Environmental Science & Technology
Critical Review
pubs.acs.org/est

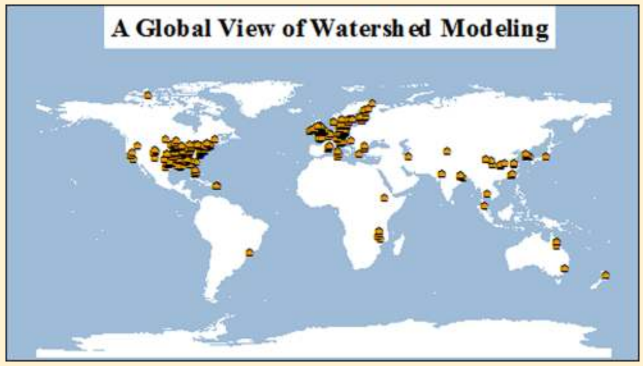
Evaluation of the Current State of Distributed Watershed Nutrient Water Quality Modeling

Christopher Wellen,*,† Ahmad-Reza Kamran-Disfani, and George B. Arhonditsis

Ecological Modeling Laboratory, Department of Physical & Environmental Sciences, University of Toronto, Toronto, Ontario Canada, M1C 1A4

Supporting Information

ABSTRACT: Watershed models have been widely used for creating the scientific basis for management decisions regarding nonpoint source pollution. In this study, we evaluated the current state of watershed scale, spatially distributed, process-based, water quality modeling of nutrient pollution. Beginning from 1992, the year when Beven and Binley published their seminal paper on uncertainty analysis in hydrological modeling, and ending in 2010, we selected 257 scientific publications which (i) employed spatially distributed modeling approaches at a watershed scale; (ii) provided predictions of flow, nutrient/sediment concentrations or loads; and (iii) reported fit to measured data. Most "best practices" (optimization, validation, sensitivity, and uncertainty analysis) are not consistently employed during model development. There are no statistically significant differences in model performance among land uses. Studies which used more than one point in space to evaluate their distributed models had significantly lower median values of the Nash-Sutcliffe Efficiency (0.70 vs 0.56, $p < 0.005$, nonparametric Mann−Whitney test), and r^2 ($p < 0.005$). This finding suggests that model calibration only to the basin outlet may mask compensation of positive and negative errors of source and transportation processes. We conclude by advocating a number of new directions for distributed watershed modeling, including in-depth uncertainty analysis and the use of additional information, not necessarily related to model end points, to constrain parameter estimation.



1. INTRODUCTION

Watershed models have been extensively used in hydrological science and environmental management research for a number of important tasks, including estimating nonpoint source pollutant inputs to receiving waterbodies and their source areas and predicting the effects of climate and land-use change on water quality.[1] Extensive research has focused on augmenting the mechanistic foundation of these watershed models and making them spatially distributed. Spatially distributed models disaggregate watersheds into multiple discrete units to represent the spatial variability of parameters and inputs.[2] However, the adequacy of earth science models for informing decision making has been questioned.[3,4] Concerns of overparameterization and equifinality have brought to the forefront of modeling efforts the development of methodologies that will obtain "the right answers for the right reasons".[5,6] Distributed, process-based models remain key tools for understanding and managing nonpoint source pollutants and the effects of land use change.[7,8,1,9] Models focused on nutrient pollution have a very long history of development and application for the purposes of management and policy, and form the focus of this paper.[2]

The documented inadequacy of many models to address important societal issues has frequently been attributed to the fact that the field has advanced without the healthy dose of introspection required to obtain good science.[3,4] For example, little work has quantitatively examined the practices of process-based watershed modeling. It is unknown to what extent "best practices" of model application are followed. While there are conventional recommendations of how "accurate" a model should be,[10] there is no sense of how well the existing class of distributed, process-based models performs across a variety of state variables, and how model development affects performance.

In this paper we quantitatively evaluate the state of distributed, process-based watershed models. We assess performance of a number of state variables associated with nutrient pollution and quantify how performance varies with model development. We also assess how often best model development practices are followed. This paper compliments more comparative, review-type approaches,[2,11] and aims to lead to concrete recommendations for the advancement of the field as a whole.

Received: October 10, 2014
Revised: February 14, 2015
Accepted: February 18, 2015
Published: February 18, 2015

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

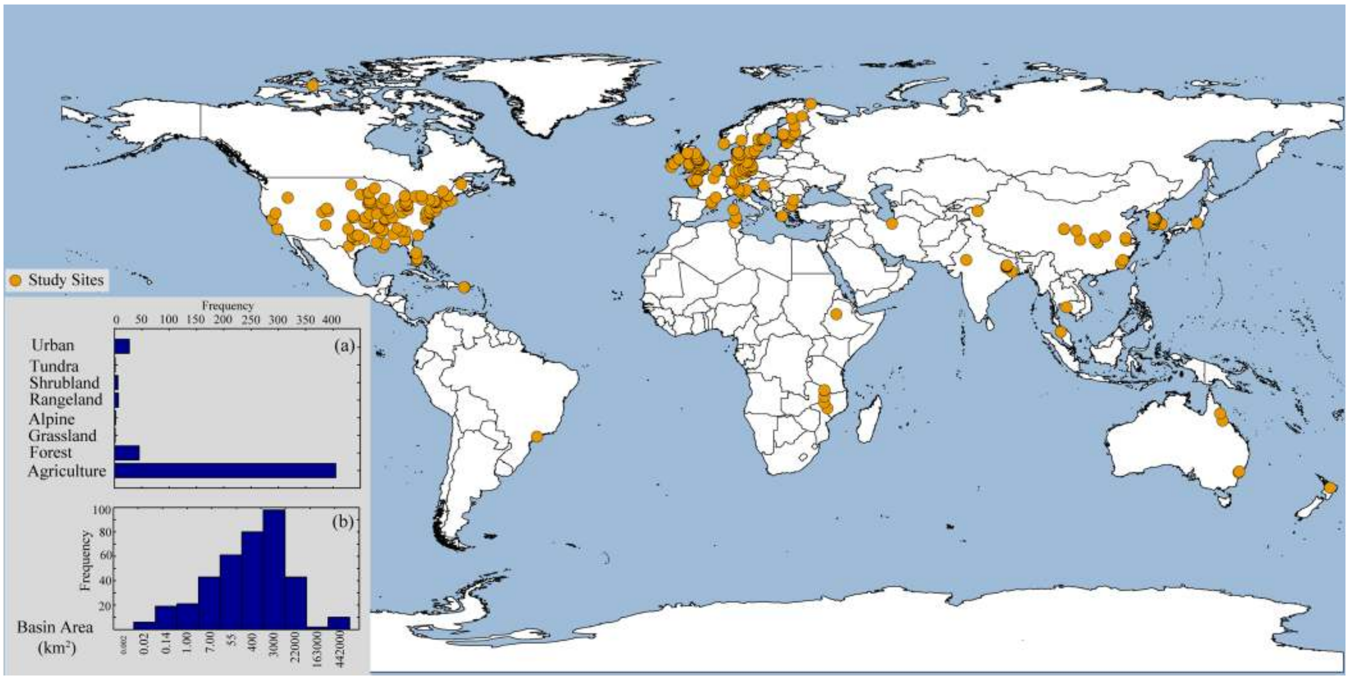


Figure 1. Map of model application watershed locations. Inset displays (a) bar graph of dominant landuse of each watershed, as identified by the authors of each study, and (b) area of model application watersheds in km^2 (note logarithmic *x*-axis).

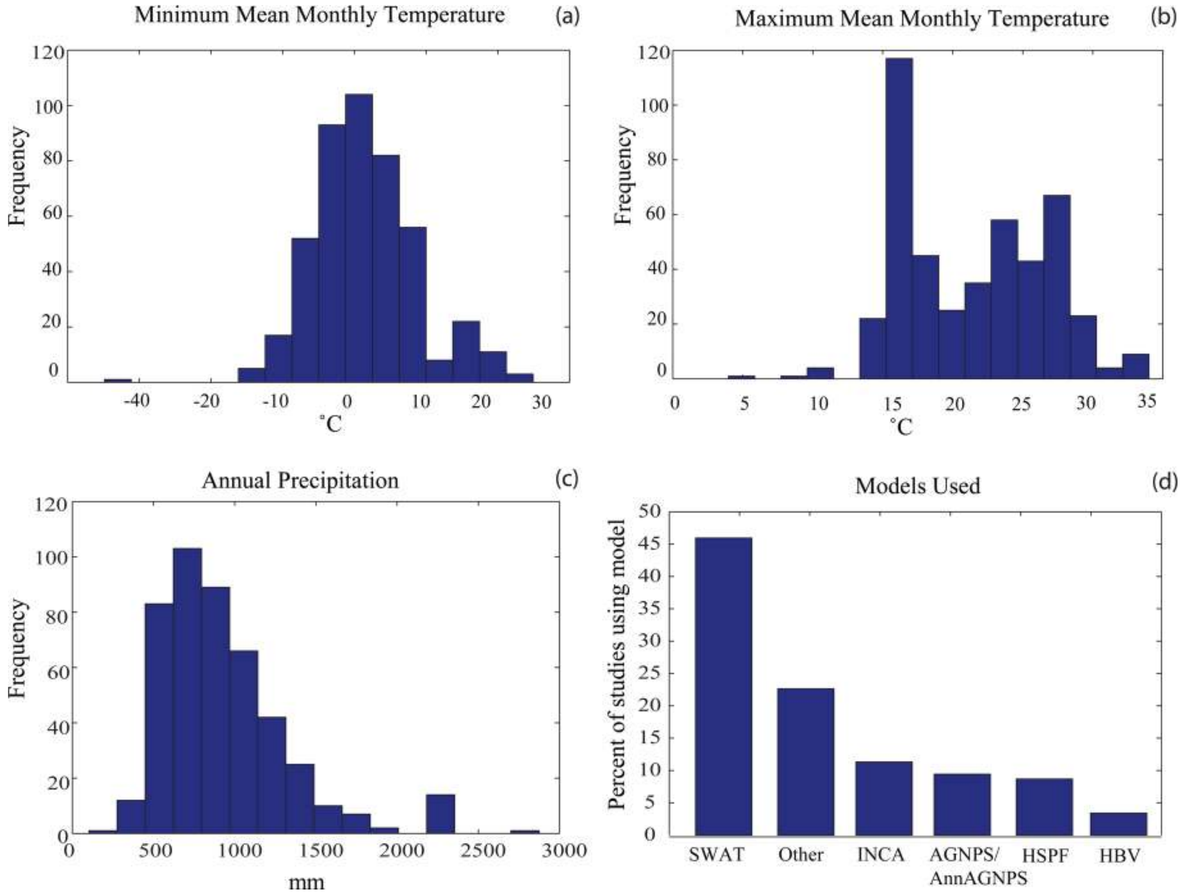


Figure 2. Histograms of (a) minimum mean monthly temperature; (b) maximum mean monthly temperature; (c); annual precipitation; and (d) the percentage of studies applying each model. Note that each study may apply more than one model.

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

2. MATERIALS AND METHODS

We sampled distributed watershed modeling work published in scientific journals between January of 1992 and July of 2010. The start of this period was the first appearance of the Generalized Likelihood Uncertainty Estimation (GLUE) methodology.[12] To be included, a publication had to (i) present watershed-scale predictions from a spatially distributed, process-based watershed model; (ii) make predictions of discharge as well as at least one nutrient or sediment concentration or load; and (iii) compare simulations to measurements from the studied system. Papers which presented sediment estimates as the sole water quality variable were included only if they used a model which could estimate nutrient concentrations or loads as well. To locate studies, we searched ISI's Web of Science database, using terms "watershed model(l)-ing", "model(l)-ing", "hydrological model-(l)-ing", plus one of: "nutrient(s)", "phosphorus", "phosphate", "nitrogen", "nitrate," or "ammonia". We found a total of 257 papers listed in the Supporting Information that fit these criteria.

We extracted metrics of fit for each simulated time series using two metrics: the coefficient of determination, $r^2 = \{\sum[(O-\bar{O}) \times (S-\bar{S})]\}^2/\{\sum(O-\bar{O})^2 \times \sum(S-\bar{S})^2)\}$ and the Nash-Sutcliffe Efficiency,[13] NSE= $(1 - \sum(O-S)^2/\sum(O-\bar{O})^2)$, where O refers to observations, S refers to simulations, $\bar{O}$ ($\bar{S}$) to the average of the observations (simulations). With both commonly used metrics, higher values indicate better fit and 1.0 indicates perfect fit. A NSE of 0 indicates a model which predicts as well as the average of the observations, and a negative NSE indicates a model which predicts more poorly than the average of the observations. The NSE penalizes for bias, where the r^2 does not penalize for linear bias. From each model application, we collected the metrics of fit, model name, method of spatial disaggregation, spatial and temporal resolution, length of time series, method of calibration, sensitivity and uncertainty analysis, basin size, land use types, latitude and longitude, the presence of point sources, and climate normals for 1980−2010.[14]

3. RESULTS

The 257 studies comprised 494 watersheds distributed globally, albeit with a preponderance of studies in the United States and Western Europe and some in China (20 watersheds, 16 studies), India (11 watersheds, 9 studies), and South Korea (9 watersheds, 8 studies; Figure 1). Notable areas with few or no studies in this database are Latin America, Northern Asia, and Africa. While there has been considerable distributed, process-based modeling in the Amazon basin, little of it focused on nutrients or sediment concentrations.[15] Only studies published in ISI indexed journals were selected for analysis. This disqualifies some studies in regional journals and all of the gray literature. Only watershed-scale studies which compare discharge and nutrients to observations are included. Most systems studied with nonpoint source models are agricultural systems, though there has been significant work in forested areas (Figure 1a). There is little published work in urban systems, despite the existence of nonpoint source process models specialized to urban areas (Storm Water Management Model).[16] More published work on this topic will allow us to assess the urban water quality modeling strategies currently in use. We found a roughly log-normal distribution of watershed sizes studied (Figure 1b). This range spans the research catchment scale through the mesoscale (10−10 000 km^2). Notable is the paucity of regional scale studies, only two (2) were based in a catchment larger than 35 000 km^2. While temperate, warm, and humid areas have been well studied with the models, we found cold and dry regions have not (Figure 2a−c). We note that five models comprise more than 80% of the literature we sampled (Figure 2d). This is in contrast with aquatic ecology, where models used are typically assembled for each study.[17,18]

3.1. Variables Simulated and Their Spatiotemporal Resolution.

The 257 studies presented 1873 simulated variables (Figure 3a). Discharge, sediment, and nitrate variables were simulated most commonly. Simulations of dissolved phosphorus species were presented by only 17% of studies. Few studies examined the individual flow paths to the stream. Only 4% of all studies presented simulations of surface runoff (overland flow), and only one study presented a simulation of groundwater flow.

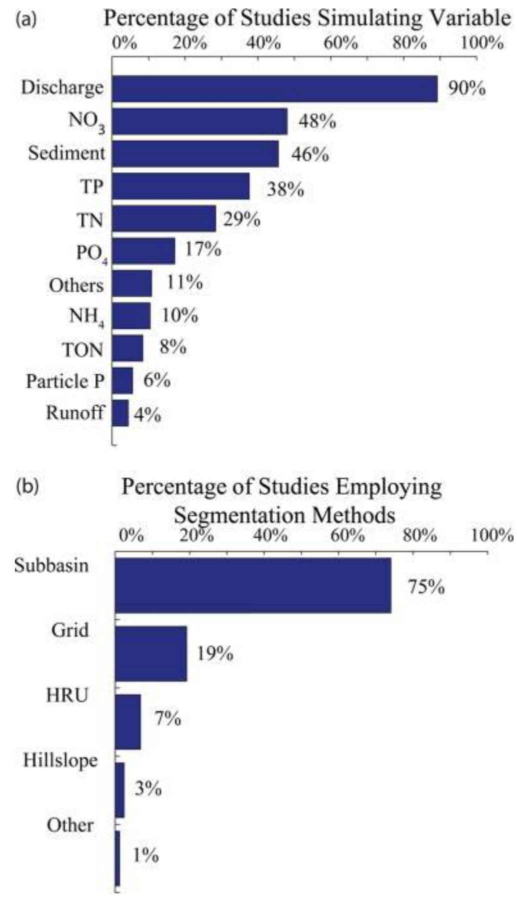


Figure 3. Bar graphs showing the percentage of studies which: (a) simulate each variable, and (b) employ particular spatial segmentation methods. The group "Others" refers to any model which was used in less than 5% of studies.

Distributed models discretize time and space. We found that 92% of studies employed a daily or monthly time step for evaluation and reporting (Supporting Information (SI) Figure S-1a). This implies that important biogeochemical or hydrological processes occurring on a time scale of hours (e.g., biological oxygen demand, first flush, snowmelt, changing source water contributions) may not be well represented in these models, even if their mathematical representation is adequate to characterize them at the required finer temporal resolution.

Detailed information in space may also be required to resolve important processes. For instance, denitrification requires the presence of nitrate and carbon and the absence of oxygen. Failure to sufficiently resolve these variables in space or time could lead

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

to a failure to simulate denitrification. By far the most common method of spatial segmentation is the subbasin, where topography is used to disaggregate discrete areas of drainage (Figure 3b). We found a log-normal distribution of the number of subbasins used, with a peak of roughly 20 subbasins (SI Figure S-1b). There was a significant but weak relationship between log transformed basin size and log transformed number of subbasins ($p < 0.01$, $r = 0.31$).

3.2. How Consistently Do Modelers Follow Best Practices? Environmental modeling textbooks typically present a sequence of best practices during model development. These include sensitivity analysis, optimization, validation, uncertainty analysis, and quantification of fit.[19] Sensitivity analysis is an assessment of how much model output varies, given a specific level of variability of the parameters and other inputs. By pinpointing parameters and inputs which exert a strong influence on model outputs, sensitivity analysis gives a sense of accuracy and precision requirements. Nevertheless, only 25% of studies reported any results of sensitivity analysis (Figure 4).

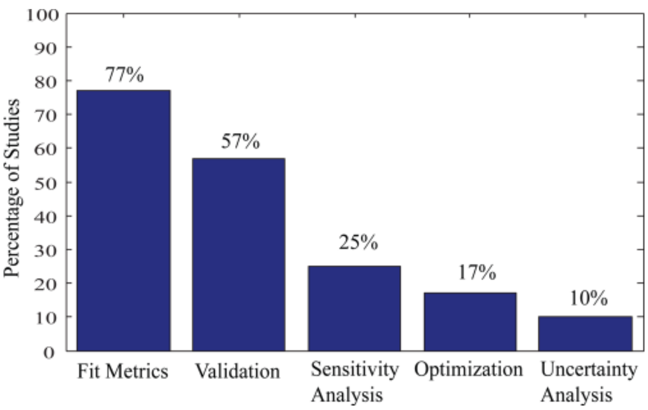


Figure 4. Percentage of studies which quantify any metrics of fit, perform a validation, perform a sensitivity analysis, perform some kind of objective optimization procedure, and perform an uncertainty analysis. We here define uncertainty analysis as any attempt to present a range of predictions based on the uncertainty of parameters, inputs, or model structure. Optimization is here defined as any attempt to objectively locate a "best fit" solution, that is, any method more advanced than manual parameter adjustment.

Considerable work has been focused on the need for uncertainty analysis in watershed modeling.[20] Uncertainty in watershed models partly stems from the use of simplified, abstract mathematics to simulate real world processes. The information we use to force, parameterize, and calibrate these models is also subject to significant uncertainty.[21−24,1] Because deterministic model applications ignore this uncertainty, many researchers have found that deterministic applications may not be particularly meaningful.[1,4,12,17,18,20−24] Tremendous effort has gone into developing uncertainty analysis frameworks in watershed modeling.[25,21,26,22,27] However, only 10% of distributed, process-based, nonpoint source pollutant modeling studies attempted to account for the uncertainty of their model predictions in a quantitative manner (Figure 4); a somewhat surprising result, given that the starting point of our study period is the publication of the Beven and Binley seminal paper.[2]

The meaning of validation has been debated extensively in the literature.[3] Rykiel[28] emphasizes that validation does not imply the model is "true", or even optimal—only that it is acceptable for a given purpose. Power[29] defines as the simplest form of validation the quantification of goodness of fit. Seventy-seven percent (77%) of watershed modeling studies in our data set quantified goodness of fit. Predictive validation refers to the ability of a model to fit data to which it was not calibrated, and is the most common validation practice. Validation was performed by the majority of studies in our database (57%, Figure 4). Structural validation refers to assessing the realism of one or more components of the model (e.g., causal relationships, relative magnitudes of fluxes).[17] Only six studies explicitly performed structural validation on any variables: corn and soy yields;[30,31] soil nutrient concentrations;[32] TN:TP ratios;[33] evapotranspiration;[34] and water table heights.[34,35]

When calibration is done manually, it is unclear if any lack of fit is due to a poor parameter choice or inadequate model structure. Optimization is any automated, objective method of selecting a parameter vector (e.g., genetic algorithms). Despite the tremendous amount of research effort invested into developing optimization techniques in the watershed modeling literature,[36] only 17% of studies reported the use of any kind of optimization technique (Figure 4).

3.3. How Well Do Process Based, Nonpoint Source Watershed Models Simulate the Real World? We present box plots of the NSE values in Figure 5 and include boxplots for the r^2 values in the SI as Figure S2. We also present performance percentiles in tabular form in SI Table S-1. The hydrometric variables tended to be simulated more accurately than the water quality variables. The median values of all of the variables presented in Figure 5a are respectable. However, the 25th percentile for some water quality variables is quite low. Negative NSE values indicate that the model predictions perform more poorly than the observed mean value. Of the 257 studies, 40 (16%) studies published at least one simulated variable with NSE < 0. There was little variability of performance across the different dominant land uses in the data set (Figure 5b). Performance did vary across models, with Agricultural NonPoint Source pollution model in both event and continuous modes (AGNPS and AnnAGNPS) being characterized by better performance, and the Integrated Catchment Model (INCA) and Hydrologiska Byråns Vattenbalansavdelning (HBV) being characterized by worse performance (Figure 5c). It should be noted that the INCA and HBV communities calibrate and assesses their models at a large number (10 or more) of discharge and water quality nodes within the basin. As we substantiate in Section 4.2, this degree of rigor would likely result in lower metrics of fit for the other models in Figure 5. Finally, model error did not vary across time step (Figure 5d).

3.4. How Does Model Development Influence Performance? Table 1 presents the correlation coefficients and p-values of relationships between performance metrics and various aspects of study design. There was no consistent relationship between the number of subelements or the subelement size and the model performance—models which have more detail in space were not on the whole more accurate than those which had less detail in space. We found either no significant relationship or a very weak positive relationship. We did find some weak relationships between various environmental covariates and performance. Generally, models performed slightly better as minimum and maximum temperatures got warmer, conditions got wetter, and elevation increased. We noted a positive relationship between fit and the number of study citations. This finding contrasts the trends reported for aquatic biogeochemical modeling,[37] although we stress that the relationship is weak.

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

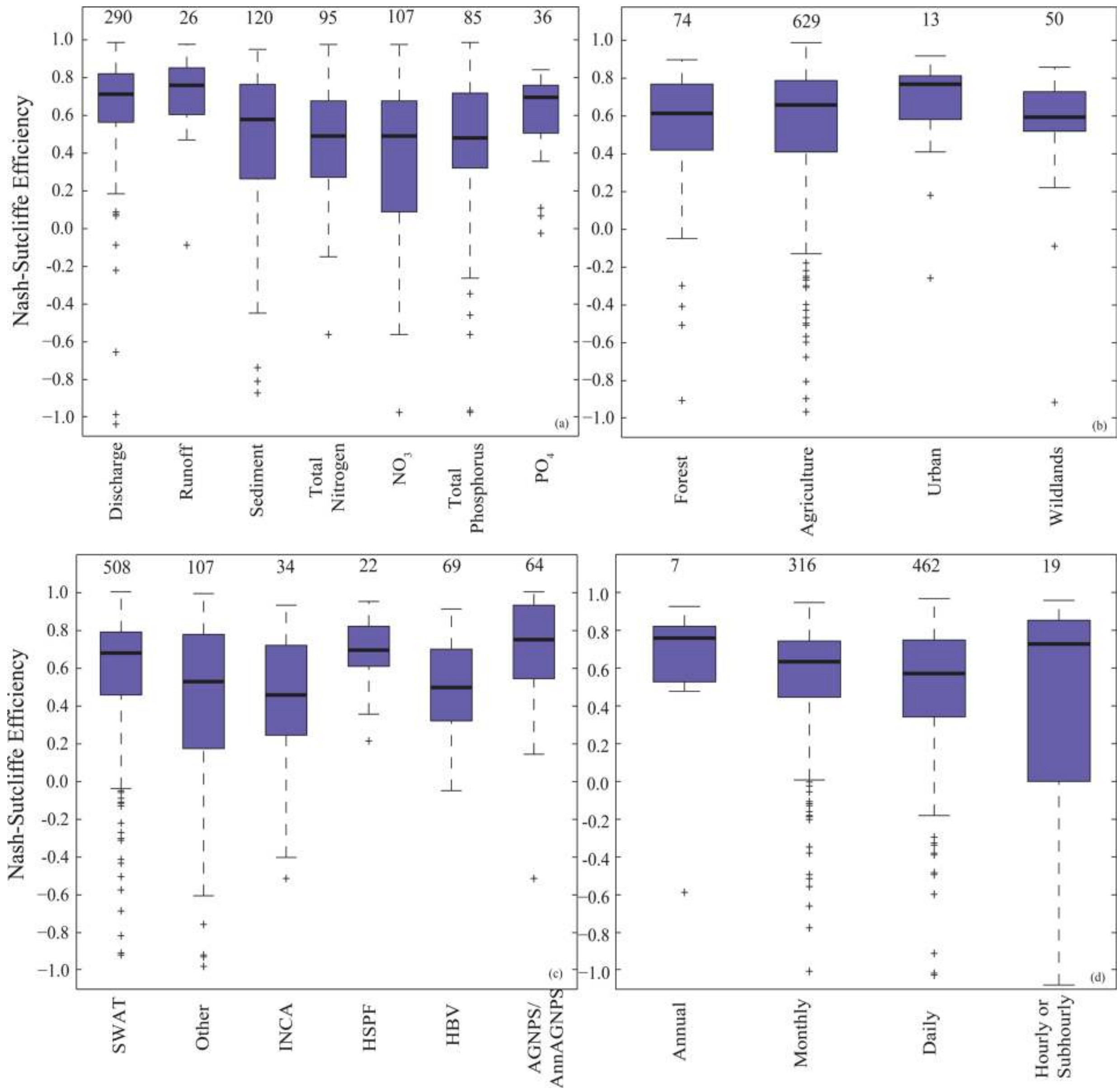


Figure 5. Box plots of Nash-Sutcliffe Efficiency for (a) selected variables; (b) dominant landuse types; (c) models used; and (d) time steps employed. Numbers above each box indicate number of samples in each group. Variables in (a) were selected to have at least 26 samples. Wildlands in (b) refers to any land other than Forest which is not dominated by human land uses, for example, grassland. Panels b–d include all variables.

4. DISCUSSION

In this paper, we sampled the process-based, distributed watershed modeling literature and assessed the state of the practice. We found that this class of models is applied globally, across a wide range of climatic and land cover conditions, albeit with a preponderance of studies in North American and European agricultural areas. Daily and monthly time steps for reporting and comparing predictions are the norm (96% of papers), as are subbasin spatial delineations (75% of papers). Model validation is common, though far from universal (57% of studies; Figure 4). However, despite the considerable amount of research effort into optimization and uncertainty analysis, only 10% of the studies we sampled attempted uncertainty analysis, and optimization is only presented in 17% of studies (Figure 4). We found a tremendous range in the performance of watershed models, with respectable median values of the Nash-Sutcliffe Efficiency (0.4–0.7), but many poor simulations reported (25th percentiles between −0.36 and 0.65; SI Table S-1). None of the study aspects reported in Table 1 were able to explain this variability. It seems that much of the variability of model fit results from aspects which are difficult to extract from papers, for example, diligence in calibration, fitness of the model for the studied system, the characteristics of the calibration data set at hand (e.g., relative representation of the baseline versus event-based conditions, sampling frequency), or the quality of the inputs (quantity and timing of nutrients applied).

4.1. A Comparison of the Most Common Watershed Models. While 41 models appeared in our database, only five (5) models appeared in more than 3% of the studies in the database: Soil–Water Assessment Tool (SWAT),[38] INCA,[7,8] Agricultural Nonpoint Source Pollution Model/Annual Agricultural Nonpoint Source Pollution Model (AGNPS/AnnAGNPS),[39] Hydrological Simulation Program-Fortran (HSPF),[40] and HBV[41]

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278–3290

Table 1. Results of Regressions between Fit Metrics and Covariates[a]

covariate	metric	*p*-value	*r*
spatial covariates			
subelement average size (km^2)	NSE	0.52	
	r^2	0.21	
number of subelements	NSE	0.11	
	r^2	<0.01	0.11
basin area (km^2)	NSE	0.34	
	r^2	0.51	
catchment area/DEM cell size	NSE	0.18	
temporal Covariates	r^2	0.42	
length (days)	NSE	<0.01	0.14
	r^2	0.12	
length (steps)	NSE	0.21	
	r^2	0.18	
environmental covariates			
minimum mean monthly temperature (°C)	NSE	<0.01	0.14
	r^2	<0.01	0.11
maximum mean monthly temperature (°C)	NSE	0.04	0.08
	r^2	<0.01	0.17
precipitation (mm)	NSE	<0.01	0.1
	r^2	<0.01	0.14
elevation (MASL)	NSE	<0.01	0.1
	r^2	<0.01	0.15
bibliographic covariates			
number of citations	NSE	0.03	0.1
	r^2	<0.01	0.13
year of publication	NSE	<0.01	0.13
	r^2	0.25	

[a]Fit (r^2) and slopes of regressions are presented only when the *p*-values were less than 0.05. Regressions were calculated for all state variables. Values of NSE less than −1 were omitted.

(Figure 2d). We synthesized some of the key elements of these models in Tables 2−5.

Most of the models in Table 2 posit that their fundamental calculation units are connected directly to a stream. This precludes incorporating any explicit representation of upland topology. Topology is pivotal to the location of biogeochemical hot spots, or locations where chemically complementary flowpaths meet.[42] Topology is also key to understanding the "fill and spill" mechanisms operating in many northern catchments.[43] While a depiction of upland topology may be unnecessary in large catchments, many of the catchments where these models have been applied are on the order of size that upland topology could be of great importance (<1 km^2; Figure 1b). Incorporating the upland-riparian topology could be a critical feature for improving the representation of biogeochemistry into these models. Some progress in this regard has already been made with SWAT. Easton et al.[35] used topographic wetness indices as the basis for their HRUs instead of landuse and soil type. This resulted in each hillslope being divided into a cascade of HRUs, and improved model fit for dissolved phosphorus.[35,44] Other work[45,46] has shown that the HRUs for SWAT can be delineated according to landscape position along a catena, or series of downslope cascading flows. This spatial framework yielded increased accuracy and decreased bias when applied to sediment yield.[47] More focus on how models represent topology may lead to improvements in catchment scale biogeochemical simulations.

The five models in Table 3 have similar conceptual models of the upland origins and routing of streamflow. Flow to the stream consists mainly of a surface component and a shallow soil component. SWAT, INCA, HSPF, and HBV include a groundwater component, whereas AnnAGNPS and HBV include a tile drainage component. There was considerable diversity of calculation procedures employed, although they all tended to be empirical, especially in the case of surface runoff. These empirical approaches allow models to be applied to large areas without extremely detailed surface data. However, they can make it difficult to determine whether these models assume that surface runoff is typically generated using infiltration excess (e.g., Hortonian) or saturation excess mechanisms,[48] as runoff increases with stored soil moisture. The models in Table 3 may not be able to locate runoff prone areas, especially in catchments where these areas expand and contract seasonally. We note that for these models, surface runoff is the main flowpath by which phosphorus, sediment, and many other important pollutants are deposited into the stream.

The models in Table 4 varied in terms of their approach to modeling overland sediment inputs. Some models use the empirical Universal Soil Loss Equation or a similar approach (SWAT, AGNPS/AnnAGNPS, and HBV). HSPF and INCA are somewhat physically based, and use calibrated relationships between overland flow volume and velocity and transport capacity, and both account for splash and sheet erosion.[49] The in-stream sediment components tended to rely on Bagnold-type models, where transport capacity is estimated as a function of flow or peak flow, and this capacity results in sediment being suspended if the load is below capacity or deposited if it is above.[50] Separate size classes are often addressed in the in-stream routing components, with each size class being characterized with a critical shear stress. Transport can only occur when this critical shear stress is exceeded. With the exception of HBV, these

Table 2. Representation of Space by Common Watershed Models[a]

	SWAT	INCA	AGNPS/AnnAGNPS	HSPF	HBV
primary disaggregation unit	subbasins	subbasins	irregular "cells" of uniform land management and soil	subbasins	subbasins
secondary disaggregation unit	HRUs on the basis of landuse, soil, and slope	landuse		pervious and impervious landuse	elevation zones
tertiary disaggregation unit		1 km^2 pixels			land use zones

[a]These five models together accounted for 83% of the studies in our sample.

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

Table 3. Representation of Streamflow Generation by Common Watershed Models

	SWAT	INCA	AGNPS/AnnAGNPS	HSPF	HBV
fluxes to stream	surface runoff, lateral flow (return flow), groundwater flow	surface runoff (INCA-P, INCA-Sed), lateral flow (return flow), groundwater flow	surface runoff, lateral flow (return flow), flow through subsurface drainage	surface runoff, lateral flow (return flow), groundwater flow	surface runoff, lateral subsurface flow, groundwater flow, flow through subsurface drainage
surface runoff	daily curve number, peak predicted by rational method.	surface flow coefficient, threshold in lateral subsurface flow.	daily curve number, TR-55 method for peak flow.	empirical outflow depth to detention storage relation, Manning equation for runoff velocity.	daily curve number
lateral subsurface flow	kinematic storage	landuse specific time constants.	Darcy's law	empirical relations	two tank linear reservoir: upper tank
groundwater flow	empirical relations	landuse specific time constants and a baseflow coefficient.	NA	empirical relations	two tank linear reservoir: lower tank
subsurface drainage	NA	NA	Hooghoudt's equation	NA	two tank linear reservoir: upper tank threshold bypass
channel flow	variable storage or Muskingum with Manning's equation. Adjustments for evaporation, diversions, and transmission.	reach time constant, empirical relation between volume and outflow.	Manning's equation numerically solved, assume trapezoidal channel.	reach outflow a function of reach volume or user-supplied demand, flow velocity with Manning equation.	equilateral triangular weighting function distributes contributions from tanks over time.

Table 4. Representation of Sediment Transport by Common Watershed Models

	SWAT	INCA	AGNPS/AnnAGNPS	HSPF	HBV
overland	MUSLE expressed in terms of runoff volume, peak flow, and USLE factors	empirical maximum sediment export rate function of overland flow. erosion from splash erosion and sheet erosion, actual precipitation drives splash erosion, includes cover factor. sheet erosion is a function of direct runoff	RUSLE describes delivery of sediment to end of hillslope, HUSLE describes delivery of that sediment to the stream, deposition is based on size distribution and particle fall velocity	transport capacity power relation of water storage and outflow. Rainfall splash detachment and wash off of the detached sediment based on transport capacity, scour estimated from overland flow using power relation with water storage and flow	empirical function of surface runoff and rainfall/snowmelt
in-stream	empirically adjusted bagnold with deposition based on particle fall velocity, resuspension allowed	Bangnold stream power, separate size class accounting	Bagnold equation determines sediment capacity using shear stress, calculated from water depth and slope, transport calculated for five different particle size classes.	noncohesive sediment using user-defined relation with flow velocity or Toffaleti or Colby method, cohesive sediment with critical shear stress and settling velocity	sediment bound P is deposited or resuspended with empirical functions of discharge rate and concentration, banks contribute sediment bound P with an empirical function of discharge and morphometry.

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

Table 5. Representation of Soil Nutrient Cycling by Common Watershed Models

	SWAT	INCA	AGNPS/AnnAGNPS	HSPF	HBV
soil nitrogen pools	soil nitrate, soil ammonium, active organic N, stable organic N, plant residue	soil nitrate, soil ammonium, groundwater nitrate, groundwater ammonium, organic nitrogen	active inorganic N, stable inorganic N, active organic N, stable organic N,	soil nitrate, solution ammonium, adsorbed ammonium, plant N above/below ground, litter N, particualte labile organic N, solution labile organic N, particulate refractory organic N, solution refractory organic N	organic N, inorganic N
soil nitrogen fluxes	plant uptake, denitrification, volatilization, nitrification, decay, residue mineralization	plant uptake, denitrification, nitrification, ammmonia mineraliazaration (decay), ammonia immobilization	plant uptake, denitrification, volatilization, nitrification, decay, residue mineralization, immobilization, leaching	plant uptake, denitrification, mineralization, immobilization, litterfall, plant N return, sorption	total nitrate retention, total organic N production
soil nitrogen kinetics	first order kinetics, rate constants depend on temperature and water availability, decay rate constants also depend on supply of nutrients	temperature dependent rate coefficients (base coefficients input)	temperature and water dependent, first order kinetics	temperature dependent, first order kinetics	temperature and N concentration dependent kinetics
soil phosphorus pools	stable mineral P, active mineral P, solution mineral P, stable organic P, active organic P, and fresh organic P	firmly bound inorganic P, inorganic P, organic P, firmly bound organic P	active mineral P, stable mineral P, solution mineral P; humic organic P, fresh organic P (residue)	plant P, adsorbed P, solution P, and organic P	dissolved phosphorus, particulate sephosphorus
soil phosphorus fluxes	plant uptake, decay, mineralization, mobilization, immobilization between active and stable pools and between solution and active pools	plant uptake, decay, mineralization, 2-way transfers between firmly bound and labile pools	plant uptake, decay, mineralization to solution, adsorption/occlusion	plant uptake, mineralization, immobilization, adsorption, desorption	long-term or daily average concentrations of dissolved and particulate P in hydrological fluxes estimated from ICECREAM model
soil phosphorus kinetics	first order kinetics, rate constants depend on temperature and water availability, decay rate constants also depend on supply of nutrients, decay process integrated across N and P cycles	temperature dependent rate coefficients (base coefficients input)	temperature and water dependent, first order kinetics	temperature dependent, first order kinetics	ICECREAM kinetics functions of soil physics and chemistry (pH, base saturation, clay content)

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

relatively robust in-stream algorithms render these models able to address questions of stream bank versus upland sources of sediment, an important emerging concern.[51]

The nitrogen and phosphorus components of the models in Table 5 had virtually the same structure. There are a small number of conceptual pools of organic and mineral forms of each nutrient, and the transformations between each pool are governed by first-order kinetics. Often the base reaction rates were included as calibration parameters, with modifications dependent on temperature. These model structures differ strikingly from models focused on biogeochemistry, for instance DNDC.[52] DNDC integrates the nitrogen and carbon cycles. For instance, denitrification can only take place when both carbon and nitrate are present. The isolation of the biogeochemical cycles from each other by the watershed models in Table 5 is another serious impediment to using them to accommodate "hot spot" phenomena, where chemically distinct waters mix and react.[42] The simplified conceptual models in Table 5 also neglect the different fates various forms of occluded phosphorus may have. Iron and Manganese bound phosphorus can become released during anoxic conditions in streambank sediments, whereas calcium bound phosphorus tends to be more stable.[53,54]

4.2. What Additional Information Should Be Incorporated into Distributed, Process Models? Only 19% of studies in our database calibrated to more than one location in space. While virtually all of the studies calibrated to streamflow, only 4% calibrated to surface runoff, and no studies calibrated or assessed the ability of their models to simulate any other hydrological fluxes. Researchers have cautioned that reliance on information from only one location in space and only one hydrological flux makes it possible that poor simulations of one flux can be compensated for by poor simulations of another flux of the opposite sign.[4,6] The danger of this fit through error compensation instead of faithful depiction of basin dynamics is heightened by the highly empirical nature of many model components. We recommend that future studies consider using three additional sources of information to further constrain their predictions and get the right answers for the right reasons: additional information in space, information on additional model fluxes, and tracer information to help constrain model sources.

Constraining simulated in-stream fluxes of water and water-borne constituents may be improved by incorporating information from a greater variety of locations in space. Research focused on inorganic nitrogen with the Hydrological Predictions for the Environment (HYPE) model suggests that calibrating a distributed, process-based model to multiple stations within a nested basin context can reduce the uncertainty of the water quality predictions and also improve the accuracy at the upstream stations.[55] Work with HBV has reached a similar conclusion.[56] To test the hypothesis that nested basin approaches can generally improve an assessment of model accuracy and uncertainty, we took the metrics of fit and categorized each into one of two categories. The first category consisted of studies where the calibration was to a nested basin, and the second to studies where the model was constrained only at the basin outlet. We conducted a nonparametric Mann−Whitney U test on the two groups and rejected the null hypothesis of no difference ($p <$ 0.005). Note that we did separate tests for values of the NSE and the r^2. For each of the metrics of fit, the median was higher when calibrating only to the basin outlet (0.70 vs 0.56 for NSE, and 0.78 vs 0.63 for r^2). Calibrating only to the basin outlet likely results in an overconfident assessment of a distributed model's ability to reproduce the internal dynamics of the basin, including source attributions and land use scenarios. INCA and HBV are usually calibrated in a nested basin context, which at least partly explains their lower performance (Figure 3). In this regard, we highlight the importance of using information from multiple locations in space to constrain model predictions. Increasing information in space could better support spatially variable parameters. Future work should seek to develop and apply frameworks which use additional information in space to allow the model parameters to vary spatially. Bayesian hierarchical frameworks are one possible approach for doing so. Bayesian hierarchical frameworks use global hyperparameters to share information across sites, while allowing parameter values some degree of site-specificity.[57]

Future work should better take advantage of the multiple criteria which watershed models can be calibrated to. The most common approach to calibrate multiple variables has traditionally been to start with hydrology, then proceed to sediment and then nutrients.[58] Yet, studies which have examined this practice find that it results in suboptimal results when compared with approaches which calibrate flow and water quality all at once.[59,56] The models included in our database simulate all the major fluxes of the hydrological cycle (e.g., evapotranspiration, groundwater flow, overland flow, return flow), yet the common practice is to calibrate only to streamflow. This implies that we do not know how realistically these models reproduce the hydrological cycle. By incorporating additional hydrological fluxes such as evapotranspiration or tile drain flow into the model evaluation, it might be possible to arrive at more credible estimates of the other hydrological fluxes. This may result in more credible estimates of pollutant export.

Model calibration can be aided by incorporating empirical information about the sources of water, sediment, and nutrients. HYPE,[34] HBV,[41] and WATFLOOD[60] explicitly account for water isotope mixing. Approaches to incorporate tracer-derived information into model calibration that do not require alterations to existing structures should be developed. There are a variety of techniques developed for drawing inferences from tracer data, including end member mixing analysis,[61] sediment fingerprinting,[51] and isotope analysis of some dissolved nutrients.[62] Source attributions estimated from these techniques could be used to constrain model predictions by calibrating the model's summary statistics to statistics of source estimates.[63] Yen et al.[64] recently used annual rates of denitrification in addition to discharge and nitrate concentration for calibration and found that doing so improved the realism of the scenario analysis. Other approaches such as approximate Bayesian computation (ABC)[65] and the Generalized Likelihood Uncertainty Estimation (GLUE)[66,67] approach could allow the incorporation of empirical source attributions with model-based estimates of source areas.

4.3. The Importance of Best Practice in Modeling. An important result from this study is that performing sensitivity analysis, uncertainty analysis, and optimization are not the norm. Similar results have been found regarding the state of aquatic biogeochemical modeling.[17,18,37] This is despite these "best practices" being described in some detail in most modeling textbooks.[19,68,69] Significant improvement to the contemporary modeling practice can be achieved simply by making standard the practice of reporting on the results of validation, sensitivity analysis, and uncertainty analysis.

There were a number of papers in our database which were focused on conducting sensitivity analyses or uncertainty analyses, mainly with the SWAT model. The importance of

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

sensitivity and uncertainty analysis are underscored by the following findings of these studies: (i) much of the sensitivity and the uncertainty of the SWAT model may stem from just one, two, or three parameters;[70,71] (ii) the SWAT model is sensitive to the spatial resolution selected, with performance increasing with spatial detail as subbasins and HRUs are coarse and remaining relatively constant after an appropriate amount of detail is found;[72] (iii) the SWAT model is sensitive to the density of rain gauges used to force it;[73] and (iv) the most sensitive parameters of the SWAT model vary with application site, and may not even be the same at water quality stations upstream of the basin outlet.[70,74,75] Many of these findings are likely due to the complex and spatially distributed nature of the SWAT model, characteristics shared with the other models in our sample. Sensitivity analysis allows modelers to defend their choice of calibration vector and should be standard reporting for distributed model results.

It was somewhat surprising to see how little uncertainty analysis had been conducted. Considerable work has been undertaken by the hydrological modeling community to develop uncertainty analysis techniques,[12,76] yet they have rarely been applied to the water quality predictions of distributed, process-based watershed scale water quality models, likely due to the models' complexity (but see Yen et al.[64,77]). One of the challenges of the contemporary modeling practice is the development of calibration techniques that can effectively accommodate the uncertainties related to the behavior of watersheds during extreme events, given that the frequency of such events is expected to increase if the current urbanization and climate change trends continue. Our work on this subject involved a novel Bayesian hierarchical framework postulating two distinct states with respect to the watershed response to precipitation; that is, precipitation depth above a certain threshold triggers an extreme state, characterized by a qualitatively different response of the watershed to precipitation. The integration of this calibration framework with the SWAT model offered reasonable end of basin fit of discharge and sediment load (NSE $\sim$ 0.7) and state-specific parameters coherently identified.[78,79] Yet, we found that the 95% credible intervals of urban and agricultural sediment export overlapped.[79] The widespread adoption of uncertainty analysis techniques to model end points (e.g., in-stream phosphate concentrations) and source apportionments (e.g., mass of phosphate exported from croplands) will help make the predictions of complex over-parameterized (but necessary) models more credible to decision makers.

5. CONCLUSIONS

We assessed the state of the art of spatially distributed, process-based watershed models with a sample of 257 papers published between 1992 and 2010. While the median performance was respectable, there was a very wide range, and performance declined as we moved from water quantity components to water quality components. The distributed watershed water quality modeling community does not consistently adhere to best practices. Doing so would constitute a methodological advancement which is well within our reach. We recommend that an examination of model best practices (error metric calculation, validation, sensitivity analysis, optimization, uncertainty analysis, and assessment at more than one station) become a typical part of the review of papers using mathematical models. While not every paper needs to employ every best practice, the onus should be on authors to explain which best practices were not employed and what the corresponding effects on their results might be.

Performance did not significantly covary with degree of spatial or temporal detail of the models employed. However, evaluating distributed models with information from more than one water quantity or quality station significantly negatively impacted their assessed performance. This suggests that the common practice of assessing a distributed model only at the basin outlet gives an overconfident assessment of its ability to reproduce within-basin dynamics.

The field of process-based, distributed watershed modeling is dominated by five models: SWAT, INCA, AGNPS/AnnAGNPS, HSPF, and, HBV, which together constitute roughly 83% of the studies in our data set. These models have similar representations of spatial variability (uplands all connect directly to streams), relevant flow paths (surface water, shallow subsurface water), and nutrient biogeochemistry (a small number of pools with reaction rates treated as calibration parameters; no interaction between nutrient cycles). These model structures pose some difficulty to accommodating contemporary ideas of observational hydrology and biogeochemistry, especially the "hot-spot" and "fill-and-spill" concepts. While these models may be appropriate for mesoscale analysis, we found that they are often used at fairly fine scales (catchments less than 10 km^2). Future work should examine how appropriate these models are at evaluating the effects of best management practices, given that these practices typically work at a field scale. While the sources of parameter sensitivity have been explored, the sources of predictive uncertainty are still relatively unknown with the class of models in this paper. This is true of model end points and even more so for internal basin dynamics, such as source attributions. Elucidating the sources and magnitudes of uncertainty in these models would constitute a key advancement in their use, and would make them more suitable decision support tools. On a final note, we believe that the publication of several recent meta-analysis/critique papers in the context of earth science modeling[17,18,20,37] is a sign of maturation of the field, as they offer the healthy dose of self-criticism and restless mindset required to address the demand for attractive and powerful management tools.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information contains a list of all 257 studies, graphs of spatiotemporal resolution (Figure S-1), boxplots of the coefficient of determination (Figure S-2), box plots comparing concentrations to loads (Figure S-3), and percentiles of metrics of fit (Table S-1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +1 647 239 5138; e-mail: wellenc@mcmaster.ca.

Present Address

†(C.W.) Watershed Hydrology Group, Department of Geography and Earth Sciences, McMaster University. Hamilton, Ontario, Canada, L8S 4L8.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Support for this study was provided by the Natural Sciences and Engineering Council of Canada through a Discovery Grant

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

awarded to George Arhonditsis. Christopher Wellen received support from Ontario Graduate Scholarships.

■ REFERENCES

(1) Rode, M.; Arhonditsis, G.; Balin, D.; Kebede, T.; Krysanova, V.; van Griensven, A.; van der Zee, S.E.A.T.M. New challenges in integrated water quality modeling. *Hydrol. Process.* **2010**, *24*, 3447−3461 DOI: 10.1002/hyp.7766.

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

(3) Oreskes, N.; Shrader-Frechette, K.; Belitz, K. Verification, validation, and confirmation of numerical models in the earth sciences. *Science* **1994**, *263* (5147), 641−646.

(4) Beven, K. A manifesto for the equifinality thesis. *J. Hydrol.* **2006**, *320* (1−2), 18−36 DOI: 10.1016/j.jhydrol.2005.07.007.

(5) Seibert, J.; McDonnell, J. J. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resour. Res.* **2002**, *38* (11), 1241 DOI: 10.1029/2001WR000978.

(6) Kirchner, J. W. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* **2006**, *42*, W03S04 DOI: 10.1029/2005WR004362.

(7) Wade, A. J.; Durand, P.; Beaujouan, V.; Wessel, W. W.; Raat, K. J.; Whitehead, P. G.; Butterfield, D.; Rankinen, K.; Lepisto, A. A nitrogen model for European catchments: INCA, new model structure and equations. *Hydrol. Earth Syst. Sci.* **2002**, *6* (3), 559−582 DOI: 10.5194/hess-6-559-2002.

(8) Wade, A. J.; Whitehead, P. G.; Butterfield, D. The Integrated Catchments model of phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogenous river systems: Model structure and equations. *Hydrol. Earth Syst. Sci.* **2002**, *6* (3), 583−606.

(9) Michalak, A. M.; Anderson, E. J.; Beletsky, D.; Boland, S.; Bosch, N. S.; Bridgeman, T. B.; Chaffin, J. D.; Cho, K.; Confesor, R.; Daloglu, I.; DePinto, J. V.; Evans, M. A.; Fahnenstiel, G. L.; He, L.; Ho, J. C.; Jenkins, L.; Johengen, T. H.; Kuo, K. C.; LaPorte, E.; Liu, X.; McWilliams, M. R.; Moore, M. R.; Posselt, D. J.; Richards, R. P.; Scavia, D.; Steiner, A. L.; Verhamme, E.; Wright, D. M.; Zagorski, M. A. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (16), 6448−6452 2013. DOI: 10.1073/pnas.1216006110.

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

(11) Gassman, P.; Reyes, M.; Green, C.; Arnold, J. The soil and water assessment tool: Historical development, applications, and future research directions. *Trans. ASABE* **2007**, *50* (4), 1211−1250.

(12) Beven, K.; Binley, A. The future of distributed models: Model calibration and uncertainty prediction. *Hydrol. Process.* **1992**, *6* (3), 279−298.

(13) Nash, J. E.; Sutcliff, J. V. River flow forecasting through conceptual models. Part 1—A discussion of principles. *J. Hydrol.* **1970**, *10* (1970), 282−290.

(14) *GHCN Global Climate*; EarthInfo: Huntington Beach, California, 2011.

(15) Strauch, M.; Bernhofer, C.; Koide, S.; Volk, M.; Lorz, C.; Makeschin, F. Using precipitation data ensemble for uncertainty analysis in SWAT streamflow simulation. *J. Hydrol.* **2012**, *414*−*415*, 413−424 DOI: 10.1016/j.jhydrol.2011.11.014.

(16) Rossman, L. Storm Water Management Model User's Manual, Version 5.0., 2010. http://nepis.epa.gov/Adobe/PDF/P100ERK4.pdf (accessed August 19, 2014).

(17) Arhonditsis, G. B.; Brett, M. T. Evaluation of the current state of mechanistic aquatic biogeochemical modeling. *Mar. Ecol-Prog. Ser.* **2004**, *271*, 13−26.

(18) Robson, B. State of the art in modeling of phosphorus in aqutic systems: Review, criticisms, and commentary. *Environ. Model. & Softw.*, in press. http://dx.doi.org/10.1016/j.envsoft.2014.01.012.

(19) Chapra, S. C. *Surface Water Quality Modeling*; Waveland Press, 1997.

(20) Pappenberger, F.; Beven, K. J. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resour. Res.* **2006**, *42*, W05302 DOI: 10.1029/2005wr004820.

(21) Wagener, T.; Gupta, H. V. Model identification for hydrological forecasting under uncertainty. *Stochem. Environ. Res. and Risk Assess.* **2005**, *19*, 378−387 DOI: 10.1007/s00477-005-0006-5.

(22) Ajami, N. K.; Duan, Q. Y.; Sorooshian, S. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour. Res.* **2007**, *43* (1), W01403 DOI: 10.1029/2005wr004745.

(23) Arhonditsis, G. B.; Papantou, D.; Zhang, W.; Perhar, G.; Massos, E.; Shi, M. Bayesian calibration of mechanistic aquatic biogeochemical models and benefits for environmental management. *J. Mar. Syst.* **2008**, *73*, 8−30 DOI: 10.1016/j.jmarsys.2007.07.004.

(24) Arhonditsis, G. B.; Perhar, G.; Zhang, W.; Massos, E.; Shi, M.; Das, A. Addressing equifinality and uncertainty in eutrophication models. *Water Resour. Res.* **2008**, *44*, W01420 DOI: 10.1029/2007WR005862.

(25) Engeland, K.; Gottschalk, L. Bayesian estimation of parameters in a regional hydrological model. *HESS* **2002**, *6* (5), 883−898.

(26) Huard, D.; Mailhot, A. A Bayesian perspective on input uncertainty in model calibration: Application to hydrological model "abc". *Water Resour. Res.* **2006**, *42* (7), W07416 DOI: 10.1029/2005WR004661.

(27) Liu, Y.; Gupta, H. V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* **2007**, *43* (7), W07401 DOI: 10.1029/2006wr005756.

(28) Rykiel, E. J. Testing ecological models: The meaning of validation. *Ecol. Model.* **1996**, *90* (3), 229−244.

(29) Power, M., The predictive validation of ecological and environmental models. *Ecol. Model. 199368*, 33−50.

(30) Hu, X.; McIsaac, G. F.; David, M. B.; Louwers, C. A. L. Modeling riverine nitrate export from and East-Central Illinois watershed using SWAT. *J. Environ. Qual.* **2007**, *36*, 996−1005.

(31) Donner, S. D.; Kucharic, C. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. *Global Biogeochem. Cycles* **2003**, *17* (3), 1085.

(32) Bastrup-Birk, A.; Gundersen, P. Water quality improvements from afforestation in an agricultural catchment in Denmark illustrated with the INCA model. *HESS* **2004**, *8* (4), 764−777.

(33) Green, M. B.; Wang, D. Watershed flow paths and stream water nitrogen-to-phosphorus ratios under simulated precipitation regimes. *Water Resour. Res.* **2008**, *44*, W12414 DOI: 10.1029/2007WR006139.

(34) Lindström, G.; Pers, C.; Rosberg, J.; Strömqvist, J.; Arheimer, B. Development of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrol. Res.* **2010**, *41*, 295.

(35) Easton, Z. M.; Fuka, D. R.; Walter, M. T.; Cowan, D. M.; Schneiderman, E. M.; Steenhuis, T. S. Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. *J. Hydrol.* **2008**, *348* (3), 279−291.

(36) Duan, Q. *Global Optimization for Watershed Model Calibration, in Calibration of Watershed Models*; Duan, Q., Gupta, H. V., Sorooshian, S., Rousseau, A. N. , Turcotte, R., Eds.; American Geophysical Union: Washington, D.C., 2013; DOI: 10.1002/9781118665671.ch6.

(37) Arhonditsis, G. B.; Adams-VanHarn, B. A.; Nielsen, L.; Stow, C. A.; Reckhow, K. H. Evaluation of the current state of mechanistic aquatic biogeochemical modeling: Citation analysis and future perspectives. *Environ. Sci. Technol.* **2006**, *40*, 6547−6554.

(38) Neitsch, S. L.; Arnold, J. G.; Kiniry, J. R.; Williams, J. R. *Soil and Water Assessment Tool Theoretical Documentation*, Version 2009; Texas Water Resources Institute Technical Report No. 406. Texas A&M

University System: College Station, TX, 2011; http://twri.tamu.edu/reports/2011/tr406.pdf (accessed January 17, 2013).

(39) Binger, R. L.; Theurer, F. D.; and Yuan, Y. AnnAGNPS Technical Processes. Version 5.2, 2011. http://www.ars.usda.gov/research/docs.htm?docid=5222 (accessed August 20, 2014).

(40) Bicknell, B. R.; Imhoff, J. C., Kittle, Jr., J. L., Donigian, Jr., A. S., and Johanson, R. C. *Hydrologic Simulation Program - FORTRAN (HSPF): User's Manual for Release 11*; U.S. EPA Environmental Research Lab: Athens, GA, 1996.

(41) Andersson, L.; Rosberg, J.; Pers, B. C.; Olsson, J.; Arheimer, B. Estimating catchment nutrient flow with the HBV-NP model: Sensitivity to input data. *Ambio* **2005**, *34*, 521−532.

(42) McClain, M. E.; Boyer, E. W.; Dent, C. L.; Gergel, S. E.; Grimm, N. B.; Groffman, P. M.; Hart, S. C.; Harvey, J. W.; Johnston, C. A.; Mayorga, E.; McDowell, W. H.; Pinay, G. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems* **2003**, *6* (4), 301−312.

(43) Oswald, C. J.; Richardson, M. C.; Branfireun, B. A. Water storage dynamics and runoff response of a boreal Shield headwater catchment. *Hydrol. Process.* **2011**, *25*, 3042−3060 DOI: 10.1002/hyp.803.

(44) Collick, A. S.; Fuka, D. R.; Kleinman, P. J. A.; Buda, A. R.; Weld, J. L.; White, M. J.; Veith, L.; Bryant, R. B.; Bolster, C. H.; Easton, Z. M. Predicting phosphorus dynamics in complex terrains using a variable source area hydrology model. *Hydrol. Process.* **2015**, *29*, 588−601 DOI: 10.1002/hyp.10178.

(45) Arnold, J. G.; Allen, P. M.; Volk, M.; Williams, J. R.; Bosch, D. D. Assessment of different representations of spatial variability on SWAT model performance. *Trans. ASABE* **2010**, *53* (5), 1433−1443.

(46) Bosch, D. D.; Arnonld, J. G.; Volk, M.; Allen, P. M. Simulation of a low-gradient Coastal Plain watershed using the SWAT landscape model. *Trans. ASABE* **2010**, *53* (5), 1445−1456.

(47) Bonuma, N.; Rossi, C.; Arnold, J.; Reichert, J.; Minella, J.; Allen, P.; Volk, M. Simulating landscape sediment transport capacity by using a modified SWAT model. *J. Environ. Qual.* **2014**, *43* (1), 55−66.

(48) Dunne, T.; Black, R. D. Partial area contributions to storm runoff in a small New England watershed. *Water Resour. Res.* **1970**, *6* (5), 1296−1311.

(49) Vaze, J.; Chiew, F. H. Study of pollutant washoff from small impervious experimental plots. *Water Resour. Res.* **2003**, *39* (6), 1160.

(50) Bagnold, R. A. Bedload transport in natural rivers. *Water Resour. Res.* **1977**, *13* (2), 303−312.

(51) Davis, C.; Fox, J. Sediment fingerprinting: Review of the method and future improvements for allocating nonpoint source pollution. *J. Environ. Eng.* **2009**, *135* (7), 490−504 http://dx.doi.org/10.1061/(ASCE)0733-9372(2009)135:7(490).

(52) Giltrap, D. L.; Li, C. and Saggar, S. DNDC: A process-based model of greenhouse gas fluxes from agricultural soils. *Agric., Ecosyst. Environ.* **2010**, *136*, 292−230.

(53) Carlyle, G. C.; Hill, A. R. Groundwater phosphate dynamics in a river riparian zone: Effects of hydrologic flowpaths, lithology, and redox chemistry. *J. Hydrol.* **2001**, *247*, 151−168.

(54) Nierdermer, A.; Robinson, J. S. Hydrological controls on soil redox dynamics in a peat-based, restored wetland. *Geoderma* **2007**, *137*, 318−326.

(55) Jiang, S.; Jomaa, S.; Rode, M. Modeling inorganic nitrogen leaching in nested mesoscale catchments in central Germany. *Ecohydrology* **2014**, DOI: 10.1002/eco.1462.

(56) Pettersson, a; Arheimer, B.; Johansson, B. Nitrogen concentrations simulated with HBV-N: New response function and calibration strategy. *Nord. Hydrol.* **2001**, *32*, 227−248.

(57) Zhang, W.; Arhonditsis, G. B. A Bayesian hierarchical framework for calibrating aquatic biogeochemical models. *Ecol. Model.* **2009**, *220*, 2142−2161.

(58) Santhi, C.; Arnold, J. G.; Williams, J. R.; Dugas, W. A.; Srinivasan, R.; Hauck, L. M. Validation of the swat model on a large river basin with point and nonpoint sources. *J. Am. Water Resour. Assoc.* **2001**, *37*, 1169−1188 DOI: 10.1111/j.1752-1688.2001.tb03630.x.

(59) Bergstrom, S.; Lindstrom, G.; Pettersson, A. Multi-variable parameter estimation to increase confidence in hydrological modeling. *Hydrol. Process.* **2002**, *16*, 413−421.

(60) Stadnyk, T. A.; Delavau, C.; Kouwen, N.; Edwards, T. W. D. Towards hydrological model calibration and validation: Simulation of stable water isotopes using the isoWATFLOOD model. *Hydrol. Process.* **2013**, *27*, 3791−3810 DOI: 10.1002/hyp.9695.

(61) Burns, D.; McDonnell, J.; Hooper, R.; Peters, N.; Freer, J.; Kendall, C.; Beven, K. Quantifying contributions to storm runoff through end-member mixing analysis and hydrologic measurements at the Panola Mountain Research Watershed (Georgia, USA). *Hydrol. Process.* **2001**, *15*, 1903−1924 DOI: 10.1002/hyp.246.

(62) McLaughlin, K.; Kendall, C.; Silva, S.; Young, M.; Paytan, A. Phosphate oxygen isotope ratios as a tracer for sources and cycling of phosphate in North San Francisco Bay, California. *J. Geophys. Res.* **2006**, *111*, G03003 DOI: 10.1029/2005JG000079.

(63) Csillery, K.; Blum, M.; Gaggiotti, O.; Francois, O. Approximate Bayesian computation (ABC) in practice. *Trend. Ecol. Evol.* **2010**, *25* (7), 410−418 DOI: 10.1016/j.tree.2010.04.001.

(64) Yen, H.; Bailey, R.; Arabi, M.; Ahmadi, M.; White, M.; Arnold, J. The role of interior watershed processes in improving parameter estimation and performance of watershed models. *J. Environ. Qual.* **2014**, *43*, 1601−1613.

(65) Vrugt, J. A.; Sadegh, M. Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resour. Res.* **2013**, *49*, 4335−4345 DOI: 10.1002/wrcr.20354.

(66) Liu, Y.; Freer, J.; Beven, K. J.; Matgen, P. Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error. *J. Hydrol.* **2009**, *367* (1−2), 93−103.

(67) Blazkova, S.; Beven, K. A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. *Water Resour. Res.* **2009**, *45*, W00B16 DOI: 10.1029/2007WR006726.

(68) Beven, K. *Rainfall-Runoff Modeling: The Primer*, 2nd ed.; Wiley-Blackwell, 2012; ISBN: 978-0-470-71459-1 .

(69) Jorgensen, S. E. *Fundamentals of Ecological Modeling*; Elsevier Science: Amsterdam, 1994.

(70) Cibin, R.; Sudheer, K. P.; Chaubey, I. Sensitivity and identifiability of stream flow generation parameters of the SWAT model. *Hydrol. Process.* **2010**, *24*, 1133−1148 DOI: 10.1002/hyp.7568.

(71) Wu, Y.; Liu, S. Automating calibration, sensitivity and uncertainty analysis of complex models using the R package Flexible Modeling Environment (FME): SWAT as an example. *Environ. Modell. Softw.* **2012**, *31*, 99−109 DOI: 10.1016/j.envsoft.2011.11.013.

(72) Chaplot, V. Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO_3-N loads predictions. *J. Hydrol.* **2005**, *312*, 207−222.

(73) Chaplot, V.; Saleh, A.; Jaynes, D. B. Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3−N loads at the watershed level. *J. Hydrol.* **2005**, *312*, 223−234.

(74) van Griensven, A.; Meixner, T.; Grunwald, S.; Bishop, T.; Diluzio, M.; Srinivasan, R. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J. Hydrol* **2006**, *324* (1−4), 10−23 DOI: 10.1016/j.jhydrol.2005.09.008.

(75) Nossent, J.; Bauwens, W. Multi-variable sensitivity and identifiability analysis for a complex environmental model in view of integrated water quantity and water quality modeling. *Water Sci. Technol.* **2012**, *65.3*, 539−549 DOI: 10.2166/wst.2012.884.

(76) Yang, J.; Reichert, P.; Abbaspour, K. C.; Xia, J.; Yang, H. Comparing uncertainty analysis techniques for a SWAT application to the Chaohe Basin in China. *J. Hydrol.* **2008**, *358*, 1−23 DOI: 10.1016/j.jhydrol.2008.05.012.

(77) Yen, H.; Wang, X.; Fontane, D.; Harmel, D.; Arabi, M. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. *J. Environ. Modell. Software* **2014**, *54*, 211−221.

(78) Wellen, C. C.; Arhonditsis, G. B.; Long, T.; Boyd, D. Accommodating environmental thresholds and extreme events in

hydrological models: A Bayesian approach. *J. Great Lakes Res.* **2014**, DOI: 10.1016/j.jglr.2014.04.002.
(79) Wellen, C. C.; Arhonditsis, G. B.; Long, T.; Boyd, D. Quantifying the uncertainty of nonpoint source attribution in distributed water quality models: A Bayesian assessment of SWAT's sediment export predictions. *J. Hydrol.* **2014**, DOI: 10.1016/j.jhydrol.2014.10.007.

DOI: 10.1021/es5049557
Environ. Sci. Technol. 2015, 49, 3278−3290

EVALUATION OF THE CURRENT STATE OF DISTRIBUTED WATERSHED NUTRIENT WATER QUALITY MODELING

Environmental Science and Technology

Manuscript Number: es5049557

SUPPORTING INFORMATION

Christopher Wellen,* Ahmad Kamran-Disfani, George B. Arhonditsis

Ecological Modelling Laboratory,

Department of Physical & Environmental Sciences, University of Toronto,

Toronto, Ontario, Canada, M1C 1A4

* Corresponding author

Current address: Watershed Hydrology Group, Department of Geography and Earth Sciences, McMaster University. Hamilton, Ontario, Canada, L8S 4L8

e-mail: wellenc@mcmaster.ca, Tel.: +1 647 239 5138

The 257 studies included in our analysis are listed below:

1. Abbaspour, K. C. *et al.* Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *J. Hydrol.* **333,** 413–430 (2007).

2. Adeuya, R. K., Lim, K. J., Engel, B. A. & Thomas, M. A. Modeling the average annual nutrient losses of two watersheds in Indiana using GLEAMS-NAPRA. *Trans. ASABE* **48,** 1739–1749 (2005).

3. Ancev, T., Stoecker, A. L., Storm, D. E. & White, M. J. The economics of efficient phosphorus abatement in a watershed. *J. Agric. Resour. Econ.* **31,** 529–548 (2006).

4. Andersson, L., Rosberg, J., Pers, B. C., Olsson, J. & Arheimer, B. Estimating catchment nutrient flow with the HBV-NP model: sensitivity to input data. *Ambio* **34,** 521–532 (2005).

5. Arabi, M., Govindaraju, R. S. & Hantush, M. M. A probabilistic approach for analysis of uncertainty in the evaluation of watershed management practices. *J. Hydrol.* **333,** 459–471 (2007).

6. Arabi, M., Govindaraju, R. S., Hantush, M. M. & Engel, B. A. Role of watershed subdivision on modeling the effectiveness of best management practices with SWAT. *J. Am. Water Resour. Assoc.* **42,** 513–528 (2006).

7. Arheimer, B. & Brandt, M. Watershed modelling of nonpoint nitrogen losses from arable land to the Swedish coast in 1985 and 1994. *Ecol. Eng.* **14,** 389–404 (2000).

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

9. Babel, M. S., Najim, M. M. M. & Loof, R. Assessment of Agricultural NonPoint Source Model for a Watershed in Tropical Environment. *J. Environ. Eng.* **130,** 1032–1041 (2004).

10. Baffaut, C. & Benson, V. W. Modeling flow and pollutant transport in a karst watershed with SWAT. *Trans. ASABE* **52,** 469–479 (2009).

11. Baginska, B., Milne-Home, W. & Cornish, P. S. Modelling nutrient transport in Currency Creek, NSW with AnnAGNPS and PEST. *Environ. Model. Softw.* **18,** 801–808 (2003).

12. Band, L. E., Tague, C. L., Groffman, P. & Belt, K. Forest ecosystem processes at the watershed scale: Hydrological and ecological controls of nitrogen export. *Hydrol. Process.* **15,** 2013–2028 (2001).

13. Bärlund, I., Kirkkala, T., Malve, O. & Kämäri, J. Assessing SWAT model performance in the evaluation of management actions for the implementation of the Water Framework Directive in a Finnish catchment. *Environ. Model. Softw.* **22,** 719–724 (2007).

14. Bärlund, I. *et al.* Three approaches to estimate inorganic nitrogen loading under varying climatic conditions from a headwater catchment in Finland. *Hydrol. Res.* **40,** 167 (2009).

15. Bärlund, I. & Kirkkala, T. Examining a model and assessing its performance in describing nutrient and sediment transport dynamics in a catchment in southwestern Finland. *Boreal Environ. Res.* **13,** 195–207 (2008).

16. Bastrup-Birk, A. & Gundersen, P. Water quality improvements from afforestation in an agricultural catchment in Denmark illustrated with the INCA model. *Hydrol. Earth Syst. Sci.* **8,** 764–777 (2004).

17. Behera, S. & Panda, R. K. Evaluation of management alternatives for an agricultural watershed in a sub-humid subtropical region using a physical process based model. *Agric. Ecosyst. Environ.* **113,** 62–72 (2006).

18. Bekele, E. G. & Nicklow, J. W. Multi-objective automatic calibration of SWAT using NSGA-II. *J. Hydrol.* **341,** 165–176 (2007).

19. Bekele, E. G. & Nicklow, J. W. Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms. *Water Resour. Res.* **41,** 1–10 (2005).

20. Benaman, J., Shoemaker, C. A. & Haith, D. A. Calibration and Validation of Soil and Water Assessment Tool on an Agricultural Watershed in Upstate New York. *J. Hydrol. Eng.* **10,** 363–374 (2005).

21. Bergman, M. J., Green, W. & Donnangelo, L. J. Calibration of Storm Loads in the South Prong Watershed, Florida, Using Basins/Hspf. *J. Am. Water Resour. Assoc.* **38,** 1423–1436 (2002).

22. Bergström, S., Lindström, G. & Pettersson, A. Multi-variable parameter estimation to increase confidence in hydrological modelling. *Hydrol. Process.* **16,** 413–421 (2002).

23. Bernal, S., Butturini, A., Riera, J. L., Vázquez, E. & Sabater, F. Calibration of the INCA model in a Mediterranean forested catchment: the effect of hydrological inter-annual variability in an intermittent stream. *Hydrol. Earth Syst. Sci.* **8,** 729–741 (2004).

24. Bhuyan, S. J., Koelliker, J. K., Marzen, L. J. & Harrington Jr, J. A. An integrated approach for water quality assessment of a Kansas watershed. *Environ. Model. Softw.* **18,** 473–484 (2003).

25. Boorman, D. B. Climate, Hydrochemistry and Economics of Surface-water Systems (CHESS): Adding a European dimension to the catchment modelling experience developed under LOIS. *Sci. Total Environ.* **314-316,** 411–437 (2003).

26. Borah, D. K., Demissie, M. & Keefer, L. L. AGNPS-based Assessment of the Impact of BMPs on Nitrate-Nitrogen Discharging into an Illinois Water Supply Lake. *Water Int.* **27,** 255–265 (2002).

27. Bosch, N. S. The influence of impoundments on riverine nutrient transport: An evaluation using the Soil and Water Assessment Tool. *J. Hydrol.* **355,** 131–147 (2008).

28. Bouraoui, F., Galbiati, L. & Bidoglio, G. Climate change impacts on nutrient loads in the Yorkshire Ouse catchment (UK). *Hydrol. Earth Syst. Sci.* **6,** 197–209 (2002).

29. Bouraoui, F., Grizzetti, B., Granlund, K., Rekolainen, S. & Bidoglio, G. Impact of climate change on the water cycle and nutrient losses in a Finnish catchment. *Clim. Change* **66,** 109–126 (2004).

30. Bouraoui, F., Benabdallah, S., Jrad, A. & Bidoglio, G. Application of the SWAT model on the Medjerda river basin (Tunisia). *Phys. Chem. Earth* **30,** 497–507 (2005).

31. Bouraoui, F. & Dillaha, T. A. ANSWERS-2000: non-point-source nutrient planning model. *J. Environ. Eng.* **126,** 1045 – 1055 (2000).

32. Bouraoui, F. & Grizzetti, B. An integrated modelling framework to estimate the fate of nutrients: Application to the Loire (France). *Ecol. Modell.* **212,** 450–459 (2008).

33. Bracmort, K. S., Arabi, M., Frankenberger, J. R., Engel, B. a & Arnold, J. G. Modeling Long-Term Water Quality Impact of Structural BMPs. *Trans. Am. Soc. Agric. Biol. Eng.* **49,** 367–374 (2006).

34. Cau, P. & Paniconi, C. Assessment of alternative land management practices using hydrological simulation and a decision support tool: Arborea agricultural region, Sardinia. *Hydrol. Earth Syst. Sci.* **11,** 1811 – 1823 (2007).

35. Cerucci, M. & Conrad, J. M. The use of binary optimization and hydrologic models to form riparian buffers. *J. Am. Water Resour. Assoc.* **39,** 1167–1180 (2003).

36. Chaplot, V. Water and soil resources response to rising levels of atmospheric CO2 concentration and to changes in precipitation and air temperature. *J. Hydrol.* **337,** 159–171 (2007).

37. Chaplot, V. Impact of DEM mesh size and soil map scale on SWAT runoff, sediment, and NO3-N loads predictions. *J. Hydrol.* **312,** 207–222 (2005).

38. Chaplot, V., Saleh, A. & Jaynes, D. B. Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO3-N loads at the watershed level. *J. Hydrol.* **312,** 223–234 (2005).

39. Chaplot, V., Saleh, A., Jaynes, D. B. & Arnold, J. Predicting water, sediment and NO3-N loads under scenarios of land-use and management practices in a flat watershed. *Water. Air. Soil Pollut.* **154,** 271–293 (2004).

40. Chen, E. & Mackay, D. S. Effects of distribution-based parameter aggregation on a spatially distributed agricultural nonpoint source pollution model. *J. Hydrol.* **295,** 211–224 (2004).

41. Chen, N. *et al.* Assessment of management practices in a small agricultural watershed in Southeast China. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* **41,** 1257–1269 (2006).

42. Cheng, H., Ouyang, W., Hao, F., Ren, X. & Yang, S. The non-point source pollution in livestock-breeding areas of the Heihe River basin in Yellow River. *Stoch. Environ. Res. Risk Assess.* **21,** 213–221 (2007).

43. Cho, J., Bosch, D., Lowrance, R., Strickland, T. & Vellidis, G. Effect of spatial distribution of rainfall on temporal and spatial uncertainty of SWAT output. *Trans. ASABE* **52,** 1545–1555 (2009).

44. Cho, J. *et al.* Effect of watershed subdivision and filter width on swat simulation of a coastal plain watershed. *J. Am. Water Resour. Assoc.* **46,** 586–602 (2010).

45. Choi, K. S. & Blood, E. Modeling developed coastal watersheds with the agricultural non-point source model. **35,** 233–244 (1999).

46. Chu, T. W., Shirmohammadi, A., Montas, H. & Sadeghi, A. Evaluation of the SWAT model's sediment and nutrient components in the Piedmont physiographic region of Maryland. *Trans. ASAE* **47,** 1523−1538 (2004).

47. Cohen, M. J. & Brown, M. T. A model examining hierarchical wetland networks for watershed stormwater management. *Ecol. Modell.* **201,** 179–193 (2007).

48. Conan, C., Bouraoui, F., Turpin, N., de Marsily, G. & Bidoglio, G. Modeling flow and nitrate fate at catchment scale in Brittany (France). *J. Environ. Qual.* **32,** 2026–2032

49. Cooper, A. B. & Bottcher, A. B. Basin Scale Modeling as Tool for Water Resource Planning. *J. Water Resour. Plan. Manag.* **119,** 306–323 (1993).

50. Cotter, A., Chaubey, I., Costello, T., Soerens, T. & Nelson, M. Water quality model output uncertainty as affected by spatial resolution of input data. *J. Am. Water Resour. Assoc.* **39,** 977–986 (2003).

51. Cruise, J. F. & Miller, R. L. D. A.-1993. Hydrologic modeling with remotely sensed databases. *Water Resour. Bull.* **29,** 997–1002 (1993).

52. Cruise, J. F. & Miller, R. L. Hydrologic modeling of land processes in Puerto Rico using remotely sensed data. *Water Resour. Bull.* **30,** 419 – 428 (1995).

53. Dalzell, B. J., Gowda, P. H. & Mulla, D. J. Modeling sediment and phosphorus losses in an agricultural watershed to meet TMDLs. *J. Am. Water Resour. Assoc.* **40,** 533–543 (2004).

54. David, M. B. *et al.* Modeling denitrification in a tile-drained, corn and soybean agroecosystem of Illinois, USA. *Biogeochemistry* **93,** 7–30 (2009).

55. Dean, S., Freer, J., Beven, K., Wade, A. J. & Butterfield, D. Uncertainty assessment of a process-based integrated catchment model of phosphorus. *Stoch. Environ. Res. Risk Assess.* **23,** 991–1010 (2009).

56. Deslandes, J., Beaudin, I., Michaud, A., Bonn, F. & Madramootoo, C. A. Influence of Landscape and Cropping System on Phosphorus Mobility within the Pike River Watershed of Southwestern Quebec: Model Parameterization and Validation. *Can. Water Resour. J.* **32,** 21–42 (2007).

57. Di Luzio, M. & Lenzi, M. A. Surface runoff, soil erosion and water quality modelling in the Alpone watershed using AGNPS integrated with a Geographic Information System. *Eur. J. Agron.* **6,** 1–14 (1997).

58. Di Luzio, M., Arnold, J. G. & Srinivasan, R. Effect of GIS data quality on small watershed stream flow and sediment simulations. *Hydrol. Process.* **19,** 629–650 (2005).

59. Di Luzio, M., Srinivasan, R. & Arnold, J. G. Integration of Watershed Tools and Swat Model Into Basins. *JAWRA J. Am. Water Resour. Assoc.* **38,** 1127–1141 (2002).

60. Dietrich, J. & Funke, M. Integrated catchment modelling within a strategic planning and decision making process: Werra case study. *Phys. Chem. Earth* **34,** 580–588 (2009).

61. Donner, S. D. & Kucharik, C. J. Evaluating the impacts of land management and climate variability on crop production and nitrate export across the Upper Mississippi Basin. *Global Biogeochem. Cycles* **17,** 1085 (2003).

62. Donner, S. D., Kucharik, C. J. & Foley, J. A. Impact of changing land use practices on nitrate export by the Mississippi River. *Global Biogeochem. Cycles* **18,** 1–21 (2004).

63. Du, B., Saleh, A., Jaynes, D. & Arnold, J. Evaluation of SWAT in simulating nitrate nitrogen and atrazine fates in a watershed with tiles and potholes. *Trans. ASAE* **49,** 949–960 (2006).

64. Durand, P. Simulating nitrogen budgets in complex farming systems using INCA: calibration and scenario analyses for the Kervidy catchment (W. France). *Hydrol. Earth Syst. Sci.* **8,** 793–802 (2004).

65. Easton, Z. M. *et al.* Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas. *J. Hydrol.* **348,** 279–291 (2008).

66. Easton, Z. M., Gérard-Marchant, P., Walter, M. T., Petrovic, A. M. & Steenhuis, T. S. Identifying dissolved phosphorus source areas and predicting transport from an urban watershed using distributed hydrologic modeling. *Water Resour. Res.* **43,** 1–16 (2007).

67. Easton, Z. M., Walter, M. T., Schneiderman, E. M., Zion, M. S. & Steenhuis, T. S. Including Source-Specific Phosphorus Mobility in a Nonpoint Source Pollution Model for Agricultural Watersheds. *J. Environ. Eng.* **135,** 25–35 (2009).

68. Easton, Z. M., Walter, M. T. & Steenhuis, T. S. Combined monitoring and modeling indicate the most effective agricultural best management practices. *J. Environ. Qual.* **37,** 1798–1809 (2003).

69. Ekstrand, S., Wallenberg, P. & Djodjic, F. Process based modelling of phosphorus losses from arable land. *Ambio* **39,** 100–115 (2010).

70. El-Kaddah, D. N. & Carey, A. E. Water quality modeling of the Cahaba River, Alabama. *Environ. Geol.* **45,** 323–338 (2004).

71. El-Sadek, A. Upscaling field scale hydrology and water quality modelling to catchment scale. *Water Resour. Manag.* **21,** 149–169 (2007).

72. Engelmann, C. J. K., Ward, A. D., Christy, A. D. & Bair, E. S. Application of the BASINS database and NPSM model on a small Ohio watershed. *J. Am. Water Resour. Assoc.* **38,** 289–300 (2002).

73. Exbrayat, J. F. *et al.* Ensemble modelling of nitrogen fluxes: Data fusion for a Swedish meso-scale catchment. *Hydrol. Earth Syst. Sci.* **14,** 2383–2397 (2010).

74. Filoso, S., Vallino, J., Hopkinson, C., Rastetter, E. & Claessens, L. Modeling Nitrogen transport in the Ipswich river, Massachusetts, using a hydrological simulation program in FORTRAN (HSPF). *J. Am. Water Resour. Assoc.* **40,** 1365–1384 (2004).

75. FitzHugh, T. W. & Mackay, D. S. Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. *J. Hydrol.* **236,** 35–53 (2000).

76. Flipo, N., Even, S., Poulin, M., Théry, S. & Ledoux, E. Modeling nitrate fluxes at the catchment scale using the integrated tool CAWAQS. *Sci. Total Environ.* **375,** 69–79 (2007).

77. Flynn, N. J., Paddison, T. & Whitehead, P. G. INCA Modelling of the Lee System: strategies for the reduction of nitrogen loads. *Hydrol. Earth Syst. Sci.* **6,** 467–484 (2002).

78. Francos, A. *et al.* Hydrological and water quality modelling in a medium-sized coastal basin. *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.* **26,** 47–52 (2001).

79. Geza, M., McCray, J. E. & Murray, K. E. Model Evaluation of Potential Impacts of On-Site Wastewater Systems on Phosphorus in Turkey Creek Watershed. *J. Environ. Qual.* **39,** 1636 (2010).

80. Geza, M., Murray, K. E. & McCray, J. E. Watershed-Scale Impacts of Nitrogen from On-Site Wastewater Systems: Parameter Sensitivity and Model Calibration. *J. Environ. Eng.* **136,** 926–938 (2010).

81. Gikas, G. D., Yiannakopoulou, T. & Tsihrintzis, V. A. Modeling of non-point source pollution in a Mediterranean drainage basin. *Environ. Model. Assess.* **11,** 219–233 (2006).

82. Gowda, P. H., Dalzell, B. J. & Mulla, D. J. Model based nitrate TMDLs for two agricultural watersheds of southeastern Minnesota. *J. Am. Water Resour. Assoc.* **43,** 254–263 (2007).

83. Gowda, P. H., Mulla, D. J. & Jaynes, D. B. Simulated long-term nitrogen losses for a midwestern agricultural watershed in the United States. *Agric. Water Manag.* **95,** 616–624 (2008).

84. Granlund, K., Rankinen, K. & Lepistö, A. Testing the INCA model in a small agricultural catchment in southern Finland. *Hydrol. Earth Syst. Sci.* **8,** 717–728 (2004).

85. Green, C. H., Starks, P. J. & Van Liew, M. W. Unit source area data: can it make a difference in calibrating the hydrologic response for watershed-scale modeling? *J. Soil Water Conserv.* **62,** 162 – 170 (2007).

86. Green, M. B. & Wang, D. Watershed flow paths and stream water nitrogen-to-phosphorus ratios under simulated precipitation regimes. *Water Resour. Res.* **44,** 1–13 (2008).

87. Grizzetti, B., Bouraoui, F., Granlund, K., Rekolainen, S. & Bidoglio, G. Modelling diffuse emission and retention of nutrients in the Vantaanjoki watershed (Finland) using the SWAT model. *Ecol. Modell.* **169,** 25–38 (2003).

88. Grizzetti, B., Bouraoui, F. & De Marsily, G. Modelling nitrogen pressure in river basins: A comparison between a statistical approach and the physically-based SWAT model. *Phys. Chem. Earth* **30,** 508–517 (2005).

89. Grunwald, S. & Qi, C. GIS-based water quality modeling in the Sandusky Watershed, Ohio, USA. *J. Am. Water Resour. Assoc.* **42,** 957–973 (2006).

90. Ham, J. H., Yoon, C. G., Jung, K. W. & Jang, J. H. Integrated modelling under uncertainty in watershed-level assessment and management. *Water Sci. Technol.* **56,** 31–39 (2007).

91. Ham, J., Yoon, C. G., Kim, H.-J. & Kim, H.-C. Modeling the effects of constructed wetland on nonpoint source pollution control and reservoir water quality improvement. *J. Environ. Sci. (China)* **22,** 834–839 (2010).

92. Hansen, J. R. *et al.* An integrated and physically based nitrogen cycle catchment model. *Hydrol. Res.* **40,** 347 (2009).

93. Harmel, R. D. & Smith, P. K. Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling. *J. Hydrol.* **337,** 326–336 (2007).

94. Hartman, M. D., Baron, J. S. & Ojima, D. S. Application of a coupled ecosystem-chemical equilibrium model, DayCent-Chem, to stream and soil chemistry in a Rocky Mountain watershed. *Ecol. Modell.* **200,** 493–510 (2007).

95. Hattermann, F. F., Krysanova, V., Habeck, A. & Bronstert, A. Integrating wetlands and riparian zones in river basin modelling. *Ecol. Modell.* **199,** 379–392 (2006).

96. Hattermann, F. F., Krysanova, V. & Hesse, C. Modelling wetland processes in regional applications. *Hydrol. Sci. J.* **53,** 1001–1012 (2008).

97. He, B. *et al.* Integrated biogeochemical modelling of nitrogen load from anthropogenic and natural sources in Japan. *Ecol. Modell.* **220,** 2325–2334 (2009).

98. He, C., Riggs, J. F. & Kang, Y. T. Integration of geographic information systems and a computer model to evaluate impacts of agricultural runnoff on water quality. *Water Resour. Bull.* **29,** 891–900 (1993).

99. Heng, H. H. & Nikolaidis, N. P. Modeling of Nonpoint Source Pollution of Nitrogen At the Watershed Scale. *J. Am. Water Resour. Assoc.* **34,** 359–374 (1998).

100. Hesse, C., Krysanova, V., Päzolt, J. & Hattermann, F. F. Eco-hydrological modelling in a highly regulated lowland catchment to find measures for improving water quality. *Ecol. Modell.* **218,** 135–148 (2008).

101. Hively, W. D., Gérard-Marchant, P. & Steenhuis, T. S. Distributed hydrological modeling of total dissolved phosphorus transport in an agricultural landscape, part II: dissolved phosphorus transport. *Hydrol. Earth Syst. Sci.* **10,** 263 – 276 (2006).

102. Hong, B., Swaney, D. P. & Weinstein, D. A. Simulating spatial nitrogen dynamics in a forested reference watershed, Hubbard Brook Watershed 6, New Hampshire, USA. *Landsc. Ecol.* **21,** 195–211 (2006).

103. Houlahan, J., Marcus, W. A. & Shirmohammadi, A. Estimating Maryland Critical Area Acts Impact on Future Nonpoint Pollution Along the Rhode River Estuary. *Water Resour. Bull.* **28,** 553–567 (1992).

104. Hu, X., McIsaac, G. F., David, M. B. & Louwers, C. A. L. Modeling riverine nitrate export from an East-Central Illinois watershed using SWAT. *J. Environ. Qual.* **36,** 996–1005 (2007).

105. Huang, J. & Hong, H. Comparative study of two models to simulate diffuse nitrogen and phosphorus pollution in a medium-sized watershed, southeast China. *Estuar. Coast. Shelf Sci.* **86,** 387–394 (2010).

106. Huang, S., Hesse, C., Krysanova, V. & Hattermann, F. From meso- to macro-scale dynamic water quality modelling for the assessment of land use change scenarios. *Ecol. Modell.* **220,** 2543–2558 (2009).

107. Huang, Z., Xue, B. & Pang, Y. Simulation on stream flow and nutrient loadings in Gucheng Lake, Low Yangtze River Basin, based on SWAT model. *Quat. Int.* **208,** 109–115 (2009).

108. Hunter, H. M. & Walton, R. S. Land-use effects on fluxes of suspended sediment, nitrogen and phosphorus from a river catchment of the Great Barrier Reef, Australia. *J. Hydrol.* **356,** 131–146 (2008).

109. Hutchins, M. G., Deflandre-Vlandas, A., Posen, P. E., Davies, H. N. & Neal, C. How do river nitrate concentrations respond to changes in Land-use? A modelling case study of headwaters in the river derwent catchment, North Yorkshire, UK. *Environ. Model. Assess.* **15,** 93–109 (2010).

110. Im, S., Brannan, K. M. & Mostaghimi, S. Simulating hydrologic and water quality impacts in an urbanizing watershed. *J. Am. Water Resour. Assoc.* **39,** 1465–1479 (2003).

111. Inamdar, S. & Naumov, A. Assessment of Sediment Yields for a Mixed-landuse Great Lakes Watershed: Lessons from Field Measurements and Modeling. *J. Great Lakes Res.* **32,** 471–488 (2006).

112. Jackson, B. M. *et al.* Catchment-scale modelling of flow and nutrient transport in the Chalk unsaturated zone. *Ecol. Modell.* **209,** 41–52 (2007).

113. Jarritt, N. P. & Lawrence, D. S. Fine sediment delivery and transfer in lowland catchments: modelling suspended sediment concentrations in response to hydrological forcing. *Hydrol. Process.* **21,** 2729 – 2744 (2007).

114. Jarvie, H. P. *et al.* Modelling nitrogen dynamics and distributions in the River Tweed, Scotland: an application of the INCA model. *Hydrol. Earth Syst. Sci.* **6,** 433–454 (2002).

115. Jeon, J. H., Yoon, C. G., Donigian, A. S. & Jung, K. W. Development of the HSPF-Paddy model to estimate watershed pollutant loads in paddy farming regions. *Agric. Water Manag.* **90,** 75–86 (2007).

116. Jeon, J. H., Yoon, C. G., Ham, J. H. & Jung, K. W. Evaluation of BASINS/WinHSPF applicability for pollutant loading estimation for a Korean watershed. *Water Sci. Technol.* **53,** 25–32 (2006).

117. Jha, M., Wolter, C. F., Schilling, K. E. & Gassman, P. W. Assessment of total maximum daily load implementation strategies for nitrate impairment of the Raccoon River, Iowa. *J. Environ. Qual.* **39,** 1317–1327 (2010).

118. Jha, M. K., Arnold, J. G. & Gassman, P. W. Water Quality Modeling for the Raccoon River Watershed Using SWAT. *Trans. ASABE* **50,** 479–493 (2006).

119. Kang, M. S., Park, S. W., Lee, J. J. & Yoo, K. H. Applying SWAT for TMDL programs to a small watershed containing rice paddy fields. *Agric. Water Manag.* **79,** 72–92 (2006).

120. Kannan, N., White, S. M. & Whelan, M. J. Predicting diffuse-source transfers of surfactants to surface waters using SWAT. *Chemosphere* **66,** 1336–1345 (2007).

121. Kaste, Ø. Contrasting Catchments in Norway By Applying the Integrated Nitrogen Model for Catchments (INCA). *Water, Air, Soil Pollut. Focus* **4,** 85–96 (2004).

122. Kaste, Ø. *et al.* Linked models to assess the impacts of climate change on nitrogen in a Norwegian river basin and fjord system. *Sci. Total Environ.* **365,** 200–222 (2006).

123. Kaur, R., Singh, O., Srinivasan, R., Das, S. N. & Mishra, K. Comparison of a subjective and a physical approach for identification of priority areas for soil and water management in a watershed - A case study of Nagwan watershed in Hazaribagh District of Jharkhand, India. *Environ. Model. Assess.* **9,** 115–127 (2004).

124. Kavvas, M. L. *et al.* Watershed Environmental Hydrology Model: Environmental Module and Its Application to a California Watershed. *J. Hydrol. Eng.* **11,** 261–272 (2006).

125. Kim, M.-K., Seo, M.-C., Kim, M.-Y., Chung, J.-B. & Kim, B.-J. Estimating nitrogen and phosphorus contents using model integrated in small agricultural watersheds. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* **39,** 1833–1842 (2004).

126. Kim, S. M., Park, S. W., Lee, J. J., Benham, B. L. & Kim, H. K. Modeling and assessing the impact of reclaimed wastewater irrigation on the nutrient loads from an agricultural watershed containing rice paddy fields. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* **42,** 305–315 (2007).

127. Kovacs, A. Comparative study of two watershed scale models to calculate diffuse phosphorus pollution. *Water Sci. Technol.* **53,** 281–288 (2006).

128. Krysanova, V. & Becker, A. Integrated modelling of hydrological processes and nutrient dynamics at the river basin scale. 131–138 (2000).

129. Krysanova, V., Müller-Wohlfeil, D. I. & Becker, A. Development and test of a spatially distributed hydrological/water quality model for mesoscale watersheds. *Ecol. Modell.* **106,** 261–289 (1998).

130. Lai, G., Yu, G. & Gui, F. Preliminary study on assessment of nutrient transport in the Taihu Basin based on SWAT modeling. *Sci. China, Ser. D Earth Sci.* **49,** 135–145 (2006).

131. Lam, Q. D., Schmalz, B. & Fohrer, N. Modelling point and diffuse source pollution of nitrate in a rural lowland catchment using the SWAT model. *Agric. Water Manag.* **97,** 317–325 (2010).

132. Langusch, J.-J. & Matzner, E. N fluxes in two nitrogen saturated forested catchments in Germany: dynamics and modelling with INCA. *Hydrol. Earth Syst. Sci.* **6,** 383–394 (2002).

133. Lee, M. *et al.* Evaluation of non-point source pollution reduction by applying Best Management Practices using a SWAT model and QuickBird high resolution satellite imagery. *J. Environ. Sci.* **22,** 826–833 (2010).

134. Lee, T. *et al.* Evaluation and spatially distributed analyses of proposed cost effective BMPs for reducing phosphorous level in Cedar Creek Reservoir, Texas. *Trans. ASABE* **53,** 1619–1627 (2010).

135. Lenhart, T., Fohrer, N. & Frede, H. G. Effects of land use changes on the nutrient balance in mesoscale catchments. *Phys. Chem. Earth* **28,** 1301–1309 (2003).

136. León, L. F., Booty, W. G., Bowen, G. S. & Lam, D. C. L. Validation of an agricultural non-point source model in a watershed in southern Ontario. *Agric. Water Manag.* **65,** 59–75 (2004).

137. León, L. F., Lam, D. C., McCrimmon, C. & Swayne, D. A. Watershed management modelling in Malawi: Application and technology transfer. *Environ. Model. Softw.* **18,** 531–539 (2003).

138. León, L. F., Soulis, E. D., Kouwen, N. & Farquhar, G. J. Nonpoint source pollution: A distributed water quality modeling approach. *Water Res.* **35,** 997–1007 (2001).

139. León, L. F., Soulis, E. D., Kouwen, N. & Farquhar, G. J. Modeling diffuse pollution with a distributed approach. *Water Sci. Technol.* **45,** 149–156 (2002).

140. Lewis, T. & Lamoureux, S. F. Twenty-first century discharge and sediment yield predictions in a small high Arctic watershed. *Glob. Planet. Change* **71,** 27–41 (2010).

141. Li, X., Ambrose, R. B. & Araujo, R. Modeling mineral nitrogen export from a forest terrestrial ecosystem to streams. *Trans. ASAE* **47,** 727 – 739 (2004).

142. Liden, R., Vasilyev, A., Stalnacke, P., Loigu, E. & Wittgren, H. B. Nitrogen source apportionment — a comparison between a dynamic and a statistical model. *Ecol. Modell.* **114,** 235–250 (1999).

143. Lindenschmidt, K.-E. & Ollesch, G. Physically-based hydrological modelling for non-point dissolved phosphorus transport in small and medium-sized river basins. *Hydrol. Sci.* **49,** 495–510 (2004).

144. Lindgren, G. A., Wrede, S., Seibert, J. & Wallin, M. Nitrogen source apportionment modeling and the effect of land-use class related runoff contributions. *Nord. Hydrol.* **38,** 317 (2007).

145. Lindström, G., Pers, C., Rosberg, J., Strömqvist, J. & Arheimer, B. Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrol. Res.* **41,** 295 (2010).

146. Lischeid, G. & Langusch, J. Comparative simulation of the nitrogen dynamics using the INCA model and a neural network analysis: implications for improved nitrogen modelling. *Hydrol. Earth Syst. Sci.* **8,** 742–750 (2004).

147. Liu, J., Zhang, L., Zhang, Y., Hong, H. & Deng, H. Validation of an agricultural non-point source (AGNPS) pollution model for a catchment in the Jiulong River watershed, China. *J. Environ. Sci.* **20,** 599–606 (2008).

148. Liu, S., Tucker, P., Mansell, M. & Hursthouse, A. Application of a Water Quality Model in the White Cart water catchment, Glasgow, UK. *Environ. Geochem. Health* **25,** 57–62 (2003).

149. Liu, S., Tucker, P., Mansell, M. & Hursthouse, A. Development and application of a catchment scale diffuse nitrate modelling tool. *Hydrol. Process.* **19,** 2625–2639 (2005).

150. Liu, Y., Yang, W. & Wang, X. Development of a SWAT extension module to simulate riparian wetland hydrologic processes at a watershed scale. *Hydrol. Process.* **22,** 2901 – 2915 (2008).

151. Liu, Z. J., Weller, D. E., Jordan, T. E., Correll, D. L. & Boomer, K. B. Integrated modular modeling of water and nutrients from point and nonpoint sources in the Patuxent river watershed. *J. Am. Water Resour. Assoc.* **44,** 700–723 (2008).

152. Liu, Z. *et al.* Application and evaluation of two nutrient algorithms of Hydrological Simulation Program Fortran in Wolf River watershed. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* **43,** 738–748 (2008).

153. Loos, S., Middelkoop, H., van der Perk, M. & van Beek, R. Large scale nutrient modelling using globally available datasets: A test for the Rhine basin. *J. Hydrol.* **369,** 403–415 (2009).

154. Lunn, R. J., Adams, R., Mackay, R. & Dunn, S. M. Development and application of a nitrogen modelling system for large catchments. *J. Hydrol.* **174,** 285–304 (1996).

155. Luo, Y. & Zhang, M. Management-oriented sensitivity analysis for pesticide transport in watershed-scale water quality modeling using SWAT. *Environ. Pollut.* **157,** 3370–3378 (2009).

156. Luzio, M. D., Arnold, J. G. & Srinivasan, R. A GIS Coupled hydrological model system for the watershed assessment of agricultural nonpoint and point sources of pollution. *Trans. GIS* **8,** 113–136 (2006).

157. Marcé, R. & Armengol, J. Modeling nutrient in-stream processes at the watershed scale using Nutrient Spiralling metrics. *Hydrol. Earth Syst. Sci.* **13,** 953 – 967 (2009).

158. Maringanti, C., Chaubey, I. & Popp, J. Development of a multiobjective optimization tool for the selection and placement of best management practices for nonpoint source pollution control. *Water Resour. Res.* **45,** (2009).

159. Marmefelt, E., Arheimer, B. & Langner, J. An integrated biogeochemical model system for the Baltic Sea. *Hydrobiologia* **393,** 45–56 (1999).

160. McIntyre, N., Jackson, B., Wade, A. J., Butterfield, D. & Wheater, H. S. Sensitivity analysis of a catchment-scale nitrogen model. *J. Hydrol.* **315,** 71–92 (2005).

161. Medici, C. *et al.* Modelling the inorganic nitrogen behaviour in a small Mediterranean forested catchment, Fuirosos (Catalonia). *Hydrol. Earth Syst. Sci.* **14,** 223–237 (2010).

162. Michaud, A. R., Beaudin, I., Deslandes, J., Bonn, F. & Madramootoo, C. A. SWAT-predicted influence of different landscape and cropping system alterations on phosphorus mobility within the Pike River watershed of south-western Québec. *Can. J. soil Sci.* **87,** 329 – 344 (2007).

163. Migliaccio, K. W., Chaubey, I. & Haggard, B. E. Evaluation of landscape and instream modeling to predict watershed nutrient yields. *Environ. Model. Softw.* **22,** 987–999 (2007).

164. Mishra, A., Kar, S. & V. P. Singh. Determination of runoff and sediment yield from a small watershed in sub-humid subtropics using the HSPF model. *Hydrol. Process.* **21,** 3035–3045 (2007).

165. Mitchell, J. K., Engel, B. A., Srinivasan, R. & Wang, S. S. Y. Validation of AGNPS for small watersheds using an integrated AGNPS/GIS system. *J. Am. Water Resour. Assoc.* **29,** 833–842 (1993).

166. Moltz, H. L. N., Lopes, V. L., Rast, W. & Ventura, S. J. Hydrologic-Economic Analysis of Best Management Practices for Sediment Control in the Santa Fe Watershed, New Mexico. *J. Hydrol. Eng.* **15,** 308–317 (2010).

167. Mostaghimi, S., Park, S. W., Cooke, R. A. & Wang, S. Y. Assessment of management alternatives on a small agricultural watershed. *Water Resour.* **31,** 1867–1878 (1997).

168. Muleta, M. K. & Nicklow, J. W. Sensitivity and uncertainty analysis coupled with automatic calibration for a distributed watershed model. *J. Hydrol.* **306,** 127–145 (2005).

169. Muleta, M. K., Nicklow, J. W. & Bekele, E. G. Sensitivity of a Distributed Watershed Simulation Model to Spatial Scale. *J. Hydrol. Eng.* **12,** 163–172 (2007).

170. Narasimhan, B., Srinivasan, R., Bednarz, S. T., Ernst, M. R. & Allen, P. M. A Comprehensive Modeling Approach for Reservoir Water Quality Assessment and Management Due To Point and Nonpoint Source Pollution. *Trans. ASABE* **53,** 1605–1617 (2010).

171. Nasr, A. *et al.* A comparison of SWAT, HSPF and SHETRAN/GOPC for modelling phosphorus export from three catchments in Ireland. *Water Res.* **41,** 1065–1073 (2007).

172. Oehler, F., Durand, P., Bordenave, P., Saadi, Z. & Salmon-Monviola, J. Modelling denitrification at the catchment scale. *Sci. Total Environ.* **407,** 1726–1737 (2009).

173. Olivera, F. *et al.* GIS-SWAT: a geodata model and GIS interface for SWAT. *J. Am. Water Resour. Assoc.* **42,** 295 – 309 (2006).

174. Ouyang, W., Hao, F. H. & Wang, X. L. Regional non point source organic pollution modeling and critical area identification for watershed best environmental management. *Water. Air. Soil Pollut.* **187,** 251–261 (2008).

175. Ouyang, W., Hao, F. H., Wang, X. L. & Cheng, H. G. Nonpoint source pollution responses simulation for conversion cropland to forest in mountains by SWAT in China. *Environ. Manage.* **41,** 79–89 (2008).

176. Ouyang, W., Skidmore, A. K., Hao, F. & Wang, T. Soil erosion dynamics response to landscape pattern. *Sci. Total Environ.* **408,** 1358–1366 (2010).

177. Ouyang, W., Wang, X., Hao, F. & Srinivasan, R. Temporal-spatial dynamics of vegetation variation on non-point source nutrient pollution. *Ecol. Modell.* **220,** 2702–2713 (2009).

178. Pandey, V. K., Panda, S. N., Pandey, A. & Sudhakar, S. Evaluation of effective management plan for an agricultural watershed using AVSWAT model, remote sensing and GIS. *Environ. Geol.* **56,** 993–1008 (2009).

179. Parajuli, P. B., Nelson, N. O., Frees, L. D. & Mankin, K. R. Comparison of AnnAGNPS and SWAT model simulation results in USDA-CEAP agricultural watersheds in south-central Kansas. *Hydrol. Process.* **23,** 748 – 763 (2009).

180. Pease, L. M., Oduor, P. & Padmanabhan, G. Estimating sediment, nitrogen, and phosphorous loads from the Pipestem Creek watershed, North Dakota, using AnnAGNPS. *Comput. Geosci.* **36,** 282–291 (2010).

181. Pettersson, A., Arheimer, B. & Johansson, B. Nitrogen concentrations simulated with HBV-N: new response function and calibration strategy. *Nord. Hydrol.* **32,** 227–248 (2001).

182. Pisinaras, V., Petalas, C., Gikas, G. D., Gemitzi, A. & Tsihrintzis, V. A. Hydrological and water quality modeling in a medium-sized basin using the Soil and Water Assessment Tool (SWAT). *Desalination* **250,** 274–286 (2010).

183. Plus, M. *et al.* Modelling water discharges and nitrogen inputs into a Mediterranean lagoon: Impact on the primary production. *Ecol. Modell.* **193,** 69–89 (2006).

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

185. Quinn, P. F., Hewett, C. J. M. & Dayawansa, N. D. K. TOPCAT-NP: a minimum information requirement model for simulation of flow and nutrient transport from agricultural systems. *Hydrol. Process.* **22,** 2565 – 2580 (2008).

186. Radcliffe, D. E., Lin, Z., Risse, L. M., Romeis, J. J. & Jackson, C. R. Modeling phosphorus in the Lake Allatoona watershed using SWAT: I. Developing Phosphorus Parameter Values. *J. Environ. Qual.* **38,** 121–129 (2009).

187. Rahman, M. & Salbe, I. Modelling impacts of diffuse and point source nutrients on the water quality of South Creek catchment. *Environ. Int.* **21,** 597–603 (1995).

188. Rankinen, K., Granlund, K. & Barlund, I. Modelling of seasonal effects of soil processes on N leaching in northern latitudes. *Nord. Hydrol.* **35,** 347–357 (2004).

189. Rankinen, K., Lepistö, A. & Granlund, K. Hydrological application of the INCA model with varying spatial resolution and nitrogen dynamics in a northern river basin. *Hydrol. Earth Syst. Sci.* **6,** 339–350 (2002).

190. Rankinen, K., Karvonen, T. & Butterfield, D. An application of the GLUE methodology for estimating the parameters of the INCA-N model. *Sci. Total Environ.* **365,** 123–139 (2006).

191. Rankinen, K., Kenttämies, K., Lehtonen, H. & Nenonen, S. Nitrogen load predictions under land management scenarios for a boreal river basin in northern Finland. *Boreal Environ. Res.* **11,** 213–228 (2006).

192. Rankinen, K., Lepistö, A. & Granlund, K. Integrated nitrogen and flow modelling (INCA) in a boreal river basin dominated by forestry: Scenarios of environmental change. *Water, Air, Soil Pollut. Focus* **4,** 161–174 (2004).

193. Rankinen, K. *et al.* Simulated nitrogen leaching patterns and adaptation to climate change in two Finnish river basins with contrasting land use and climatic conditions. *Hydrol. Res.* **40,** 177 (2009).

194. Ranzini, M. *et al.* Integrated Nitrogen CAtchment model (INCA) applied to a tropical catchment in the Atlantic Forest , São Paulo , Brazil. *Hydrol. Earth Syst. Sci.* **11,** 614 – 622 (2007).

195. Ribarova, I., Ninov, P. & Cooper, D. Modeling nutrient pollution during a first flood event using HSPF software: Iskar River case study, Bulgaria. *Ecol. Modell.* **211,** 241–246 (2008).

196. Richards, C. E., Munster, C. L., Vietor, D. M., Arnold, J. G. & White, R. Assessment of a turfgrass sod best management practice on water quality in a suburban watershed. *J. Environ. Manage.* **86,** 229–245 (2008).

197. Rodda, H. J. E., Demuth, S. & Shankar, U. The application of a GIS-based decision support system to predict nitrate leaching to groundwater in southern Germany. *Hydrol. Sci. J.* **44,** 221–236 (1999).

198. Rode, M. & Frede, H.-G. Modification of AGNPS for Agricultural Land and Climate Conditions in Central Germany. *J. Environ. Qual.* **26,** 165 – 172 (1997).

199. Rode, M. & Lindenschmidt, K. E. Distributed sediment and phosporus transport modeling on a medium sized catchment in Central Germany. *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.* **26,** 635–640 (2001).

200. Rollo, N. & Robin, M. Relevance of watershed modelling to assess the contamination of coastal waters due to land-based sources and activities. *Estuar. Coast. Shelf Sci.* **86,** 518–525 (2010).

201. Rossipisa, P., Preti, F., Rossi, M., Ventura, F. & Mazzanti, B. Water, soil and chemical losses: field experiments and model analysis. *Water Sci. Technol.* **39,** 93–102 (1999).

202. Sahu, M. & Gu, R. R. Modeling the effects of riparian buffer zone and contour strips on stream water quality. *Ecol. Eng.* **35,** 1167–1177 (2009).

203. Saleh, A. & Du, B. Evaluation of Swat and Hspf Within Basins Program for the Upper North Bosque River Watershed in Central Texas. *Am. Soc. Agric. Eng.* **47,** 1039–1050 (2004).

204. Salvetti, R. *et al.* Modelling the point and non-point nitrogen loads to the Venice Lagoon (Italy): the application of water quality models to the Dese-Zero basin. *Desalination* **226,** 81–88 (2008).

205. Salvetti, R., Azzellino, A. & Vismara, R. Diffuse source apportionment of the Po river eutrophying load to the Adriatic sea: Assessment of Lombardy contribution to Po river nutrient load apportionment by means of an integrated modelling approach. *Chemosphere* **65,** 2168–2177 (2006).

206. Santhi, C. *et al.* Validation of the SWAT model on a large river basin with point and nonpoint sources. *J. Am. Water Resour. Assoc.* **37,** 1169–1188 (2001).

207. Santhi, C., Arnold, J. G., Williams, J. R., Hauck, L. M. & Dugas, W. A. Application of a watershed model to evaluate management effects on point and nonpoint source pollution. *Trans. ASABE* **44,** 1559–1570 (2001).

208. Schilling, K. E. & Wolter, C. F. Modeling nitrate-nitrogen load reduction strategies for the Des Moines river, Iowa using SWAT. *Environ. Manage.* **44,** 671–682 (2009).

209. Schoumans, O. F. *et al.* Evaluation of the difference of eight model applications to assess diffuse annual nutrient losses from agricultural land. *J. Environ. Monit.* **11,** 540–553 (2009).

210. Scott, J. T., Doyle, R. D., Prochnow, S. J. & White, J. D. Are watershed and lacustrine controls on planktonic N2 fixation hierarchically structured? *Ecol. Appl.* **18,** 805–819 (2008).

211. Setegn, S. G., Dargahi, B., Srinivasan, R. & Melesse, A. M. Modeling of sediment yield from Anjeni-gauged watershed, Ethiopia using SWAT model. *J. Am. Water Resour. Assoc.* **46,** 514 – 526 (2010).

212. Shamshad, A., Leow, C. S., Ramlah, A., Wan Hussin, W. M. A. & Mohd. Sanusi, S. A. Applications of AnnAGNPS model for soil loss estimation and nutrient loading for Malaysian conditions. *Int. J. Appl. Earth Obs. Geoinf.* **10,** 239–252 (2008).

213. Sharma, K. D. & Singh, S. Satellite remote sensing for soil erosion modelling using the ANSWERS model. *Hydrol. Sci. J.* **40,** 259–272 (1995).

214. Shen, Z., Hong, Q., Yu, H. & Liu, R. Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Sci. Total Environ.* **405,** 195–205 (2008).

215. Shen, Z., Hong, Q., Yu, H. & Niu, J. Parameter uncertainty analysis of non-point source pollution from different land use types. *Sci. Total Environ.* **408,** 1971–1978 (2010).

216. Shirinian-Orlando, A. A. & Uchrin, C. G. Modeling the Hydrology and water quality using BASINS/HSPF for the upper Maurice River watershed, New Jersey. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* **42,** 289–303 (2007).

217. Shirmohammadi, A., Chu, T. W. & Montas, H. J. Modeling at catchment scale and associated uncertainties. *Boreal Environ. Res.* **13,** 185–193 (2008).

218. Silgram, M. *et al.* Subannual models for catchment management: evaluating model performance on three European catchments. *J. Environ. Monit.* **11,** 526–539 (2009).

219. Singh, R., Tiwari, K. N. & Mal, B. C. Hydrological studies for small watershed in India using the ANSWERS model. *J. Hydrol.* **318,** 184–199 (2006).

220. Tague, C. L. & Band, L. E. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. *Earth Interact.* **8,** 1–42 (2004).

221. Tague, C., McMichael, C., Hope, A., Choate, J. & Clark, R. Application of the RHESSys model to a California semiarid shrubland watershed. *J. Am. Water Resour. Assoc.* **40,** 575–589 (2004).

222. Talebizadeh, M., Morid, S., Ayyoubzadeh, S. A. & Ghasemzadeh, M. Uncertainty analysis in sediment load modeling using ANN and SWAT model. *Water Resour. Manag.* **24,** 1747–1761 (2010).

223. Taskinen, A. & Bruen, M. Incremental distributed modelling investigation in a small agricultural catchment: 2. Erosion and phosphorus transport. *Hydrol. Process.* **21,** 92 – 102 (2007).

224. Tattari, S., Koskiaho, J., Bärlund, I. & Jaakkola, E. Testing a river basin model with sensitivity analysis and autocalibration for an agricultural catchment in SW Finland. *Agric. Food Sci.* **18,** 428–439 (2009).

225. Tian, Y., Huang, Z. & Xiao, W. Reductions in non-point source pollution through different management practices for an agricultural watershed in the Three Gorges Reservoir Area. *J. Environ. Sci.* **22,** 184–191 (2010).

226. Tolson, B. A. & Shoemaker, C. A. Cannonsville Reservoir Watershed SWAT2000 model development, calibration and validation. *J. Hydrol.* **337,** 68–86 (2007).

227. Tong, S. T. Y., Liu, A. J. & Goodrich, J. A. Assessing the water quality impacts of future land-use changes in an urbanising watershed. *Civ. Eng. Environ. Syst.* **26,** 3–18 (2009).

228. Tripathi, M. P., Panda, R. K. & Raghuwanshi, N. S. Development of effective management plan for critical subwatersheds using SWAT model. *Hydrol. Process.* **19,** 809–826 (2005).

229. Tripathi, M. P., Panda, R. K. & Raghuwanshi, N. S. Identification and prioritisation of critical sub-watersheds for soil conservation management using the SWAT model. *Biosyst. Eng.* **85,** 365–379 (2003).

230. Tsihrintzis, V. & Hamid, R. Runoff quality prediction from small urban catchments using SWMM. *Hydrol. Process.* **12,** 311–329 (1998).

231. Tzoraki, O. & Nikolaidis, N. P. A generalized framework for modeling the hydrologic and biogeochemical response of a Mediterranean temporary river basin. *J. Hydrol.* **346,** 112–121 (2007).

232. Ullrich, A. & Volk, M. Influence of different nitrate-N monitoring strategies on load estimation as a base for model calibration and evaluation. *Environ. Monit. Assess.* **171,** 513–527 (2010).

233. Vachaud, G. & Chen, T. Sensitivity of computed values of water balance and nitrate leaching to within soil class variability of transport parameters. *J. Hydrol.* **264,** 87–100 (2002).

234. Vaché, K. B., Eilers, J. M. & Santelmann, M. V. Water Quality Modeling of Alternative Agricultural Scenarios in the U . S . Corn Belt 1. *J. Am. Water Resour. Assoc.* **38,** 773–787 (2003).

235. van Griensven, A. *et al.* A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J. Hydrol.* **324,** 10–23 (2006).

236. Veith, T. L., Sharpley, A. N. & Arnold, J. G. Modeling a small, northeastern watershed with detailed, field-level data. *Trans. ASABE* **51,** 471–483 (2008).

237. Wade, A. J., Butterfield, D., Griffiths, T. & Whitehead, P. Eutrophication control in river-systems: an application of INCA-P to the River Lugg. *Hydrol. Earth Syst. Sci.* **11,** 584–600 (2007).

238. Wade, A. J. *et al.* A nitrogen model for European catchments: INCA. New model structure and equations. *Hydrol. Earth Syst. Sci.* **6,** 559 – 582 (2002).

239. Wade, A. J., Jackson, B. M. & Butterfield, D. Over-parameterised, uncertain 'mathematical marionettes' - How can we best use catchment water quality models? An example of an 80-year catchment-scale nutrient balance. *Sci. Total Environ.* **400,** 52–74 (2008).

240. Wade, A. J. *et al.* Modelling instream nitrogen variability in the Dee catchment, NE Scotland. *Sci. Total Environ.* **265,** 229–252 (2001).

241. Walton, R. S. & Hunter, H. M. Isolating the water quality responses of multiple land uses from stream monitoring data through model calibration. *J. Hydrol.* **378,** 29–45 (2009).

242. Wang, S. H. *et al.* An integrated modeling approach to total watershed management: Water quality and watershed assessment of Cheney Reservoir, Kansas, USA. *Water. Air. Soil Pollut.* **164,** 1–19 (2005).

243. Wang, S. *et al.* A Modeling Approach to Water Quality Management of an Agriculturally Dominated Watershed, Kansas, USA. *Water. Air. Soil Pollut.* 1–14 (2009). doi:10.1007/s11270-009-0003-2

244. Wang, X., Saleh, A., McBroom, M. W., Williams, J. R. & Yin, L. Test of APEX for nine forested watersheds in East Texas. *J. Environ. Qual.* **36,** 983–995 (2007).

245. White, J. D. *et al.* A combined watershed–water quality modeling analysis of the Lake Waco reservoir: I. Calibration and confirmation of predicted water quality. *Lake Reserv. Manag.* **26,** 147–158 (2010).

246. White, K. L. & Chaubey, I. Sensitivity Analysis, Calibration, and Validations for a Multisite and Multivariable SWAT Model. *J. Am. Water Resour. Assoc.* **41,** 1077–1089 (2005).

247. White, M. J. *et al.* A quantitative phosphorus loss assessment tool for agricultural fields. *Environ. Model. Softw.* **25,** 1121–1129 (2010).

248. Whitehead, P. G., Heathwaite, A. L., Flynn, N. J., Wade, A. J. & Quinn, P. F. Evaluating the risk of nonpoint source pollution from biosolids : integrated modelling of nutrient losses at field and catchment scales. *Hydrol. Earth Syst. Sci.* **11,** 601–613 (2007).

249. Whitehead, P. G., Lapworth, D. J., Skeffington, R. A. & Wade, A. Excess nitrogen leaching and C/N decline in the Tillingbourne catchment, southern England : INCA process modelling for current and historic time series. *Hydrol. Earth Syst. Sci.* **6,** 455–466 (2002).

250. Whitehead, P. G., Wilson, E. J., Butterfield, D. & Seed, K. A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part 2 - application to large river basins in south Wales and eastern England. *Sci. Total Environ.* **210/211,** 559–583 (1998).

251. Whitehead, P., Hill, T. J. & Neal, C. Impacts of forestry on nitrogen in upland and lowland catchments: a comparison of the River Severn at Plynlimon in mid-Wales and the Bedford Ouse in south-east England using the INCA Model. *Hydrol. Earth Syst. Sci.* **8,** 533–544 (2004).

252. Williams, M. R. *et al.* An integrated modelling system for management of the Patuxent River estuary and basin, Maryland, USA. *Int. J. Remote Sens.* **27,** 3705–3726 (2006).

253. Wu, J., Yu, S. & Zou, R. A Water Quality-• Based Approach for Watershed Wide Bmp Strategies. *JAWRA J. Am. Water Resour. Assoc.* **42,** 1193–1204 (2006).

254. Yang, Q. *et al.* Assessing the impacts of flow diversion terraces on stream water and sediment yields at a watershed level using SWAT model. *Agric. Ecosyst. Environ.* **132,** 23–31 (2009).

255. Yoon, K. S., Yoo, K. H. & Soileau, J. M. Nonpoint source (NPS) model simulation of tillage effects on water quality. *J. Environ. Sci. Heal. Part A Environ. Sci. Eng. Toxicol.* **32,** 1491–1506 (1997).

256. Yuan, Y. P., Bingner, R. L. & Rebich, R. A. Evaluation of AnnAGNPS nitrogen loading in an agricultural watershed. *J. Am. Water Resour. Assoc.* **39,** 457–466 (2003).

257. Yuan, Y., Locke, M. A. & Bingner, R. L. Annualized Agricultural Non-Point Source model application for Mississippi Delta Beasley Lake watershed conservation practices assessment. *J. Soil Water Conserv.* **63,** 542–551 (2008).

FIGURE CAPTIONS

Figure S-1: Panel (a) shows the percentage of studies which employ particular timesteps. Panel (b) shows the number of subbasins used by studies which employ the subbasin spatial segmentation method. Note that the relationship between basin size and number of subbasins used was significant but very weak ($p<0.01$, $r^2 = 0.10$).

Figure S-2: Box plots of coefficient of determination (r^2) for: a) selected variables; b) dominant landuse types; c) models used; and d) time steps employed. Numbers above each box indicate number of samples in each group. Variables in (a) were selected to have at least 20 samples. Wildlands in (b) refers to any land other than Forest which is not dominated by human land uses, e.g. grassland. Panels b – d include all variables.

Figure S-3: Box plots of Nash-Sutcliffe Efficiency comparing concentration and load of selected variables.

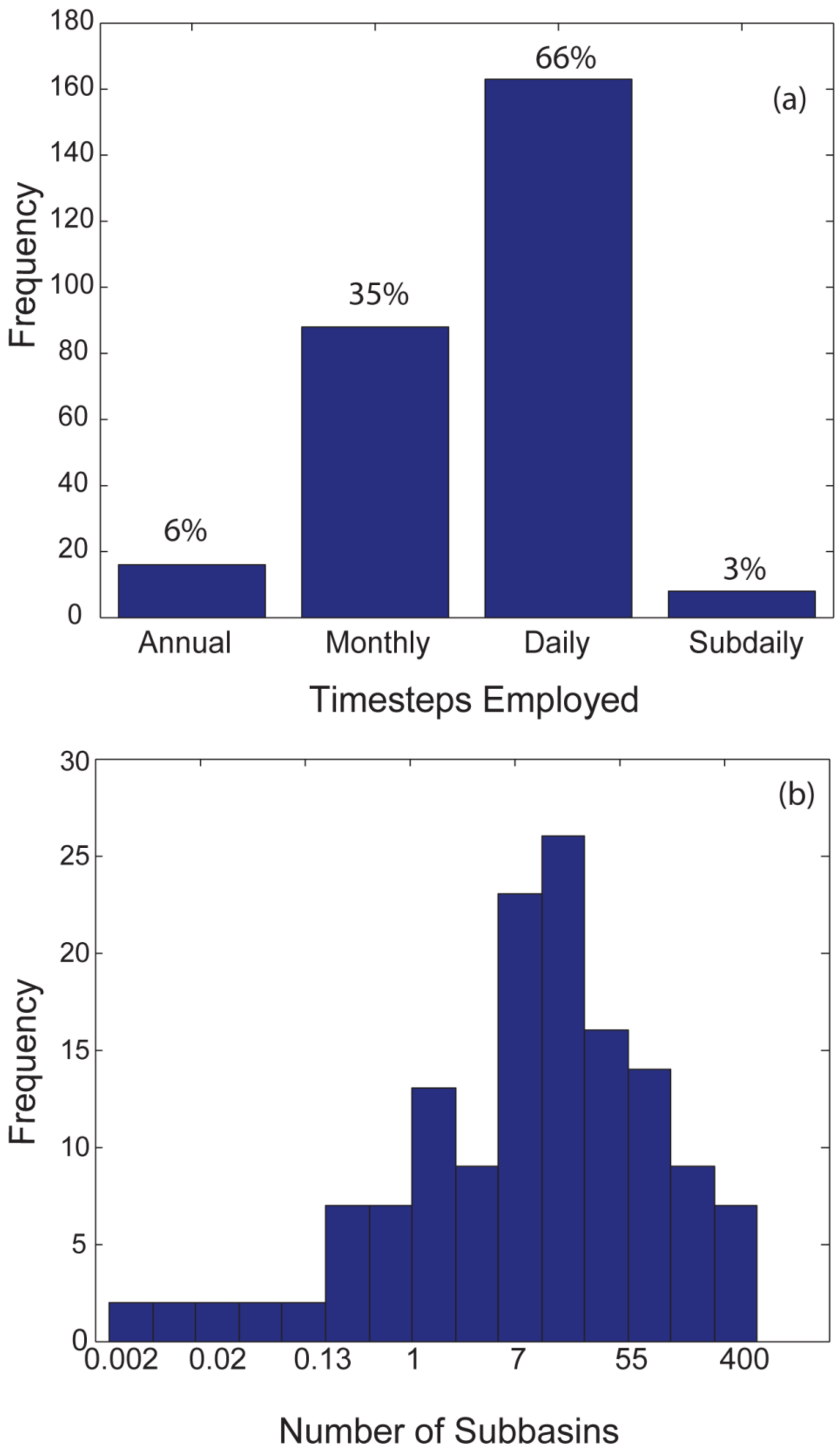


Figure S-1: Bar graphs. Panel (a) shows the percentage of studies which employ particular timesteps. Panel (b) shows the number of subbasins used by studies which employ the subbasin spatial segmentation method. Note that the relationship between basin size and number of subbasins used was significant but very weak ($p<0.01$, $r^2 = 0.10$).

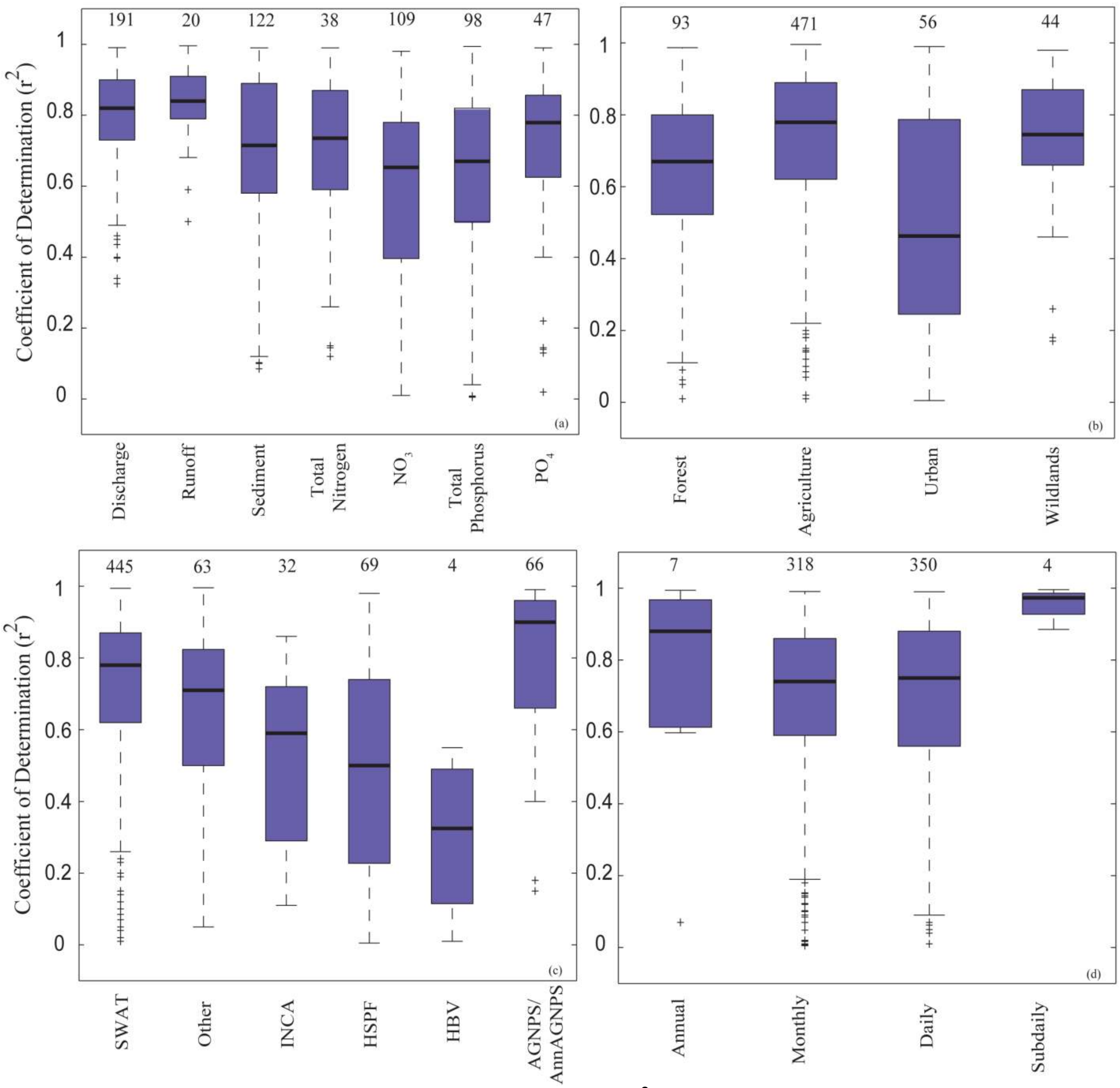


Figure S-2: Box plots of coefficient of determination (r^2) for: a) selected variables; b) dominant landuse types; c) models used; and d) time steps employed. Numbers above each box indicate number of samples in each group. Variables in (a) were selected to have at least 20 samples. Wildlands in (b) refers to any land other than Forest which is not dominated by human land uses, e.g. grassland. Panels b – d include all variables.

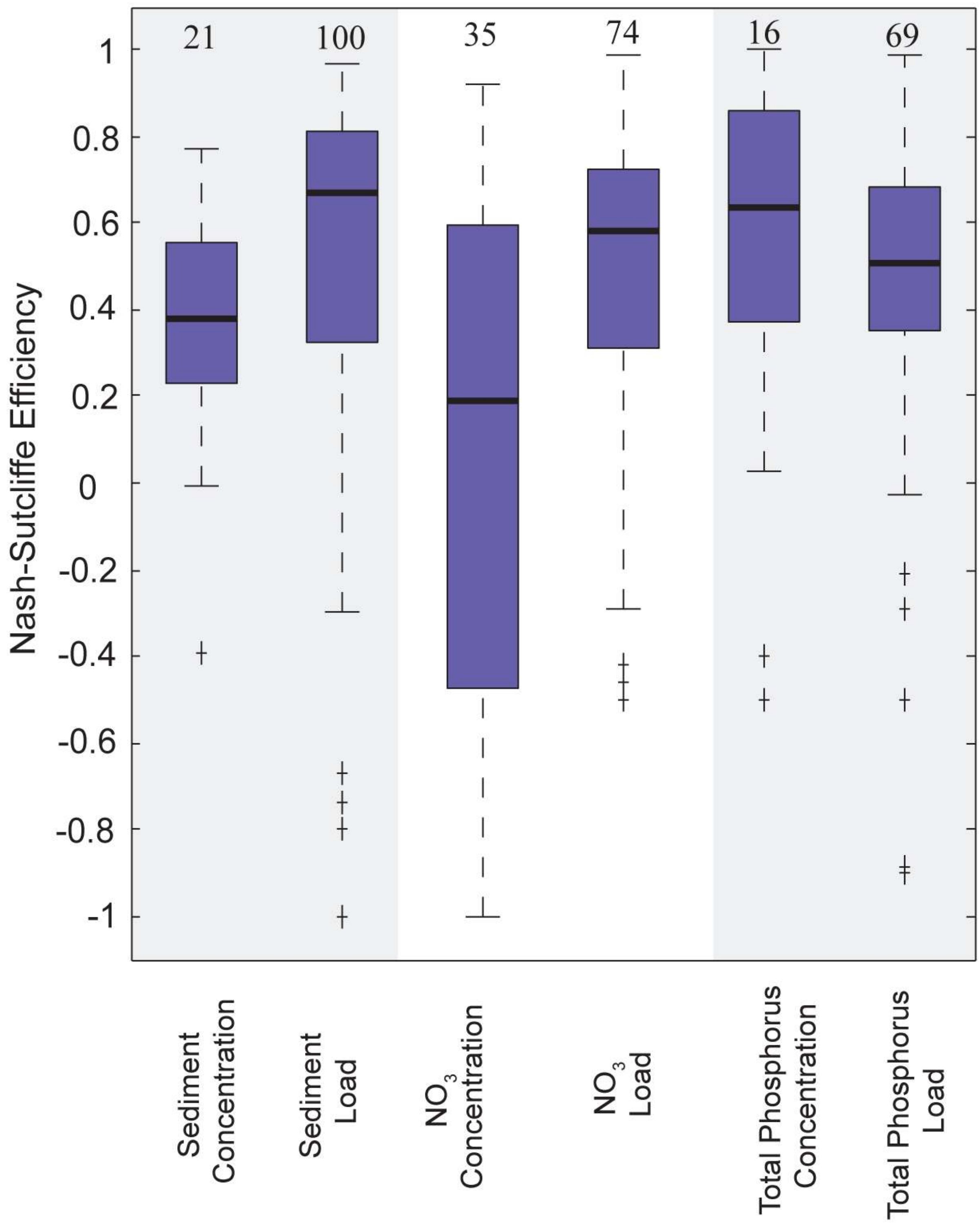


Figure S-3: Box plots of Nash-Sutcliffe Efficiency comparing concentration and load of selected variables.

Table S-1: Percentiles of Nash-Sutcliffe Efficiency of selected water quality variables during calibration and validation periods. Both concentration and load simulations are included. C refers to calibration period, V to validation period.

Variable	Calibration	N	2.5^{th} Percentile	25^{th} Percentile	Median	75^{th} Percentile	97.5^{th} Percentile
Discharge	C	167	0.24	0.61	0.74	0.84	0.94
Discharge	V	121	-0.37	0.56	0.71	0.84	0.93
Runoff	C	13	0.6	0.65	0.79	0.88	0.96
Runoff	V	13	-0.04	0.63	0.75	0.86	0.99
Sediment	C	75	-0.53	0.25	0.56	0.77	0.91
Sediment	V	46	-0.83	0.32	0.66	0.83	0.93
Total Nitrogen	C	69	-0.02	0.4	0.54	0.7	0.85
Total Nitrogen	V	24	-0.46	0.17	0.49	0.69	0.99
NO_3	C	67	-3.62	-0.36	0.35	0.68	0.9
NO_3	V	44	-0.26	0.3	0.51	0.7	0.82
Total Phosphorus	C	51	-0.59	0.25	0.51	0.74	0.97
Total Phosphorus	V	32	-0.72	0.37	0.55	0.66	0.95
PO_4	C	19	0.11	0.54	0.74	0.78	0.86
PO_4	V	15	0.02	0.56	0.71	0.75	0.81