (1986). *Conceptual modelling of rainfall-runoff systems*. Master thesis, Department of Engineering Hydrology, National University of Ireland, Galway.
- Liang, G.C. (1992). *A note on the revised SMAR model*. Memorandum to the River Flow Forecasting Workshop Group, Department of Engineering Hydrology, National University of Ireland, Galway. (Unpublished).
- Mainstone, C.P. and Parr, W. (2002). Phosphorus in rivers – ecology and management. *The Science of the Total Environment*, 282-283, 25-47.
- Nash, J.E. and Sutcliffe, J.V. (1970). River flow forecasting through conceptual models. Part 1. A discussion of principles. *Journal Hydrology*, 10, 282-290.
- Nasr, A., Taskinein, A. and Bruen, M. (2005). Developing an independent, generic, phosphorus modelling component for use with grid-oriented, physically based distributed catchment models. *Water Science and Technology*, 51(3-4), 135-142.
- Nasr, A. (2004). *Modelling of phosphorus loss from land to water: a comparison of SWAT, HSPF and SHETRAN/GOPC for three Irish catchments*. PhD thesis, Department of Civil Engineering, University College Dublin, Ireland.
- Tan, B.Q. and O'Connor, K.M. (1996). Application of an empirical infiltration equation in the SMAR conceptual model. *Journal of Hydrology*, 185, 275-295.
- Viney, N.R., Sivapalan, M. and Deeley, D. (2000). A conceptual model of nutrient mobilisation and transport applicable at large catchment scales. *Journal of Hydrology*, 240, 23-44.
- Wade, A.J., Whitehead, P.G. and Butterfield, D. (2002). The integrated catchments model of phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogeneous river systems: model structure and equations. *Hydrology and Earth System Science*, 6(3), 583-606.