HYDROLOGICAL PROCESSES
Hydrol. Process. **28**, 5624–5636 (2014)
Published online 7 October 2013 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10066

How much complexity is needed to simulate watershed streamflow and water quality? A test combining time series and hydrological models

Huicheng Chien* and D. Scott Mackay

Department of Geography, State University of New York at Buffalo, 105 Wilkeson Quadrangle, Buffalo, NY 14261, USA

Abstract:

Modelled hydrologic processes are represented in a set of numerical equations; the complexity of which can be measured by the total number of variables needed. A single dominant hydrologic process could control the hydrologic response of a watershed, and so the identification of the corresponding dominant variable(s) would aid in identifying a parsimonious model and in collecting more reliable data. By accounting for both model complexity and serial correlation in the variables, a model is used to identify the dominant variables for representing watershed scale streamflow, sediment transport and phosphorus yields. Long-term water quantity and quality data were used to show that rainfall and non-linear soil water storage were the dominant variables for weekly streamflow, suspended sediment and particulate phosphorus. Model accuracy did not consistently improve when other statistically significant variables were included. The results suggest that improved model performance may not justify the added model complexity. As such, identification of dominant variables would be the priority for developing parsimonious hydrologic models, especially at watershed scales. Copyright © 2013 John Wiley & Sons, Ltd.

KEY WORDS complex hydrologic model; dominant variable; parsimonious statistical model; long-term hydrologic data; time series model; water quantity and quality model

Received 1 April 2013; Accepted 9 September 2013

INTRODUCTION

Watershed scale process models of water quantity and quality vary in complexity, number and quantity of processes represented and data required (Merritt *et al.*, 2003; Sivapalan, 2003). The expectation is that by adding complexity to watershed models, one should be able to represent more complex watersheds and their processes (Doherty and Christensen, 2011). In fact, predictions of water quantity and quality at the watershed outlet are not substantially improved by using complex models instead of simple ones (Sivapalan, 2003). This is inevitable, as most watershed scale hydrologic models incorporate process knowledge acquired piecewise in small-scale studies (Klemes, 1983), packaging it in a way that requires a large amount of data and many hard to constrain parameters (Sivakumar, 2008). Modellers have sought sophisticated methods for calibrating hydrologic models to obtain these parameters (Beven, 2001; Duan *et al.*, 1992b; Gupta *et al.*, 1998; Wagner *et al.*, 2004; Yadav *et al.*, 2007). However, few true insights are possible if the calibration exercise cannot identify the simplest model structure needed to represent each process.

Hydrologic processes are represented in a model as a set of numerical equations comprised of variables, which include environmental drivers, flux variables and state variables. The number of these variables within a model is one measure of that model's complexity. As such, the goal of model simplification is to reduce the number of variables by identifying just the key or dominant variable (s), and developing a focused model and measurement plan that emphasizes just these variables (Grayson and Blöschl, 2000a; Woods, 2002b; Young *et al.*, 1996). Sometimes, a hydrologic response is controlled by a single dominant hydrologic process (Grayson and Blöschl, 2000b; Sivakumar, 2004; Sivakumar, 2008; Sivapalan *et al.*, 2003a; Woods, 2002a), which can make the task of simplifying a model easier. When many processes are needed, previous researchers have employed top–down modelling to identify the appropriate model structure and necessary model complexity (Fenicia *et al.*, 2008; Klemes, 1983; Sivapalan *et al.*, 2003a). In top–down modelling, one starts with dominant hydrologic processes (Sivakumar, 2008), adding model complexity only when it is supported by available data. However,

*Correspondence to: Huicheng Chien, Now at Department of Geography, State University of New York at New Paltz, 1 Hawk Drive, New Paltz, NY 12561, USA.
E-mail: hcchien420@gmail.com

it is difficult to ascertain the dominant variables because most of these variables are to some extent correlated with each other (Brodie and Dunn, 2010), and generally serially correlated.

The goal of this research is to investigate the dominant variables and their relative importance for watershed scale hydrologic responses including streamflow, sediment transport and phosphorus yields, accounting for both model complexity and serial correlation in the variables. We had two focusing questions. First, what are the dominant variables driving streamflow, suspended sediment (SS), particulate phosphorus (PP) and soluble reactive phosphorus (SRP) yields? Second, how do the dominant variables influence model performance?

METHODS

Study site

The watershed used for this study was Rock Creek, a tributary of the Sandusky River, in Seneca County, northwestern Ohio, USA (Figure 1). The watershed outlet, with US Geological Survey gauging station 04197170, is located at Rock River at Tiffin, Ohio. Rock Creek watershed drains an area of 89.6 km^2. The geology of the Rock Creek watershed, part of Sandusky River watershed, varies from Devonian limestone and shale to Silurian dolostone (Forsyth, 1975). The Rock Creek watershed lies at the interface between till plain and lake plain environments. The till plains distributed on the south and east of Tiffin are characterized by extensive flat to very gently rolling topography with heavy till soils. The lake plains, distributed at the north of Tiffin area, were formed by the recession of glacial lakes. Dominant soils typically have silt loam and silty clay loam textures (Richards *et al.*, 2002; Riddle, 2006).

Rock Creek watershed is situated in a humid continental climate zone that features cold winters and hot summers. Average low and high monthly temperatures in Tiffin, Ohio, during the years 1983–2007 ranged from −8.24 °C to 2.64 °C in January and from 21.07 °C to 25.12 °C in July. For the same period, the average annual precipitation was around 957.97 mm. The wettest month (June) and driest month (February) had average monthly precipitation of 100.27 and 51.97 mm, respectively. The mean monthly discharge of the Rock River near Tiffin for water year 1983–2007 ranged from 1.55 m^3s^{-1} during high flow in March to 0.26 m^3s^{-1} during low flow in September. Analysis of land use based on National Land Cover Dataset (Homer *et al.*, 2004) has shown that the main land uses within Rock Creek Watershed include agriculture (78.88%), urban (8.27%), forest (11.26%) and water(0.26%).

Hydrologic and water quality data

Daily streamflow, SS, total phosphorus (TP) and SRP from 1983 to 2007 were obtained from the Water Quality Laboratory at Heidelberg College. Three samples per day during storm events and one sample per day during interstorm periods were collected using autosamplers (Richards and Baker, 2002; Richards *et al.*, 2001; Richards *et al.*, 2002). When there was more than one sample in a day, we calculated daily weighted average concentrations of the fluxes using relative flow at each time step during the diurnal period. Streamflow and concentration data were approximately log-normally distributed (Gilbert, 1987), and so to prepare the data for time series analysis, the data were log transformed. Our log-transformed data were approximately normally distributed. In addition, outliers were removed from the log-transformed streamflow and water quality data using Rosner's test (Rosner, 1983). This test is designed to avoid masking one outlier by another, assuming that

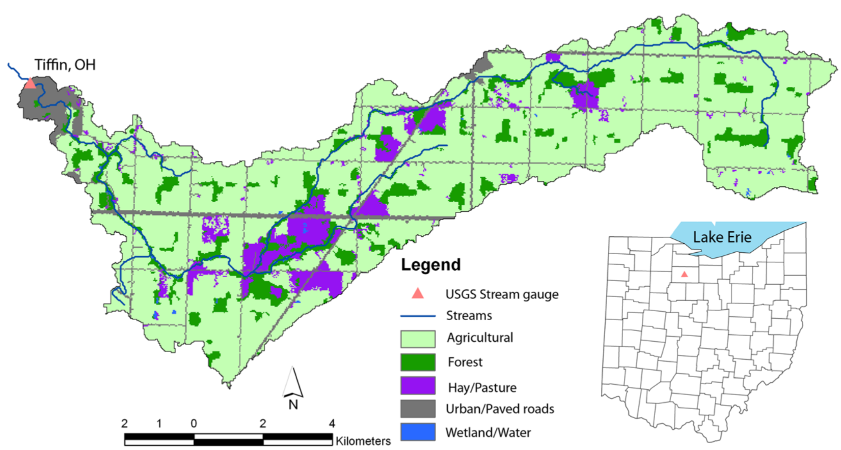

Figure 1. Study site – Rock Creek watershed at Tiffin, OH

observations are normally distributed. Masking occurred when an outlier was undetected because it was very close in value to another outlier. The removed outliers were treated as missing data. All daily water quality concentration data were transformed to unit volume by multiplying the flow volume and dividing by the watershed area (kg ha^{-1} for SS; g ha^{-1} for P). PP was calculated by subtracting the SRP from TP.

Model variables

Observed daily maximum temperature, minimum temperature and precipitation were obtained from the Tiffin National Weather Service Cooperation station (#338313), located about 0.37 km northeast of the watershed outlet. Hourly solar radiation data were simulated by the National Renewable Energy Laboratory (NREL, 1992; NREL, 2007) and summed to get daily solar radiation. Several predictive variables were derived from the observed weather data and model results. Cooling degree days base 5 °C (C_{DD5}), cooling degree days base 10 °C (C_{DD10}) and cooling degree days base 15 °C (C_{DD15}) were calculated as follows:

$$C_{DDn} = \frac{(T_{max} + T_{min})}{2} - T_n \tag{1}$$

where $T_{\max}$, $T_{\min}$ and T_n represent maximum, minimum and base temperature (°C), respectively. In addition to precipitation, net precipitation (P_{net}) was calculated as

$$P_{net} = P - E_a \tag{2}$$

where P is daily precipitation, and E_a is daily actual evapotranspiration. E_a is a function of soil moisture and potential evapotranspiration (E_p). E_a was calculated as

$$E_a = \theta_{rel} {*} E_p \tag{3}$$

where θ_{rel} is relative soil moisture; θ_{rel} is defined as follows:

$$\theta_{rel} = \begin{cases} \dfrac{\theta - \theta_{pwp}}{\theta_{fc} - \theta_{pwp}}, & \theta < \theta_{fc} \\ 1, & \theta \geq \theta_{fc} \end{cases} \tag{4}$$

where θ is the ratio of water volume to soil volume, θ_{fc} is the field capacity and θ_{pwp} is the permanent wilting point of the root-zone soil. θ_{fc} and θ_{pwp} are functions of soil texture. For silt loam, s = 0.22 and $\theta_{pwp} = 0.12$ (Dingman, 2002). E_p was calculated using the Priestley and Taylor Equation (Priestley and Taylor, 1972). Because E_a depends on rooting zone soil moisture (Capehart and Carlson, 1994), the soil water limits to E_a were calculated using a soil depth of 150 mm.

In addition to the environmental drivers, other model variables including soil moisture and runoff were calculated at the watershed scale. Soil moisture was calculated as

$$S_{(t)} = S_{(t-1)} + P_{(t)} - E_{a(t)} - Q_{(t)} \tag{5}$$

where $S_{(t)}$ is the soil moisture at the end of the day (millimetre), and $S_{(t\text{-}1)}$ is the soil moisture of the previous day (millimetre). P, E_a and Q are daily total precipitation (millimetre), actual evapotranspiration (millimetre) and water yield (millimetre), respectively, at day t. The total water (Q) from soil profile could be from saturation excess flow (Q_{se}) and subsurface flow (Q_{ss}). Q_{se} was calculated using the Soil Conservation Service curve number method (USDA-SCS, 1972). Subsurface flow (Q_{ss}) only occurred when soil moisture was over the field capacity and was calculated as

$$Q_{ss} = \begin{cases} K_h^* \left(\dfrac{\theta}{\varnothing}\right)^c, & if\, \theta > \theta_{fc} \\ 0, & if\, \theta \leq \theta_{fc} \end{cases} \tag{6}$$

where K_h^* is saturated hydraulic conductivity, $\varnothing$ is soil porosity and c is a constant set to 13.6 for silt loam in this study.

Selection of dominant variables

A statistical time series model, seasonal autoregressive integrated moving average (SARIMA), was used in this study to account for serial correlation in all hydrologic fluxes. Time series analysis required continuous time series data. Due to equipment failure, daily water quantity and quality data at times had missing data. The proportion of days having data gaps was 6.49% and 14.52% in sediment and phosphorus daily time series, respectively, for the period from 1982 to 2007. Because data gaps in the water quality data precluded using a daily time scale, all hydrologic response data as well as model variables were aggregated into weekly mean data for time series analysis. Weekly temporal resolution for daily data aggregation is selected because it was the optimal temporal resolution to comprise all data gaps especially in water quality data. When there was missing data within a week, the mean of the available data was used to represent the weekly mean data. To understand the temporal influence on the dominant variable identification for watershed hydrologic responses, daily time series of streamflow and SS was also analysed. A regression relationship between streamflow and SS was used to fill the missing data of daily SS time series (Chen and Mackay, 2004).

A SARIMA model is typically expressed as SARIMA (p,d,q)(P,D,Q). Three basic components include the order of autoregressive (p), integrated (d) and moving average (q) components. The P, D and Q terms respectively represent seasonal autoregressive, seasonal integrated and seasonal moving average components. When several possible SARIMA models were available for a hydrologic response time series data, we selected the model that minimized the root mean square error (RMSE) when the model was applied to the calibration period. We also applied Akaike's information criterion (AIC) (Akaike, 1974) as measure of parsimony of the model parameters. AIC was calculated as

$$AIC = n \ln \sigma_a^2 + M \ln(n) \tag{7}$$

where M is the number of parameters in the SARIMA model, and σ_a^2 is model residual variance. A smaller AIC meant that the model was more parsimonious. Consequently, a SARIMA model with both small RMSE and AIC was preferred over models with poorer fit or added complexity.

After the optimal SARIMA model was selected for each hydrologic response, we used it with stepwise regression to select the environmental variables according to RMSE and AIC. Environmental variables were selected for inclusion by first identifying significant lags in the cross correlation analysis. In addition to the original environmental variables, variables were squared to test for nonlinear effects. The final variables chosen for inclusion in the time series model were added to the model one by one and were selected based on comparisons of the SARIMA models with and without inclusion of the extra variables (Alberdi *et al.*, 1998; Trawinski and Mackay, 2008).

Each model was developed and tested by dividing the data set into two periods: Data between 1983–2002 water years were used to construct a SARIMA model, and data for 2003–2007 water years were then used to validate the model by comparing the predictions and observations, which was called the forecasting model in this study. As a test of the sensitivity of our models to periods used for SARIMA model development, we repeated the process such that the SARIMA model development and validation were conducted during 1995–2007 and 1983–1994 water years, respectively. During model validation, goodness-of-fit of the models' predictive capability was assessed using RMSE and index of agreement (d) (Willmott, 1981). Lower confidence interval and upper confidence interval were provided and calculated as

$$\begin{cases} (LCI)_t = (FIT)_t - T_{1-\alpha/2,\quad df(SEP)_t} & t = 1, 2, \ldots N \\ (UCI)_t = (FIT)_t + T_{1-\alpha/2,\quad df(SEP)_t} & t = 1, 2, \ldots N \end{cases} \tag{8}$$

Table I. Characteristics of environmental variables

Weekly environmental variables	Units	Minimum	Maximum	Mean	Standard deviation
Precipitation	mm	0.00	16.04	2.60	2.71
Snowfall	mm	0.00	6.28	0.20	0.62
Rainfall	mm	0.00	16.04	2.38	2.74
Surface runoff	mm	0.00	7.55	0.26	0.81
Snow melt	mm	0.00	5.80	0.22	0.75
Maximum temperature (T_{max})	°C	−13.25	35.08	15.68	10.79
Minimum temperature (T_{min})	°C	−24.76	22.38	5.01	9.29
Mean temperature (T_{ave})	°C	−19.01	27.86	10.34	9.98
Cooling degree days 5 (CDD_5)	Degree days	0.00	22.86	7.68	7.10
Cooling degree days 10 (CDD_{10})	Degree days	0.00	17.86	4.72	5.29
Cooling degree days 15 (CDD_{15})	Degree days	0.00	12.86	2.42	3.34
Solar radiation (R_n)	Watt m^{-2}	26.91	328.65	154.94	76.00
Evapotranspiration (E_a)	mm	0.15	7.37	2.06	1.28
Net precipitation (P_{net})	mm	−5.57	13.78	0.54	2.73
Soil water	mm mm^{-1}	0.18	0.38	0.30	0.05
Precipitation × precipitation	mm^2	0.00	257.23	14.08	29.09
Snowfall × snowfall	mm^2	0.00	39.41	0.42	2.33
Rainfall × rainfall	mm^2	0.00	257.23	13.17	28.76
Surface runoff × surface runoff	mm^2	0.00	57.05	0.73	3.50
Snow melt × snow melt	mm^2	0.00	33.68	0.61	2.71
$T_{max} \times T_{max}$	$°C^2$	0.00	1230.57	362.20	328.73
$T_{min} \times T_{min}$	$°C^2$	0.00	613.15	111.41	118.78
$T_{ave} \times T_{ave}$	$°C^2$	0.00	776.03	206.59	208.21
$CDD_5 \times CDD_5$	Degree $days^2$	0.00	522.46	109.39	133.91
$CDD_{10} \times CDD_{10}$	Degree $days^2$	0.00	318.88	50.24	72.50
$CDD_{15} \times CDD_{15}$	Degree $days^2$	0.00	165.31	17.00	30.25
$R_n \times R_n$	$(Watt\ m^{-2})^2$	724.19	108 013.92	29 777.19	24 953.24
$E_a \times E_a$	mm^2	0.02	54.27	5.89	6.99
$P_{net} \times P_{net}$	mm^2	0.00	189.85	7.75	16.67
Soil water × soil water	mm mm^{-1}	0.03	0.15	0.09	0.03

Table II. Weekly time series seasonal autoregressive integrated moving average models for streamflow, suspended sediment, particulate phosphorus and soluble reactive phosphorus for 1983–2002 water years and 1993–2007 water years

	Weekly streamflow	Weekly suspended sediment	Weekly particulate phosphorus	Weekly soluble reactive phosphorus
	SARIMA (1,0,2)(1,0,1)	SARIMA (1,0,2)(1,0,1)	SARIMA (1,0,2)(1,0,1)	SARIMA (2,0,0)(1,0,1)
Non-seasonal lags	AR 1 MA 1 MA 2	AR 1 MA 1 MA 2	AR 1 MA 1 MA 2	AR 1 AR 2 —
Seasonal lags	Seasonal AR 1 Seasonal MA 1	Seasonal AR 1 Seasonal MA 1	Seasonal AR 1 Seasonal MA 1	Seasonal AR 1 Seasonal MA 1
Regression coefficients	Rainfall lag0 Soil water × soil water lag0 Precipitation × CDD_5 lag0	Rainfall lag0 Soil water × soil water lag0 Maximum temperature lag0	Rainfall lag0 Soil water × soil water lag0 Maximum temperature lag0	Rainfall lag0 Solar radiation lag0 Surface runoff lag1

AR, autoregression; MA, moving average; SARIMA, seasonal autoregressive integrated moving average.

where *FIT* and *SEP* are predicted values and standard errors of the predicted values, respectively, at time series *t*. $T_{1-\alpha/2,\ df}$ is the $(1-\alpha/2)$th percentile of a t distribution with df degrees of freedom, and α is the specified confidence level ($\alpha = 0.05$).

The SARIMA models developed and verified were focused on weekly time series dataset including streamflow, SS, PP and SEP. In terms of daily time series data, SARIMA models for streamflow and SS were developed and verified in forecasting mode only. All analyses were performed using SPSS trends version 13.0 and version 16.0 (SPSS Inc., Chicago, IL).

RESULTS

Statistics of environmental variables

Weekly total simulated water yields compared favourably with weekly total observed streamflow for 1983–2002 (calibration) and 2003–2007 (validation) water years, with goodness-of-fit (*d*) of 0.68 and 0.64, respectively. Table I shows summary statistics for each environmental variable from 1983 to 2007 water years. The weekly maximum temperature, minimum temperature and average temperature were 15.68, 5.01 and 10.34 °C, respectively. Weekly mean total C_{DD5}, C_{DD10} and C_{DD15} were 7.10, 5.29 and 3.34, respectively. Weekly mean precipitation, rainfall, snowfall, snow melt and surface runoff were 2.60, 2.38, 0.20, 0.22 and 0.26 mm, respectively. The mean of evapotranspiration, net precipitation and soil water were 2.06, 0.54 and 0.3.

There were significant correlations between all hydrologic responses and precipitation, snowfall, rainfall, snow melt and surface runoff at various lags (Table S1). Maximum temperature, minimum temperature, mean temperature, C_{DD5}, net precipitation and soil water also had correlation at different lags with all hydrologic responses. The non-linear terms for precipitation and rainfall correlated with streamflow, PP and SRP at several lags, but at lag 0 only for SS. Streamflow and PP were non-linearly correlated to soil water at lag 0, whereas SS and SRP were non-linearly related to soil water at several lags. All of the significant lags of these variables were tested for statistical significance in the SARIMA model. Squared variables for maximum temperature, C_{DD5}, C_{DD10} and C_{DD15} were not significant at any lag.

Dominant variables for streamflow, suspended sediment, particulate phosphorus and soluble reactive phosphorus

The basic SARIMA models for weekly streamflow, SS and PP for 1982–2002 water years and for 1992–2007 water years were (1,0,2)(1,0,1), which had one autoregressive, two moving average, one seasonal autoregressive and one seasonal moving average parameters

(Table II). The basic SARIMA models for daily streamflow and SS were (1,0,4)(1,0,1) and (2,0,2)(1,0,1), respectively (Table S2). For weekly SRP, the basic SARIMA time series model for 1982–2002 water years and 1992–2007 water years could be expressed as (2,0,0)(1,0,1) (Table II).

Rainfall at lag 0 (P_{rain}_lag0) was the dominant variable for weekly streamflow, SS, PP and SRP. Soil water × soil water at lag 0 (SW^2_lag0) was the second relatively import variable for streamflow, SS and PP. For daily streamflow and SS time series, SW^2_lag0 was the dominant variable, and Rainfall at lag 1 (P_{rain}_lag1) was the second relatively import variable (Supplement). For weekly SRP, solar radiation at lag 0 (R_n_lag0) was the second relatively important variable. A third relatively important variable for weekly SS and PP was maximum temperature at lag 0 (T_{max}_lag0). Precipitation × C_{DD5} at lag 0 (PC_{DD5}_lag0) and surface runoff at lag 1 (Q_{se}_lag1) were the third relatively important variables for weekly streamflow and SRP, respectively (Table II).

Table III shows the changes of RMSE and AIC of univariate and multivariate weekly streamflow, SS, PP and SRP time series models for 1982–2002 and 1993–2007 water years. The RMSE and AIC of the univariate streamflow time series model for 1982–2002 water years was 0.95 and 2931.16, respectively. Three environmental variables contributed significantly to the weekly streamflow time series model. When P_{rain}_lag0 was included, the RMSE and AIC were lowered to 0.58 and 2437.98, respectively. By including SW^2_lag0, the RMSE improved to 0.498, and AIC decreased to 2260.99. By also including PC_{DD5}_lag0, both RMSE and AIC improved a small yet significant amount (Table III(a)). Comparing all environmental variables for SS in 1982–2002 water years, P_{rain}_lag0 was relatively more important than the other variables, and it decreased the RMSE from 4.67 to 3.04 and AIC from 4577.19 to 4131.23. When SW^2_lag0 was included in the SS model, RMSE and AIC further decreased to 2.59 and 3990.93, respectively. In addition to P_{rain}_lag0 and SW^2_lag0, Model 3 further included T_{max}_lag0 and only slightly decreased the RMSE and AIC to 2.52 and 3953.67, respectively (Table III(b)).

When P_{rain}_lag0, SW^2_lag0 and T_{max}_lag0 (Table II) were added into the weekly PP model one by one, the RMSE and AIC decreased from 3.23 and 4198.01 to 1.79 and 3584.19, respectively (Table III(c)). The best one-variable model (Model 1) for weekly SRP time series was P_{rain}_lag0 with an RMSE and AIC of 1.92 and 3734.58, respectively. The best two variable model (Model 2) included P_{rain}_lag0 and R_n_lag0 with RMSE and AIC of

Table III. (a) Weekly Streamflow (SF), (b) weekly suspended sediment, (c) weekly particulate phosphorus and (d) weekly soluble reactive phosphorus multivariate seasonal autoregressive integrated moving average model diagnostics for 1983–2002 and 1993–2007 water years

		1982–2002		1993–2007	
	Variables added	RMSE	AIC	RMSE	AIC
(a) Weekly SF					
Model 0	No predictor	0.95	2931.15	1.04	2270.21
Model 1	Rainfall lag0	0.58	2437.98	0.58	1828.62
Model 2	Soil water × soil water lag0	0.50	2260.99	0.49	1694.99
Model 3	Precipitation × CDD_5 lag0	0.49	2226.54	0.48	1669.98
(b) Weekly SS					
Model 0	No predictor	4.67	4577.19	5.12	3518.12
Model 1	Rainfall lag0	3.03	4131.23	3.16	3151.64
Model 2	Soil water × soil water lag0	2.59	3990.93	2.18	3064.54
Model 3	Maximum temperature lag0	2.52	3953.67	2.75	3049.50
(c) Weekly PP					
Model 0	No predictor	3.23	4198.01	3.63	3256.10
Model 1	Rainfall lag0	2.12	3757.58	2.22	2896.18
Model 2	Soil water × soil water lag0	1.84	3612.86	1.96	2771.89
Model 3	Maximum temperature lag0	1.79	3584.19	1.91	2747.43
(d) Weekly SRP					
Model 0	No predictor	2.68	4051.89	3.24	3223.95
Model 1	Rainfall lag0	1.92	3734.58	2.42	2968.91
Model 2	Solar radiation lag0	1.91	3665.71	2.31	2909.37
Model 3	Surface runoff lag1	1.82	3625.37	2.18	2866.26

RMSE, root mean square error; AIC, Akaike's information criterion; SF, streamflow; SS, suspended sediment; PP, particulate phosphorus; SRP, soluble reactive phosphorus.

1.91 and 3665.71, respectively. Model 3 included P_{rain}_lag0, R_n_lag0 and Q_{se}_lag1 and had RMSE and AIC of 1.82 and 3625.37, respectively. Similar results were found for 1993–2007 water years. RMSE and AIC decreased when relatively important variables were included in each hydrologic response time series model (Table III(d)).

SW^2_lag0 and P_{rain}_lag1 were top two dominant variable for daily streamflow and SS time series (Table S2). When the two variables were included in Model 2 for daily streamflow and SS, the AIC decreased to 11 543.54 and 13 547.22, respectively. Similar results were found for RMSE, which decreased to 0.31 and 0.41 for daily streamflow and SS, respectively (Table S3).

Model performance with the dominant variables

Figures 2–5 show forecasting (2003–2007) and hind casting (1983–1992) validation of weekly streamflow, SS, PP and SRP, respectively. Table IV shows the goodness-of-fit measures for the forecasting during the validation using *d* and RMSE. The *d* (RMSE) was 0.46 (1.15), 0.30 (2.49), 0.29 (2.06) and 0.34 (2.04) for streamflow, SS, PP and SRP, respectively, before adding predictors. The *d*

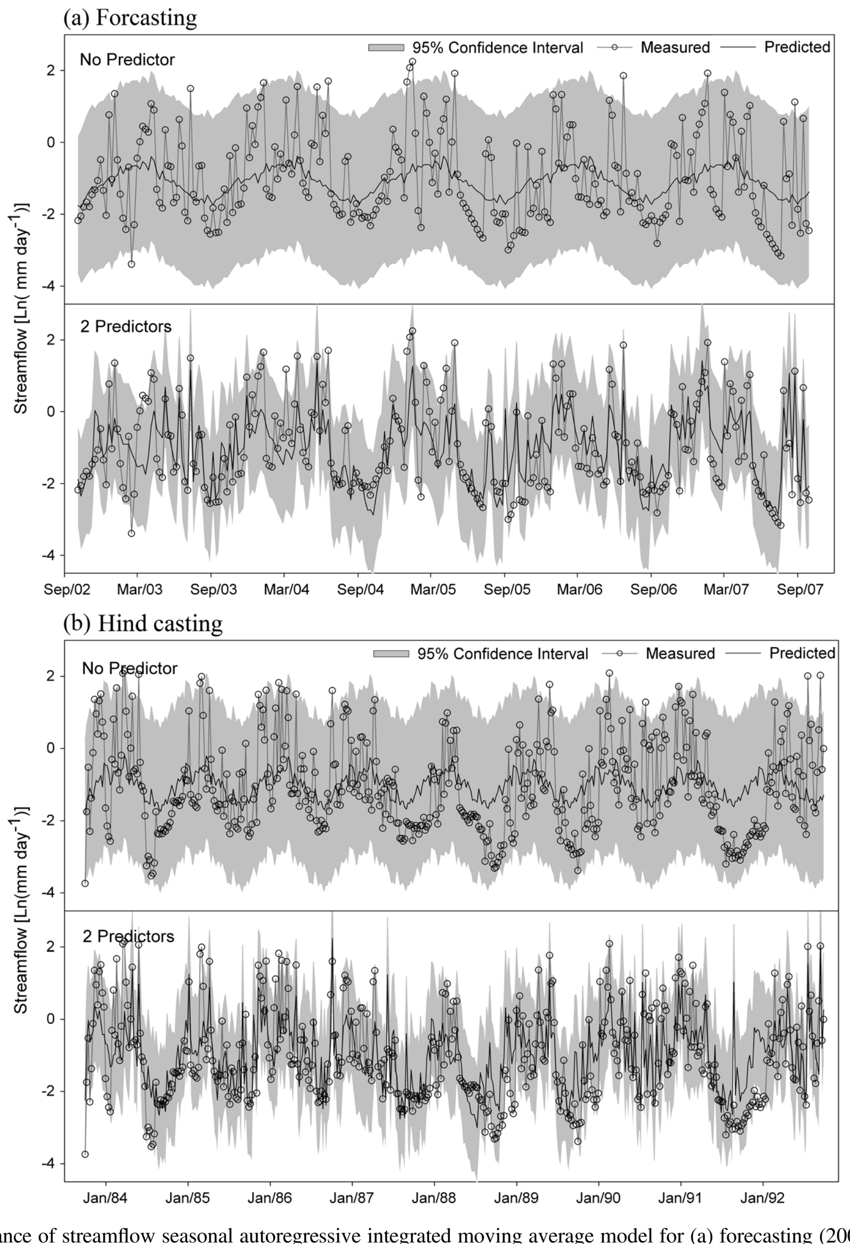

Figure 2. Model performance of streamflow seasonal autoregressive integrated moving average model for (a) forecasting (2003–2007 water years) and (b) hind casting (1983–1992 water years) without and with predictors. Two-predictor model includes P_{rain}_lag0 and SW^2_lag0. Solid line represents best-fit simulations with two predictors. Grey line with circles represents observation. Grey area is 95% confidence intervals

(RMSE) of predictions with predictors significantly improved to 0.85 (0.81), 0.79 (1.95), 0.80 (1.58) and 0.69 (1.68) for streamflow, SS, PP and SRP, respectively, when three predictors were added into the models. As shown in Table IV, the results were similar when the model was used in a hind casting mode.

DISCUSSION

Answers to focus questions

Our results show that watershed scale streamflow had three dominant variables including P_{rain}_lag0, SW^2_lag0 and PC_{DD5}_lag0. P_{rain}_lag0 and SW^2_lag0, and P_{rain}_lag0 and SW^2_lag0 were the top three dominant variables driving SS and PP. The dominant variables driving watershed scale SRP were P_{rain}_lag0, R_n_lag0 and Q_{se}_lag1. When the number of identified dominant variables is a measure of model complexity, our results suggest that the parsimonious hydrologic models could well represent streamflow, SS, PP and SRP at the watershed scale in forecasting (2003–2007) and hind casting (1983–1992) validation periods.

Dominant variables for streamflow, suspended sediment, particulate phosphorus and soluble reactive phosphorus

For weekly streamflow, SS and PP, P_{rain}_lag0 and SW^2_lag0 were the top two common dominant variables.

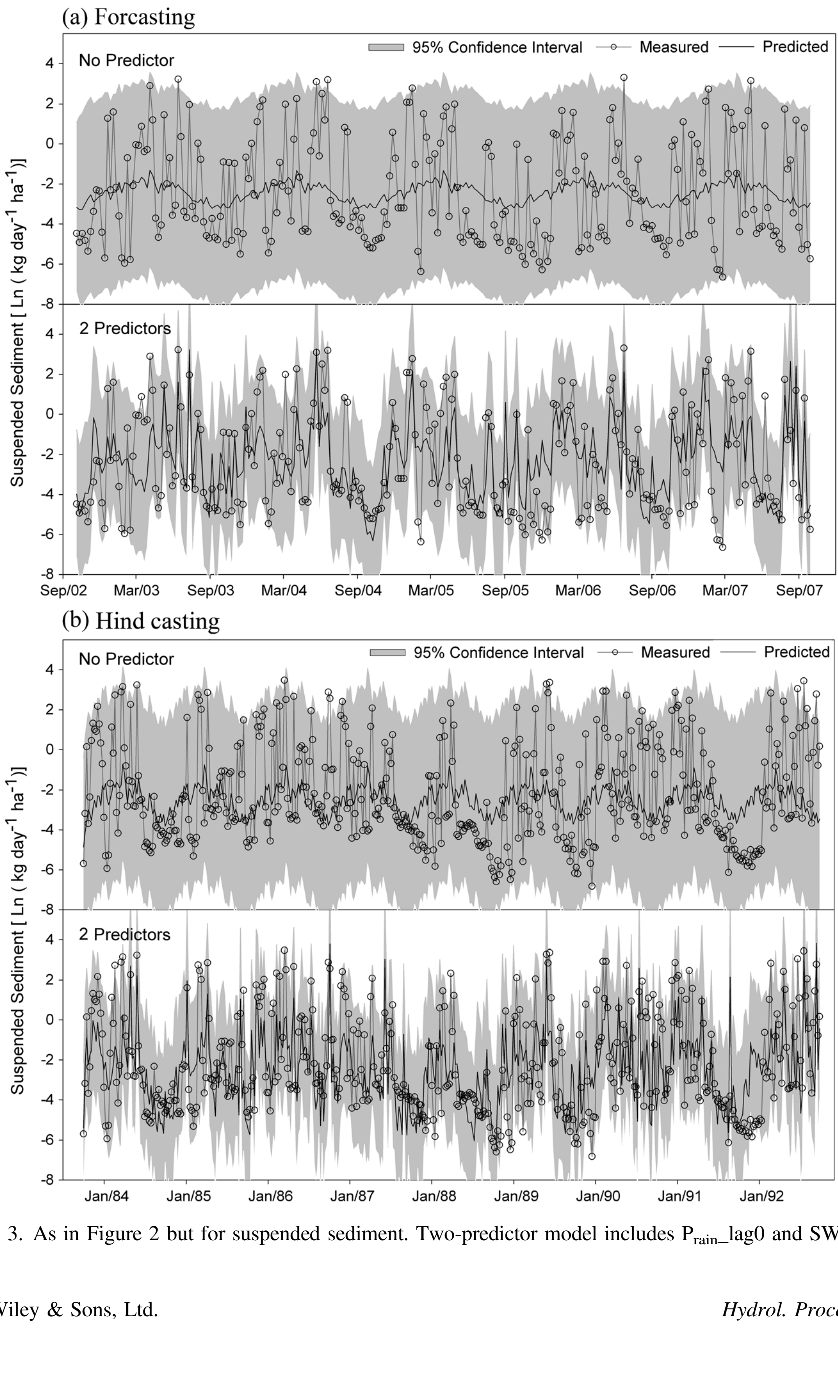

Figure 3. As in Figure 2 but for suspended sediment. Two-predictor model includes P_{rain}_lag0 and SW^2_lag0

It is notable that modelled surface runoff did not improve predictions of SS and PP, and yet surface runoff is widely accepted as a proxy for the energy associated with sediment and phosphorus transport (Hairsine and Rose, 1992; Knighton, 1998) and is widely used as such in models (Beasley *et al*., 1980; Laflen *et al*., 1991; Williams, 1975; Wischmeier and Smith, 1965). Theory on the role of surface runoff as a proxy for sediment transport is firmly grounded at the plot scale, but this theory does not easily scale up to the watershed due to internal heterogeneity of sinks and sources. The added uncertainty of modelling surface runoff as a means of modelling sediment transport is supported by our results and suggests that watershed scale models need improved methods for representing sediment transport capacity rather than relying on plot-scale soil detachment logic (Lane *et al*., 1997).

It is well known that most phosphorus moves in particulate form, attached to the sediment (Bottcher *et al*., 1981; David and Gentry, 2000; Hart *et al*., 2004; Haygarth and Sharpley, 2000; Prairie and Kalff, 1986; Sonzogni *et al*., 1982). The identical dominant variables for SS and PP suggest that dominant variables controlling sediment are useful to predict the PP. The results also support the assumption made by many water quality models such as SWAT (Arnold *et al*., 1998) and AnnAGNPS (Bingner *et al*., 2011), that PP is predictable from SS.

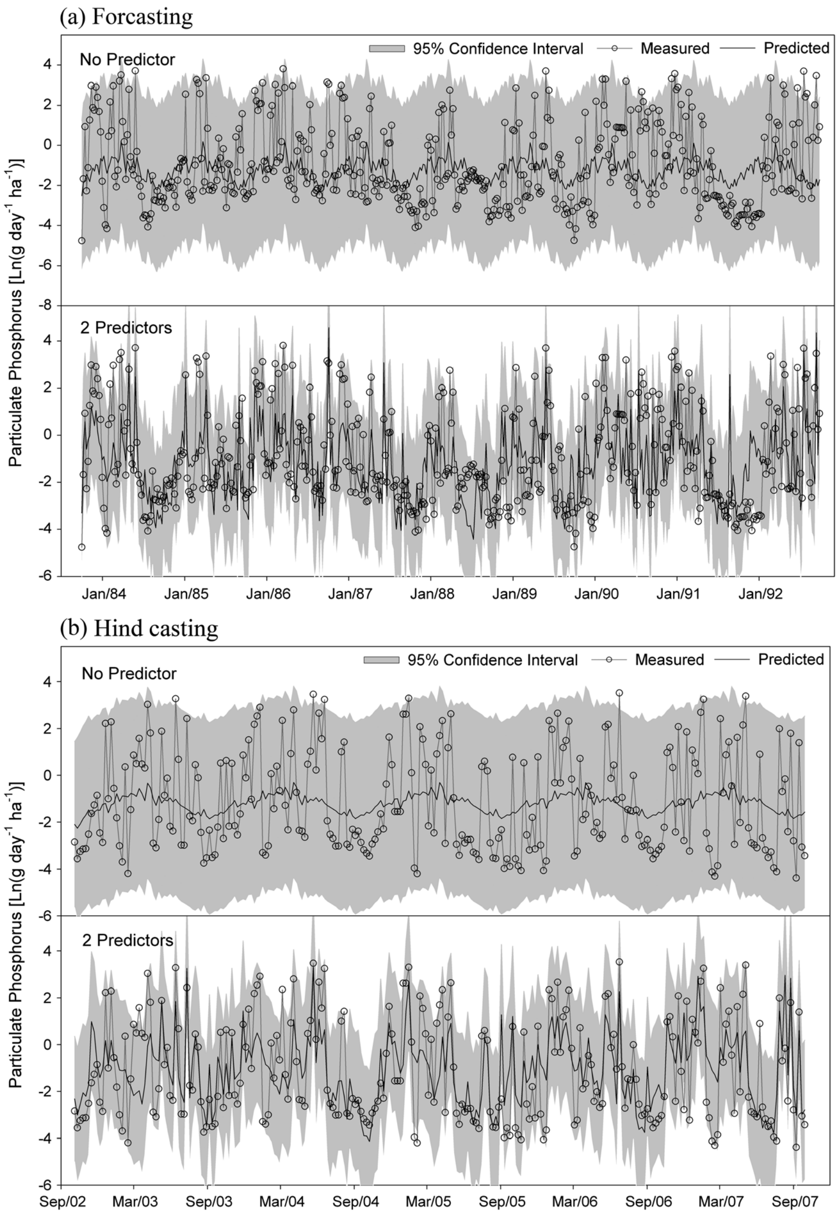

Figure 4. As in Figure 2 but for particulate phosphorus. Two-predictor model includes P_{rain}_lag0 and SW^2_lag0

As shown in Table IV, model performance was marginally improved when the third variable was added. There were diminishing returns with adding additional variables, even when they individually were significantly correlated with the respective hydrologic flux. Increasing model complexity neither made for better predictive performance nor did it reduce prediction uncertainty, as has been suggested by past researchers, e.g. Boyle *et al.*, 2001; Loague and Freeze, 1985 and Van Dijk, 2011. Indeed, Paudel and Jawitz (2012) found that model performance in predicting TP decreased as model complexity increased. This unintended consequence of added model complexity may be attributed to increased uncertainty related to a large number of variables and non-uniqueness in the parameters (Schoups *et al.*, 2008).

Effect of temporal resolution

Temporal resolution of hydrologic response had an effect on the identified dominant variables. For the daily streamflow and SS, SW^2_lag0 and P_{rain}_lag1 were the first and second relatively import variables, respectively. However, rainfall was relatively more important than soil water for weekly streamflow and SS. Soil storage was relatively more important for shorter time series, likely because dry days could be reflected in daily time steps. This is consistent with

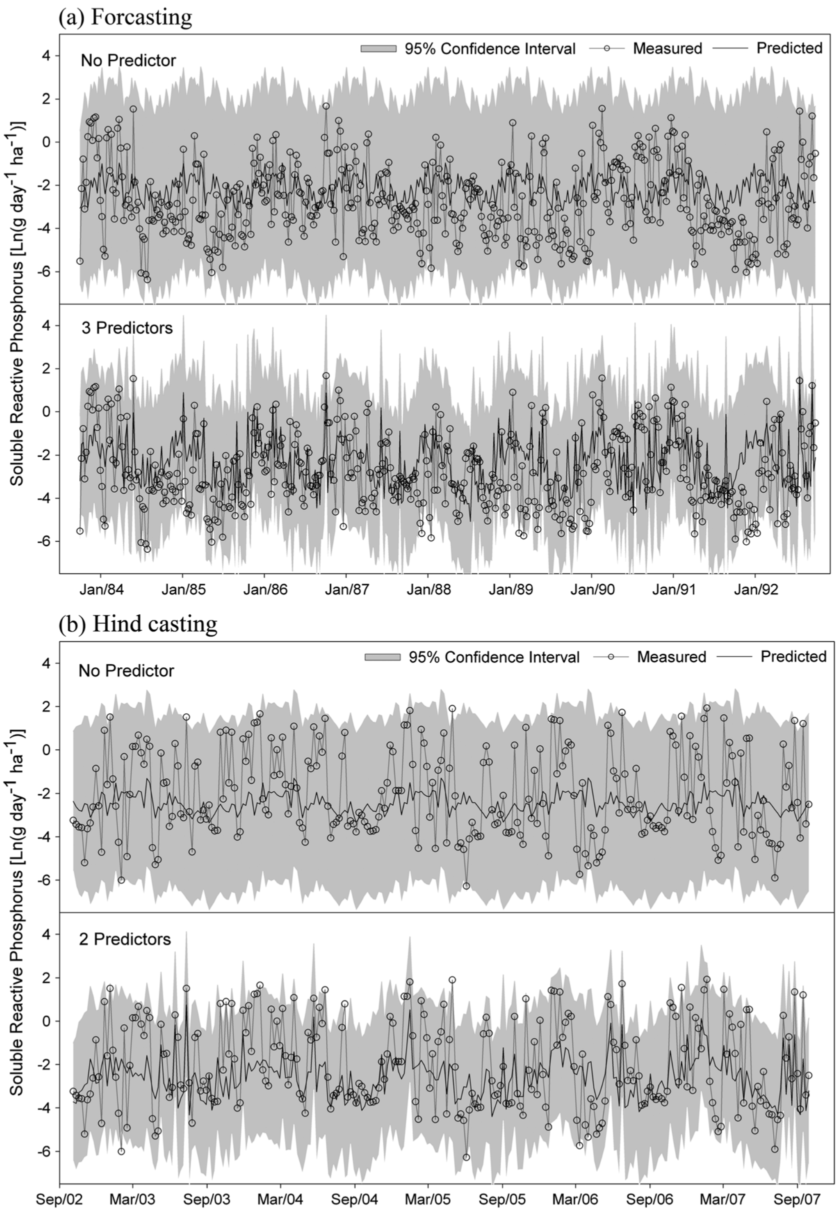

Figure 5. As in Figure 2 but for soluble reactive phosphorus. Two-predictor model includes Prain_lag0 and R_n_lag0

Table IV. Goodness-of-fit measures using *d* and root mean square error for time series model validations for forecasting (2003–2007 water years) and hind casting (1983–1992 water years)

		2003–2007 water years				1983–1992 water years			
		No predictor	1 predictor	2 predictors	3 predictors	No predictor	1 predictor	2 predictors	3 predictors
Streamflow	*d*	0.46	0.78	0.85	0.85	0.42	0.73	0.83	0.85
	RMSE	1.15	0.88	0.81	0.81	1.20	0.99	0.88	0.84
SS	*d*	0.30	0.72	0.79	0.79	0.35	0.57	0.70	0.70
	RMSE	2.49	2.00	1.90	1.95	2.44	2.00	1.84	1.83
PP	*d*	0.29	0.72	0.80	0.80	0.35	0.69	0.79	0.80
	RMSE	2.06	1.66	1.54	1.58	2.03	1.67	1.54	1.54
SRP	*d*	0.34	0.63	0.64	0.69	0.39	0.61	0.66	0.70
	RMSE	2.04	1.73	1.75	1.68	1.88	1.80	1.63	1.59

RMSE, root mean square error; SS, suspended sediment; PP, particulate phosphorus; SRP, soluble reactive phosphorus.

previous studies that showed higher mean soil moisture status when longer temporal resolution precipitation data were used (Abu Bakar and Lu, 2011; Ishidaira *et al.*, 2003). Many of the continuous time watershed models are designed to simulate hydrologic responses at daily or shorter time steps (Merritt *et al.*, 2003; Sudheer *et al.*, 2007). However, continuous water quality data are rarely available on daily time steps except during intensive studies. The results in this study suggest that greater attention should be made to use water quality models with structures based on daily data because model skill at an aggregated time step is no indication of model veracity at shorter time steps.

The merits of long-term and detailed data

The 25-year long data set used in this study is indeed rare, especially because sediment and phosphorous data are rarely measured for more than a year or two in a given watershed. This is problematic, as our results demonstrated and previous studies pointed out that data used for model development and calibration should be in sufficient detail to represent the various real system phenomena including wet, dry and average years experienced by a watershed (Duan *et al.*, 1992a; Duan *et al.*, 1994; Richards, 2004; Sorooshian *et al.*, 1993). Moreover, long-term data are needed to reduce the individual year variations in hydrologic responses that result from extreme weather and other factors (Richards, 2004). Hydrologic variables rarely exhibit stationary behaviour and may contain long-term trends caused by global-scale climate change, and by land-use changes at the local or regional scale (Sivapalan *et al.*, 2003b). By implication, without long-term data, even slow response land-surface processes cannot be detected (Gentine *et al.*, 2012).

CONCLUSIONS

We have shown that rainfall and soil water storage are sufficient for modelling streamflow, SS and PP. Among a list of environmental drivers, flux variables or state variables, only three dominant variables were statistically significant to each hydrologic response in Rock Creek watershed. Other variables may be individually significant to streamflow, SS, PP and SRP, but they were serially correlated with the dominant variables. Consequently, as model complexity increased, model performance did not significantly improve. This suggests that complex models for water quantity and quality should be simplified by reformulating them to use just the dominant variables.

ACKNOWLEDGEMENTS

We thank the Water Quality Laboratory at Heidelberg College in Tiffin, Ohio, for the long-term hydrologic responses data. Primary funding for this research was provided by Nutrient Science grant R-830669 from the US 640 Environmental Protection Agency (EPA) Science to Achieve Results (STAR) program. Additional funding was provided by National Science Foundation (NSF) grant EAR-0405306. Statements made in this manuscript reflect the views of the authors and do not necessarily reflect the views of EPA or NSF.

REFERENCES

Abu Bakar AA, Lu M. 2011. Effects of temporal resolution on river flow forecasting with simple interception model within a distributed hydrological model. *International Journal of Research and Reviews in Applied Sciences* **8**(1): http://www.arpapress.com/Volumes/Vol8Issue1/IJRRAS_8_1_02.pdf

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**(6): 716–723.

Alberdi JC, Diaz J, Montero JC, Miron I. 1998. Daily mortality in Madrid community 1986–1992: relationship with meteorological variables. *European Journal of Epidemiology* **14**(6): 571–578.

Arnold JG, Srinivasan R, Muttiah RS, Williams JR. 1998. Large area hydrologic modeling and assessment - Part 1: Model development. *Journal of the American Water Resources Association* **34**(1): 73–89.

Beasley D, Huggins L, Monke E. 1980. Answers: a model for watershed planning. *Transactions of the ASAE* **23**(4): 938–944.

Beven K. 2001. Rainfall-runoff Modeling – the Primer. John Wiley & Sons Ltd: Chichester, England.

Bingner RL, Heurer FD, Yuan Y. 2011. AnnAGNPS technical processes documentation (Version 5.2). USDA-NRCS, Washington, D.C. http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/water/quality/?&cid=stelprdb1043591

Bottcher AB, Monke EJ, Huggins LF. 1981. Nutrient and sediment loadings from a subsurface drainage system. *Transactions of the ASAE* **24**(5): 1221–1226.

Boyle DP, *et al*. 2001. Toward improved streamflow forecasts: value of semidistributed modeling. *Water Resources Research* **37**(11): 2749–2759.

Brodie IM, Dunn PK. 2010. Commonality of rainfall variables influencing suspended solids concentrations in storm runoff from three different urban impervious surfaces. *Journal of Hydrology* **387**(3–4): 202–211.

Capehart WJ, Carlson TN. 1994. Estimating near-surface soil moisture availability using a meteorologically driven soil-water profile model. *Journal of Hydrology* **160**(1–4): 1–20.

Chen E, Mackay DS. 2004. Effects of distribution-based parameter aggregation on a spatially distributed agricultural nonpoint source pollution model. *Journal of Hydrology* **295**(1–4): 211–224.

David MB, Gentry LE. 2000. Anthropogenic inputs of nitrogen and phosphorus and riverine export for Illinois, USA. *Journal of Environmental Quality* **29**(2): 494–508.

Dingman SL. 2002. Physical Hydrolog (2nd edn.), Prentice-Hall, Inc.: Upper Saddle River, NJ.

Doherty J, Christensen S. 2011. Use of paired simple and complex models to reduce predictive bias and quantify uncertainty. *Water Resources Research* **47**: W12534. DOI: 10.1029/2011WR010763

Duan Q, Sorooshian S, Gupta V. 1992a. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* **28**(4): 1015–1031.

Duan QY, Sorooshian S, Gupta V. 1992b. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research* **28**(4): 1015–1031.

Duan Q, Sorooshian S, Gupta VK. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology* **158**(3–4): 265–284.

Fenicia F, Savenije HHG, Winsemius HC. 2008. Moving from model calibration towards process understanding. *Physics and Chemistry of the Earth* **33**(17–18): 1057–1060.

Forsyth JL. 1975. The Geological Setting of the Sandusky River Basin, Ohio. In Sandusky River Basin Symposium, May 2-3, 1975, Tiffin, Ohio, Baker DB, Prater BL, Jackson WB (eds). International Reference Group on Great Lakes Pollution from Land Use Activities, International Joint Commission. 13–60.

Gentine P, Troy TJ, Lintner BR, Findell KL. 2012. Scaling in surface hydrology: progress and challenges. *Journal of Contemporary Water Research & Education* **147**(1): 28–40.

Gilbert RO. 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold International: NY, Chapter 13.

Grayson R, Blöschl G (eds). 2000a. Spatial Patterns in Catchment Hydrology: Observations and Modeling. Cambridge University Press: Cambridge, UK.

Grayson R, Blöschl G. 2000b. Summary of pattern comparison and concluding remarks. In Spatial Patterns in Catchment Hydrology: Observations and Modeling, Grayson R, Blöschl G (eds). Cambridge University Press: Cambridge, UK; 355–367.

Gupta HV, Sorooshian S, Yapo PO. 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resources Research* **34**(4): 751–763.

Hairsine PB, Rose CW. 1992. Modeling water erosion due to overland flow using physical principles 1. sheet flow. *Water Resources Research* **28**(1): 237–243.

Hart MR, Quin BF, Nguyen ML. 2004. Phosphorus runoff from agricultural land and direct fertilizer effects: a review. *Journal of Environmental Quality* **33**(6): 1954–1972.

Haygarth PM, Sharpley AN. 2000. Terminology for phosphorus transfer. *Journal of Environmental Quality* **29**(1): 10–15.

Homer C, Huang C, Yang L, Wylie B, Coan M. 2004. Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering and Remote Sensing* **70**(7): 829–840.

Ishidaira H, *et al*. 2003. Effect of spatial and temporal resolution of precipitation data on the accuracy of the long-term runoff simulation. Weather Radar Information and Distributed Hydrological Modelling, IAHS Publ. 282: 186–193.

Klemes V. 1983. Conceptualization and scale in hydrology. *Journal of Hydrology* **65**(1–3): 1–23.

Knighton D. 1998. Fluvial Forms and Processes: a New Perspective. Arnold: a member of the Hodder Headlie Group: London.

Laflen JM, Lane LJ, Foster GR. 1991. WEPP: a new generation of erosion prediction technology. *Journal of Soil and Water Conservation* **46**(1): 34–38.

Lane LJ, Hernandez M, Nichols M. 1997. Processes controlling sediment yield from watersheds as functions of spatial scale. *Environmental Modelling & Software* **12**(4): 355–369.

Loague KM, Freeze RA. 1985. A comparison of rainfall-runoff modeling techniques on small upland catchments. *Water Resources Research* **21**(2): 229–248.

Merritt WS, Letcher RA, Jakeman AJ. 2003. A review of erosion and sediment transport models. *Environmental Modelling & Software* **18**(8–9): 761–799.

NREL. 1992. National solar radiation database user's amnnual 1961–2990 update. National Renewable Energy Laboratory, http://rredc.nrel.gov/solar/pubs/NSRDB/NSRDB_index.html.

NREL. 2007. National solar radiation database 1991–2005 update: user's manual. National Renewable Energy Laboratory, Technical Report NREL/TP-581-41364.

Paudel R, Jawitz JW. 2012. Does increased model complexity improve description of phosphorus dynamics in a large treatment wetland? *Ecological Engineering* **42**: 283–294.

Prairie YT, Kalff J. 1986. Effect of catchment size on phosphorus export. *Water Resources Bulletin* **22**(3): 465–470.

Priestley CHB, Taylor RJ. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review* **100**(2): 81–92.

Richards RP. 2004. Improving total maximum daily loads with lessons learned from long-term detailed monitoring. *Journal of Environmental Engineering-ASCE* **130**(6): 657–663.

Richards RP, Baker DB. 2002. Trends in water quality in LEASEQ rivers and streams (northwestern Ohio), 1975–1995. *Journal of Environmental Quality* **31**(1): 90–96.

Richards RP, *et al*. 2001. Storm discharge, loads, and average concentrations in northwest Ohio rivers, 1975–1995. *Journal of the American Water Resources Association* **37**(2): 423–438.

Richards RP, Calhoun FG, Matisoff G. 2002. The Lake Erie agricultural systems for environmental quality project: an introduction. *Journal of Environmental Quality* **31**(1): 6–16.

Riddle CM. 2006. Sandusky River – Tiffin Watershed action plan. Sandusky River Watershed Coalition.

Rosner B. 1983. Percentage points for a generalized ESD many-outlier procedure. *Technometrics* **25**(2): 165–172.

Schoups G, van de Giesen NC, Savenije HHG. 2008. Model complexity control for hydrologic prediction. *Water Resources Research* **44**: W00B03.

Sivakumar B. 2004. Dominant processes concept in hydrology: moving forward. *Hydrological Processes* **18**(12): 2349–2353.

Sivakumar B. 2008. Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stochastic Environmental Research and Risk Assessment* **22**(6): 737–748.

Sivapalan M. 2003. Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection? *Hydrological Processes* **17**(5): 1037–1041.

Sivapalan M, Blöschl G, Zhang L, Vertessy R. 2003a. Downward approach to hydrological prediction. *Hydrological Processes* **17**(11): 2101–2111.

Sivapalan M, *et al*. 2003b. IAHS decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques* **48**(6): 857–880.

Sonzogni WC, Chapra SC, Armstrong DE, Logan TJ. 1982. Bioavailability of phosphorus inputs to lakes. *Journal of Environmental Quality* **11**(4): 555–563.

Sorooshian S, Duan Q, Gupta VK. 1993. Calibration of rainfall-runoff models: application of global optimization to the Sacramento Soil Moisture Accounting Model. *Water Resources Research* **29**(4): 1185–1194.

Sudheer KP, Chaubey I, Garg V, Migliaccio KW. 2007. Impact of time-scale of the calibration objective function on the performance of watershed models. *Hydrological Processes* **21**(25): 3409–3419.

Trawinski PR, Mackay DS. 2008. Meteorologically conditioned time-series predictions of West Nile virus vector mosquitoes. *Vector-Borne and Zoonotic Diseases* **8**(4): 505–521.

USDA-SCS. 1972. National Engineering Handbook, Hydrology Section 4, Chapter 4-10. US Department of Agriculture, Soil Conservation Service: Washington, DC.

Van Dijk AIJM. 2011. Model-data fusion: using observations to understand and reduce uncertainty in hydrological models 19th International Congress on Modelling and Simulation, Perth, Australia, 12-16 December 2011(http://www.mssanz.org.au/modsim2011)

Wagner T, *et al*. 2004. Predictions in ungauged basins as a catalyst for multidisciplinary hydrology. *EOS, Transactions American Geophysical Union* **85**(44): 451–453.

Williams JR. 1975. Sediment-yield prediction with universal equation using runoff energy factor. In Present and prospective technology for prediction sediment yield and sources: Proceedings of the sediment yield workshop, USDA Sedimentation Lab, Oxford, MS, November 28–30, 1972. ARS-S-40, pp. 244–252.

Willmott CJ. 1981. On the validation of models. *Physical Geography* **2**: 184–194.

Wischmeier WH, Smith DD. 1965. Predicting rainfall-erosion losses from cropland east of the rocky mountains. Agriculture Handbook 282, USDA_ARS.

Woods R. 2002a. Seeing catchments with new eyes. *Hydrological Processes* **16**(5): 1111–1113.

Woods R. 2002b. Seeing catchments with new eyes. Spatial patterns in catchment hydrology: observations and modelling Rodger Grayson, Günter Blöschl (Eds.) Cambridge University Press, 416 pp. GBP 65.00, USD$95, AUD215 (ISBN 0-521-63316-8) Published 2000. *Hydrological Processes* **16**(5): 1111–1113.

Yadav M, Wagener T, Gupta H. 2007. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources* **30**(8): 1756–1774.

Young P, Parkinson S,Lees M. 1996. Simplicity out of complexity in environmental modelling: Occam's razor revisited. *Journal of Applied Statistics* **23**(2–3): 165–210.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site:

Table S1: Significant leading indicator environmental variables from cross-correlation analysis. Correlations are significant at the 0.05 level.

Table S2: Daily time series SARIMA models for streamflow and suspended sediment (SS) for 1983–2002 water years.

Table S3: Daily streamflow and daily SS multivariate SARIMA model diagnostics for 1983–2002 water years.