Water Resources Research

Supporting Information for

Integrating Water Quality Data with a Bayesian Network Model to Improve Spatial and Temporal Phosphorus Attribution: Application to the Maumee River Basin

Zihan Wei[1]*, Sarfaraz Alam[1,2*], Miki Verma[3*], Margaret Hilderbran[2], Yuchen Wu[4], Brandon Anderson[2], Daniel E. Ho[2†], Jenny Suckale[1†]

[*]Joint First Authors
[†]Equal co-supervision
[1]Department of Geophysics, Stanford University, Stanford, CA, USA
[2]Regulation, Evaluation and Governance Lab, Stanford University, Stanford, CA, USA
[3]Symbolic Systems Program, Stanford University, Stanford, CA, USA
[4]Department of Statistics, Stanford University, Stanford, CA, USA

Contents of this file

Text S1 to S2

Figures S1 to S2

Tables S1 to S2

Introduction

This supporting information document presents text, figures, and tables that provide additional details about the methods and results outlined in the manuscript.

Text S1. Hydrologic model set-up.

We used the Soil and Water Assessment Tool (SWAT), a semi-distributed, physically based hydrologic model, to simulate the hydrologic processes in the Maumee River Basin. The SWAT model solves the water balance equation at its smallest calculation unit, known as a hydrologic response unit (HRU), to quantify water flux and changes in storage. Each HRU is determined from the unique combinations of land use, soil, and slope data. The key strength of the SWAT model is that it can represent the physical hydrologic processes and model agricultural and water management changes, all while being computationally faster than commonly used distributed hydrologic models like Variable Infiltration Capacity (VIC).

We used topography, soil, land use, and meteorological time series data to set up the SWAT model for the Maumee River Basin. Table S1 lists the type, source, and resolution of each data set used for the SWAT model. We followed four key steps for model development.

First, we delineated the watersheds using topographic data (via a digital elevation model, or DEM). The elevation data was used to compute flow direction and flow accumulation (i.e., the number of grids contributing flow to each grid). Streams generally have relatively higher flow accumulation value (or higher number of grids upstream contributing flow), which then used to separate stream networks. Based on the threshold area for flow accumulation, the stream network density was determined. Using smaller thresholds yielded denser networks. We tested several thresholds in an attempt to obtain stream network density resembling the USGS HUC-12 watersheds, ultimately using a threshold of 3000 ha, or 30 km^2. However, we note that the areas we obtained are not identical to HUC-12 watersheds.

Furthermore, to obtain simulated outputs at locations with USGS water quality and flow measurements, we added outlet points at these locations. In a few cases, USGS monitor locations are not exactly on the streamlines due to errors in delineated stream network locations. The typical approach in this scenario is to snap the monitor locations to the streamline, which we did for distances up to 100 m from the stream. We note that positioning outlets subdivides a subwatershed into two, which in some cases resulted in the creation of much smaller subwatersheds. Furthermore, we identified two channels in the NHDPlus stream network that are undirected cycles (loops independent of edge direction); these generally occur when there are bypasses or irrigation channels. Because we seek to aggregate the channel contributions at subwatershed scale, we collapsed these loops into single edges. In summary, the watershed delineation process yields subwatersheds with outlets located at the water quality monitors and stream junctions. For reference, these monitor and junction nodes are later used to simulate pollutant transport through the stream network, while the subwatersheds are used as source nodes.

Second, we used the land use map from USDA Cropland Data Layer (Han et al., 2012), the soil map from SSURGO (Soil Survey Staff, 2015), and slope information derived from DEM to determine HRUs. SWAT used these three datasets to find unique combinations of land parcels, which are defined to be the HRUs. All simulation in SWAT is first computed at the HRU level, then aggregated at the subwatershed level.

Third, we forced the model with temperature and precipitation data from PRISM (PRISM Climate Group, 2014) to simulate the model at daily time steps from 2014 through 2020. We used 2014 as the spinning period (or warming period, which is necessary for model stability), so the simulation output is available from 2015 to 2020.

Fourth, we calibrated the model using SWAT-simulated streamflow as the calibration variable and the USGS streamflow data as the 'observed' data. The objective function for calibration was to maximize Kling-Gupta Efficiency (KGE). Details about calibration and validation are provided in the following section (Text S2).

Figure S1 shows the elevation, land use, and soil maps used as inputs to the SWAT model.

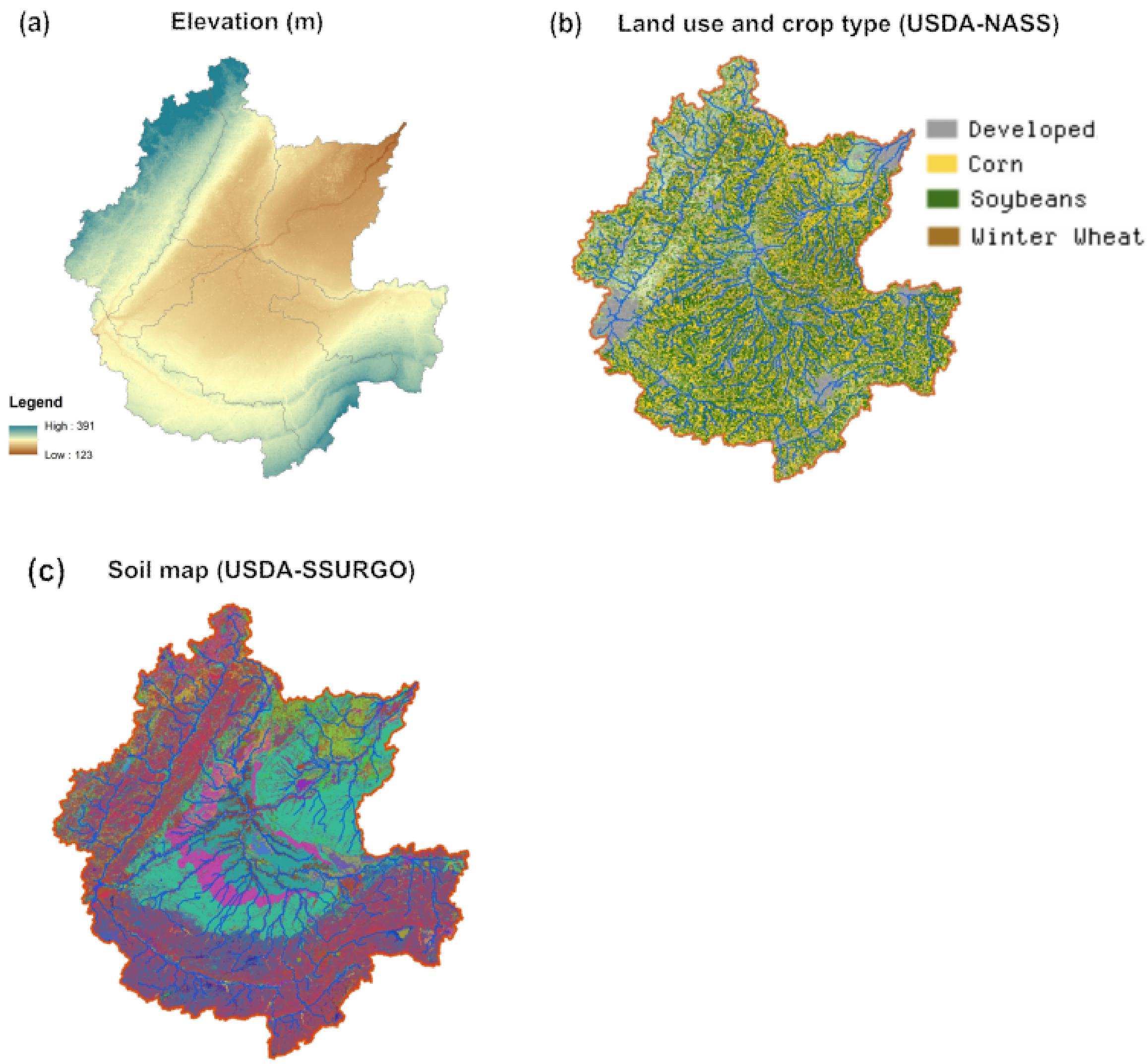

Figure S1. Input data to the SWAT hydrologic model. (a) 30-m elevation map from Shuttle Radar Topographic Mission (SRTM). (b) 30-m land use and crop type map from USDA-NASS. Legend for land use includes only the dominant land use types; others are not shown for concision. (c) 10-m SSURGO soil map from USDA. Legend for soil type, which consists of a large number of soil types, is not shown for brevity.

Text S2. SWAT calibration and validation.

When calibrating the SWAT model, we search for the model parameter values for which model simulation most closely matches the in-situ measurements. Because we use simulated flow and runoff for source attribution, we calibrate the model using streamflow, or river discharge. Hydrologic model calibration is generally suggested to be treated as a multi-objective problem using either multi-site or multi-variable measurements or multi-response function (Gupta et al., 2009; van Griensven & van Bauwens, 2003; Madsen, 2003). In this study we performed multi-site calibration. For calibration, we chose Kling-Gupta-Efficiency (KGE) as the objective function (Kling et al., 2012). KGE includes correlation (r), variability (α), and bias error (β) in its goodness-of-fit criterion (Gupta et al., 1998; Kling et al., 2012). This goodness-of-fit criterion measures the match between simulated and observed values on a scale ranging from negative infinity to 1, where 1 indicates a perfect match.

The objective function for optimization is:

$$KGE_{mean} = \sum_{j=1}^{k} \frac{1}{k} \left(1 - \sqrt{((r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2)_j}\right)$$

Where,

$$\alpha = \frac{\sigma_{sim}/\mu_{sim}}{\sigma_{obs}/\mu_{obs}}$$

and

$$\beta = \frac{\mu_{sim}}{\mu_{obs}}$$

and k is the total number of streamflow measurement sites; r is the regression coefficient, α is a measure of relative variability (variability ratio); and β is the bias ratio (the ratio of the simulated and observed means, μ_{sim} and μ_{obs}, respectively). σ is the standard deviation. We calibrated the streamflow for three years: 2015, 2017, and 2019. The validation periods were the alternate years: 2016, 2018, and 2020.

We used the Dynamically Dimensioned Search Algorithm (DDS), a widely used method for hydrologic calibration, to optimize SWAT model parameters (Lin et al., 2017; Tolson and Shoemaker, 2007). The key advantage of DDS over commonly-used global search algorithms (e.g., the shuffled complex evolution algorithm) is the ability to dynamically adjust search space by successively decreasing parameter dimension until iterations reach a user-defined limit. For this study, we calibrated seven parameters and iterated 3000 times, using the tool Ostrich that has a built-in DDS algorithm (Matott, 2017). Parameter selection was based on the most commonly-used parameters for streamflow calibration (Abbaspour et al., 2015; Zambrano-Bigiarini & Rojas, 2013), as well as our experiments to identify most sensitive parameters. The calibration parameters used in our SWAT model are listed in Table S1.

Table S1. Calibrated parameters for the SWAT model. Here, R indicates that an existing parameter value is multiplied by (1+ a given value), while V indicates that the existing parameter value is replaced by a given value.

Parameter	*Definition*	*Type of change*	*Range*	*Fitted value*
CN2	Curve number for moisture condition II	R	-0.25 _ 0.25	0.04
ALPHA_BF	Baseflow alpha factor for bank storage (days)	V	0 _ 1	0.99
SURLAG	Surface runoff lag coefficient	V	0.01 _ 2	0.3
GW_DELAY	Groundwater re-evaporation factor	V	0.01 _ 50	0.07
SOL_AWC	Available soil water capacity (mm H2O/mm soil)	R	-0.8 _ 0.2	0.04
SOIL_K	Saturated hydraulic conductivity (mm/h)	R	-0.8 _ 0.2	0.0017
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur (mm)	V	-1000 _ 1000	657

Table S2 shows the performance metrics. We find the KGE values for the calibration and validation periods are 0.78 and 0.82, respectively. R^2 for the calibration and validation periods are 0.87 and 0.83, respectively. The KGE and R^2 values testing the match between simulated and observed flow indicate overall satisfactory model performance.

Table S2. Performance metrics for calibration and validation periods.

Evaluation criterion	*Calibration period*	*Validation period*
KGE	0.78	0.82
R^2	0.87	0.83

Figure S2 compares simulated and observed streamflows at multiple USGS sites.

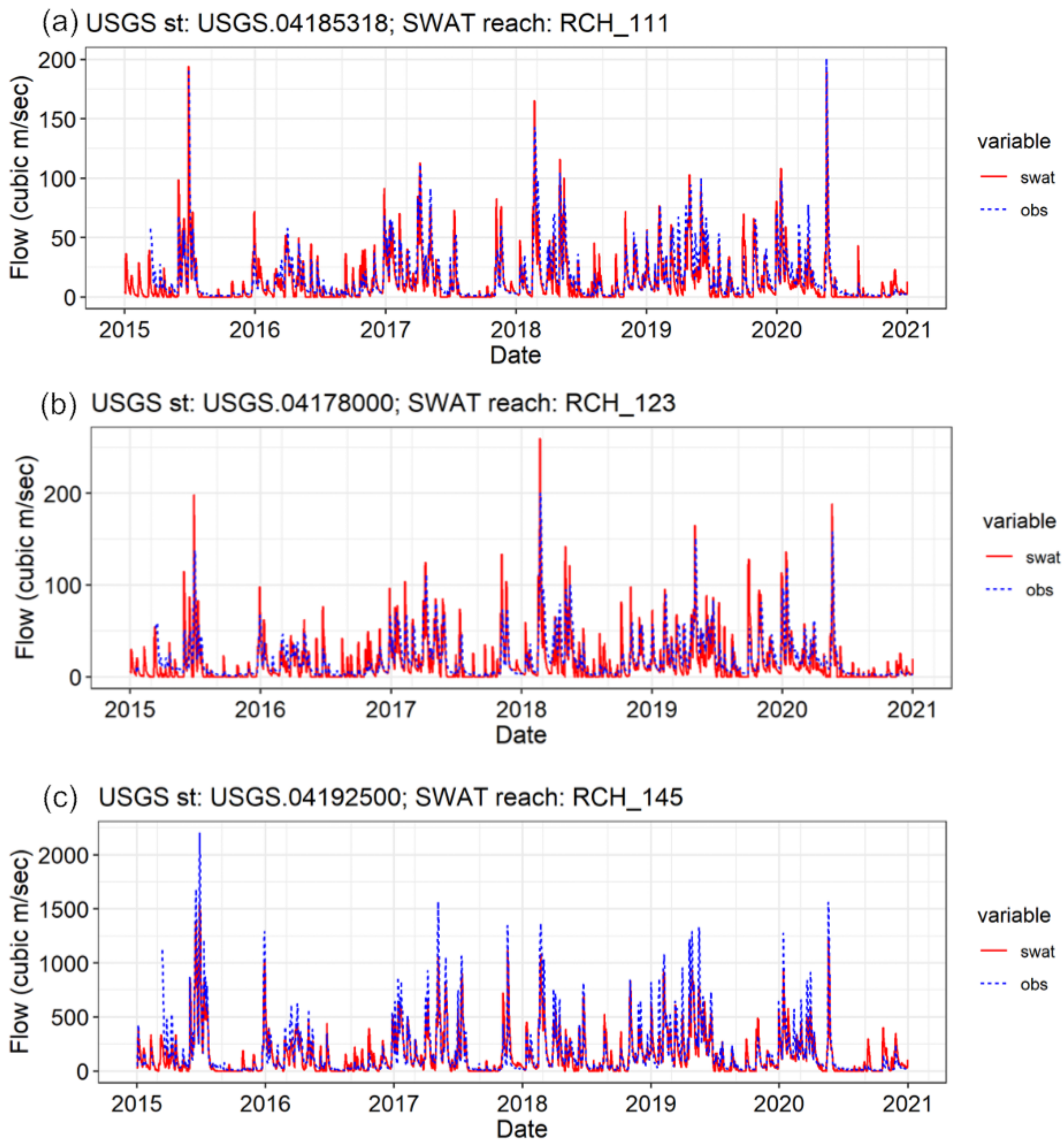

Figure S2. SWAT-simulated vs. USGS observed flow at selected sites.

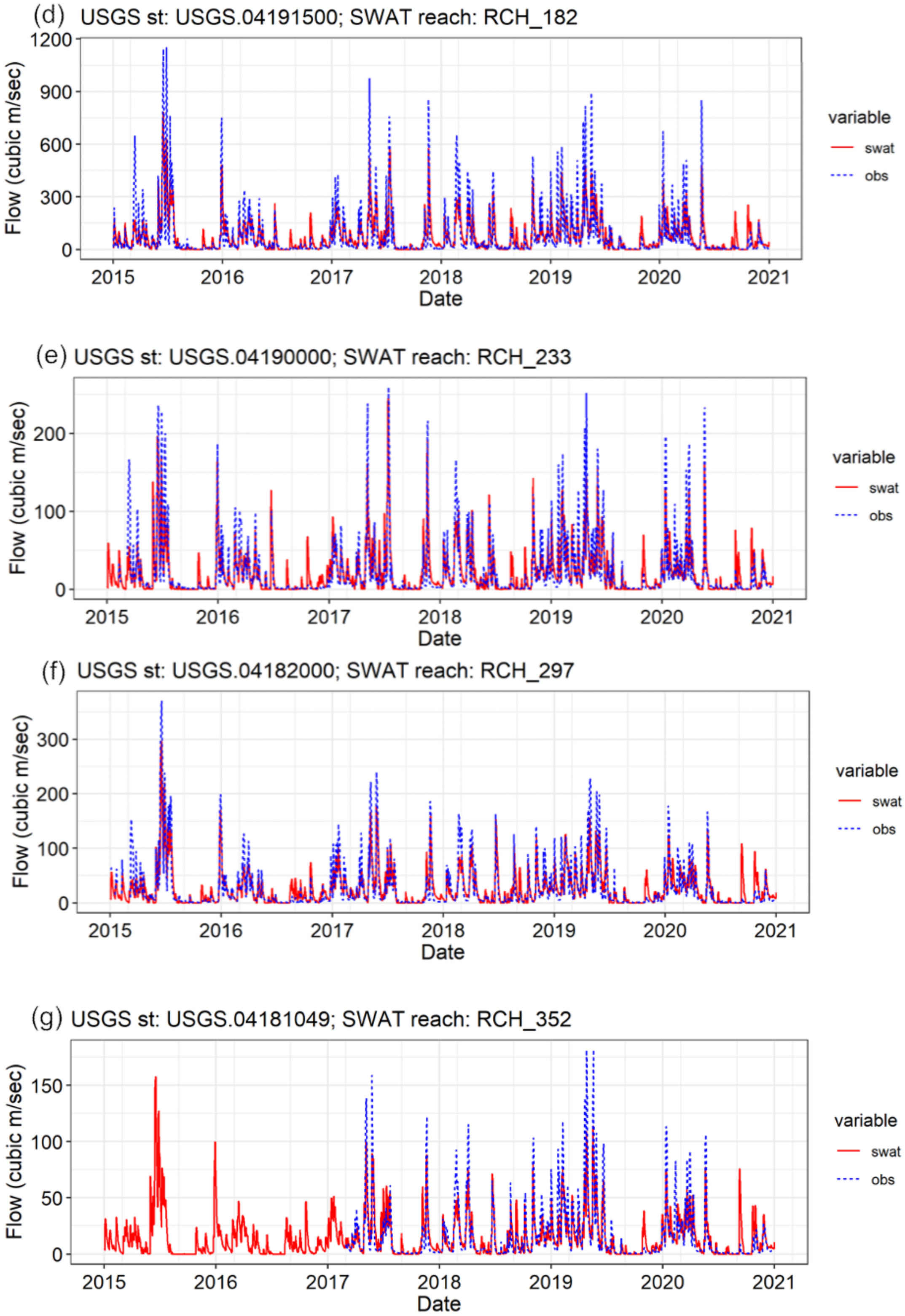

Figure S2. SWAT-simulated vs. USGS observed flow at selected sites (continued).

References

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. and Kløve, B. (2015). A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology*, *524*, 733-752. doi:10.1016/j.jhydrol.2015.03.027

Gupta, H. V., Sorooshian, S. and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resources Research*, *34*(4), 751-763. doi:10.1029/97WR03495

Han, W., Yang, Z., Di, L. and Mueller, R. (2012). CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, *84*, 111-123. doi:10.1016/j.compag.2012.03.005

Kling, H., Fuchs, M. and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424-425*, 264-277. doi:10.1016/j.jhydrol.2012.01.011

Lin, F. R., Wu, N. J., Tu, C. H. and Tsay, T. K. (2017). Automatic calibration of an unsteady river flow model by using dynamically dimensioned search algorithm. *Mathematical Problems in Engineering*, *2017*(7), 1-19. doi:10.1155/2017/7919324

Matott, L. S. (2017). *OSTRICH: an optimization software tool, documentation and user's guide, version 17.12.19*. University at Buffalo Center for Computational Research. http://www.civil.uwaterloo.ca/envmodelling/Ostrich.html.

PRISM Climate Group. (2014). PRISM Gridded Climate Data. Oregon State University. https://prism.oregonstate.edu

Soil Survey Staff. (2015). Soil Survey Geographic (SSURGO) Database for Maumee River Basin. Natural Resources Conservation Service, United States Department of Agriculture. https://data.nal.usda.gov/dataset/soil-survey-geographic-database-ssurgo

Tolson, B. A. and Shoemaker, C. A. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, *43*(1). doi:10.1029/2005WR004723

Zambrano-Bigiarini, M. and Rojas, R. (2013). A model-independent Particle Swarm Optimisation software for model calibration. *Environmental Modelling & Software*, *43*, 5-25. doi:10.1016/j.envsoft.2013.01.004