



# 1 SWAT Modeling of Water Quantity and Quality in the Tennessee River Basin:

# 2 Spatiotemporal Calibration and Validation

- 3 Gangsheng Wang<sup>1,2,\*</sup>, Henriette I. Jager<sup>1,2</sup>, Latha M. Baskaran<sup>1</sup>, Tyler F. Baker<sup>3</sup>, Craig C.
- 4 Brandt<sup>4</sup>
- 5
- 6 <sup>1</sup>Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA
- 7 <sup>2</sup>Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA
- 8 <sup>3</sup>Tennessee Valley Authority, Knoxville, TN 37902 USA
- 9 <sup>4</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831 USA
- 10

## 11 \*Corresponding Author: Gangsheng Wang

- 12 Bldg 4500N, Room F129S, MS-6301
- 13 Oak Ridge National Laboratory
- 14 Oak Ridge, TN 37831-6301
- 15 <u>wangg@ornl.gov</u>
- 16
- 17 Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-
- 18 00OR22725 with the US Department of Energy. The United States Government retains and the
- 19 publisher, by accepting the article for publication, acknowledges that the United States
- 20 Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or
- 21 reproduce the published form of this manuscript, or allow others to do so, for United States
- 22 Government purposes. The Department of Energy will provide public access to these results of
- 23 federally sponsored research in accordance with the DOE Public Access Plan
- 24 (http://energy.gov/downloads/doe-public-access-plan).





#### 25 Abstract

26 Model-data comparisons are always challenging, especially when working at a large spatial 27 scale and evaluating multiple response variables. We implemented the Soil and Water Assessment Tool (SWAT) to simulate water quantity and quality for the Tennessee River Basin. 28 29 We developed three innovations to overcome hurdles associated with limited data for model 30 evaluation: 1) we implemented an auto-calibration approach to allow simultaneous calibration 31 against multiple responses, including intermediate response variables, 2) we identified empirical 32 spatiotemporal datasets to use in our comparison, and 3) we compared functional patterns in 33 landuse-nutrient relationships between SWAT and empirical data. Comparing monthly SWAT-34 simulated runoff against USGS data produced satisfactory median Nash-Sutcliffe Efficiencies of 35 0.83 and 0.72 for calibration and validation periods, respectively. SWAT-simulated water quality responses (sediment, TP, TN, and inorganic N) reproduced the seasonal patterns found in 36 37 LOADEST data. SWAT-simulated spatial TN loadings were significantly correlated with 38 empirical SPARROW estimates. The spatial correlation analyses indicated that SWAT-modeled 39 runoff was primarily controlled by precipitation; sedimentation was controlled by topography; 40 and NO<sub>3</sub> and soluble P were highly influenced by land management, particularly the proportion 41 of agricultural lands in a subbasin.

42

43 Keywords: model calibration, validation, reservoir, runoff, SWAT, Tennessee River, water
44 quality





#### 45 **1 Introduction**

46 The Energy Independence and Security Act (EISA) of 2007 set a target for US production of 47 over 36 billion gallons of renewable fuels annually by 2022 (EISA, 2007). Because agricultural 48 development has historically been associated with impacts on water quality (Dodds and Oakes, 49 2008), converting the lands needed to meet EISA targets heightened concerns for the nation's 50 rivers and lakes, as well as for downstream estuaries. The health of waters in the Tennessee 51 River Basin (TRB) is of particular interest because this region supports one of the most 52 biologically diverse river fauna in North America (Haag and Williams, 2014; Keck et al., 2014). 53 Previous studies have shown higher nutrient and sediment loadings in non-forested, human 54 influenced watersheds in the TRB (Scott et al., 2002). Evaluating changes in water quality 55 associated with large-scale regional shifts in land-use and management requires process-based modeling of hydrology and nutrient dynamics (Wellen et al., 2015). Process-based models are 56 57 favored whenever projections beyond historical conditions are needed because these models 58 incorporate the processes leading to change and do not require extrapolation of statistical relationships beyond the range represented in the data. 59

60 Process-oriented models like the Soil & Water Assessment Tool (SWAT) (Arnold and Fohrer, 61 2005; Srinivasan et al., 1998) incorporate current understanding of linkages between watershed 62 properties and water quality responses, but they are also difficult to calibrate (Wang and Chen, 63 2012). Although evaluation of multiple responses simulated by spatially-distributed processbased models over time and space is strongly encouraged (Cao et al., 2006; Wellen et al., 2015), 64 65 such comprehensive evaluations are limited by the availability of spatial and long-term temporal 66 data. This challenge is compounded for models applied at a regional scale because monitoring 67 efforts tend to be local in scale and of short duration, especially for water quality (Hoos and





68 McMahon, 2009). As such, we see a role for empirical models in the calibration and validation

69 of regional-scale models.

70 Empirical models have previously been fitted to spatial and temporal nutrient loads in the US 71 (Saad et al., 2011). The monthly instream nutrient fluxes were estimated using LOADEST 72 (LOAD ESTimator) developed by the United States Geological Survey (USGS) (Runkel et al., 73 2004; USGS, 2015). LOADEST assists in developing regression models for estimating nutrient 74 loads or fluxes over a user-defined time interval based on functions of streamflow, time, and 75 additional user-specified variables (Runkel et al., 2004). The SPARROW (SPAtially Reference 76 Regressions On Watershed attributes) is also a model developed by USGS that relates water 77 quality measurements to characteristics of watersheds (Hoos and McMahon, 2009; Saad et al., 78 2011) to estimate nutrient loads/fluxes. Both models represent empirical relationships most 79 important during the historical period and smooth out the noise inherent in fine-resolution 80 temporal water-quality measurements.

81 When seeking regional surveys suitable for calibration, data may not be available for exactly 82 the outputs produced by the model. However, flexibility in assimilating data can be achieved by 83 comparing against intermediate or synthetic response variables. Several SWAT calibration tools 84 are available, e.g., SWAT-CUP 2012 (Abbaspour, 2014), the Auto-Calibration tool (Van 85 Griensven, 2005), and the R-SWAT-FME framework (Wu and Liu, 2014); however, these tools 86 do not include intermediate or synthetic response variables to compare against. This limitation 87 prevented us from calibrating SWAT using the final water quantity and quality responses if the 88 corresponding observations (or datasets) were not available.

This paper presents solutions to the aforementioned challenges, including fitting regionalscale SWAT model when there is limited spatial and long-term water quality data available and





91 representation of reservoirs in a highly regulated watershed. We describe efforts to implement 92 SWAT modeling of water quantity and quality for the TRB including the configuration of 22 93 reservoirs. We incorporate the Shuffled Complex Evolution algorithm (Duan et al., 1992) into 94 SWAT2012 to enable auto-calibration of the model against multiple hydrologic (i.e., water 95 quantity) and water quality response variables (including intermediate and synthetic response 96 variables) at multiple sites. We calibrate and validate SWAT water quantity and quality against 97 empirically modeled datasets available from the USGS throughout the conterminous US. We 98 also used functional validation to compare primary drivers controlling runoff and water quality 99 in the process-based SWAT model and the empirical models. Functional validation goes beyond 100 adding a stamp of approval (i.e., validation), instead comparing relationships to understand 101 differences and guide future modeling or data collection efforts. The approach described here can 102 be applied in other regions of the US where the required empirical models have been developed.

103

## 104 **2 Materials and Methods**

### 105 **2.1 Study Area**

106 The Tennessee River Basin (TRB), a tributary basin of the Mississippi River Basin, is located 107 in the southeastern part of the United States (USGS, 2014b) (Fig. 1). There are significant 108 physiographic differences in the eastern and western portions of the basin (Price and Leigh, 109 2006). Forest cover is the dominant natural vegetation in the basin. In the western portion, 110 alluvial plains produced rich soils. The middle of the basin, which was historically covered by 111 bottomland forest and prairie, now supports high percentages of pasture and cropland. Eastward, 112 the geology becomes more mountainous and dominated by limestone with sandstone ridges. The 113 easternmost portion of the basin lies in the rugged Blue Ridge and Southern Appalachian





114 provinces with relatively poor soils (Price and Leigh, 2006). The TRB area has a subtropical 115 climate (warm, humid summers, mild winters) (Sagona, 2003). December through early May is 116 the major flood season (TVA, 2014). Since the 1930's, the TRB has been impounded by a series of dams (reservoirs), most of which are managed by the Tennessee Valley Authority (TVA). 117 Main-stem Tennessee River dams are operated in "run-of-river" mode to support river navigation 118 119 and generate hydroelectric power. Dams on the tributaries function as storage impoundments and 120 are used primarily for flood control (TVA, 2014). Kentucky Dam is 35 km (22 mi) upstream 121 from Paducah, Kentucky, where the Tennessee River flows northwest into the Ohio River (Fig. 122 1).

#### 123 2.2 Watershed Delineation and Definition of SWAT Hydrologic Response Units

124 SWAT (Version 2012/Revision 627) was used to model water quantity and water quality 125 (Arnold et al., 2012). The Digital Elevation Model (DEM) data (1-arc-second, c.a. 30 m) for TRB 126 was downloaded from the National Elevation Dataset 127 (http://nationalmap.gov/elevation.html). We conducted watershed delineation in ArcSWAT 128 (Winchell et al., 2013) based on (i) USGS-defined 8-digit Hydrologic Unit Codes (Jager et al., 129 2015) (HUC8, Fig. 1), and (ii) major stream gages and reservoirs (Fig. 1). Watershed delineation of the TRB using the DEM resulted in a drainage area of 106,124 km<sup>2</sup>. Twenty-two (22) 130 131 reservoirs were included in the SWAT setup (Fig. 1). SWAT includes a reservoir module that 132 can represent these waterbodies in the watershed (Chen et al., 2015; Wang and Xia, 2010). The 133 reservoir outflow may be calculated by one of the four methods provided by SWAT (Arnold et 134 al., 2012): (i) average annual release rate for uncontrolled reservoir; (ii) measured monthly 135 outflow; (iii) simulated controlled outflow with target release; and (iv) measured daily outflow.





- 136 The last method (i.e., IRESCO = 3, measured daily outflow) was adopted in this study and TVA
- 137 provided daily reservoir outflow rates from 1985 to 2013.
- 138 Hydrological Response Units (HRUs) represent unique combinations of soil type, slope, and
- 139 land use or land cover. STATSGO soil map units (Soil Survey Staff, 1994) that comprised more
- 140 than 10% of a subbasin were retained. We discretized slope into four categories: <1%, 1–2%, 2–
- 141 5%, and >5%. We used the 2009 Cropland Data Layer (CDL-2009) (USDA-NASS, 2014) to
- 142 represent land use/land cover (Jager et al., 2015). Natural vegetation in TRB is dominated by
- 143 forest (59.4%) and grassland (11.7%). The major crops in TRB are hay (non-Alfalfa, 9.7%),
- soybeans (1.7%), and corn (1.5%). We retained land-use classes that comprised more than 2%
- area of a subbasin. This protocol created a total of 4,026 distinct HRUs in 55 subbasins.

### 146 2.3 Meteorological Forcings

We downloaded historical meteorological observation from DAYMET (Thornton et al., 1997) estimated for the center of each HUC8 (Fig. 1) over the period 1980–2014 (35 years). Daily meteorological variables include total precipitation (mm), maximum and minimum temperatures (°C), and solar radiation (MJ m<sup>-2</sup> d<sup>-1</sup>). Two additional variables (wind speed and relative humidity) were estimated by the SWAT model's climate generator (Gassman et al., 2007). The mean annual precipitation (MAP) on HUC8 units ranged from 1129 to 1715 mm with an average of 1433 mm during 1980-2014.

### 154 **2.4 Model Calibration**

Existing auto-calibration routines in the SWAT model are not designed to calibrate against intermediate response variables (e.g., HUC8 runoff and  $NO_3+NO_2$ ). For this effort, we incorporated the Shuffled Complex Evolution (SCE) algorithm (Duan et al., 1992) into the source code of SWAT2012 model to implement auto-calibration (Fig. 2). SCE is a stochastic





159 optimization algorithm that has been widely used in calibration of hydrological models including 160 SWAT (Wang and Xia, 2010; Zhang et al., 2009). We calibrated 39 parameters (Table 1) 161 governing the hydrologic (i.e., water quantity) and water quality processes in SWAT. The 39 162 parameters were selected based on the sensitivity analyses in previous studies (Abbaspour et al., 2007; Baskaran et al., 2010; Bekele and Nicklow, 2007; Santhi et al., 2001; Wang et al., 2014; 163 164 Wu and Liu, 2012). Generally, these parameters were calibrated step by step. The hydrologic 165 parameters (No. 1-14) were first calibrated against hydrologic response variables (i.e., water 166 quantity variables, e.g., streamflow or runoff). The second step was to calibrate the water quality 167 parameters (No. 15-39) using water quality measurements (e.g., sediment, nitrogen, and/or 168 phosphorus), where a subset of the parameters might be calibrated depending on the response 169 variables as described below.

Fifteen types of calibrations with regard to various response variables (See Supplement Table S1) were defined in our current auto-calibration tool. The first five types correspond to five hydrologic response variables: daily streamflow, monthly streamflow, daily reservoir storage, daily soil water content, and monthly runoff on subbasin or HUC8; the next five types include monthly nutrient (sediment, nitrogen, phosphorus) fluxes (metric tons per month); and the last five types refer to instream monthly nutrient concentration (mg/L). Other response variables could also be defined and added to this calibration framework.

177 Criteria used to assess model performance include:

178 (1) Nash-Sutcliffe Efficiency (NSE, Eq. 1):

179 
$$NSE = 1 - \frac{\sum_{i=1}^{n} \left[ Y_{sim}(i) - Y_{obs}(i) \right]^{2}}{\sum_{i=1}^{n} \left[ Y_{obs}(i) - \overline{Y}_{obs} \right]^{2}}$$
(1)





- 180 where  $Y_{obs}$  and  $Y_{sim}$  are the observed and simulated data, respectively;  $\overline{Y}_{obs}$  is the mean value of
- 181 observations; and *n* and *i* denote the number of data points and the *i*th data, respectively. *NSE* is
- 182 less than or equal to 1 and may be negative (Moriasi et al., 2007).
- 183 (2) Percent Bias (PBIAS, Eq. 2):

184 
$$PBIAS = \frac{\overline{Y}_{obs} - \overline{Y}_{sim}}{\overline{Y}_{obs}} \times 100\%$$
(2)

185 where PBIAS (Moriasi et al., 2007) denotes the deviation of predicted mean value ( $\overline{Y}_{sim}$ ) from

186 observed mean value ( $\overline{Y}_{obs}$ ) as a percentage of  $\overline{Y}_{obs}$ .

According to Moriasi et al. (2007), model simulation is satisfactory if NSE > 0.5 and if  $|PBIAS| \le 25\%$  for streamflow (runoff),  $|PBIAS| \le 55\%$  for sediment, and  $|PBIAS| \le 70\%$  for nitrogen (N) and phosphorus (P).

190 The overall objective function is the weighted average of individual objective functions:

191 
$$F = \sum_{j=1}^{k} (w_j \cdot f_j)$$
 (3)

where *F* denotes the overall objective function of *k* individual objective functions;  $f_j$  is the *j*th objective function that could be calculated as *NSE* or *|PBIAS*| of interested response variable; and  $w_i$  is the weighting factor for each  $f_i$ .

## 195 **2.5 Calibration and Validation of Runoff**

Because streamflow (discharge) at a station within the TRB is largely a measure of the outflow from the upstream reservoir(s) and because observed reservoir outflow was used in this study, we calibrated hydrologic parameters based on runoff (i.e., total water yield) instead of streamflow. We used the USGS computed monthly runoff (1985–1995) in HUC8(USGS, 2014a)





200 as reference for SWAT calibration, with one year (1985) for model spin-up and 10 years (1986– 201 1995) for model calibration. Another 18 years (1996-2013) of data were used for model 202 validation. The USGS HUC8 runoff estimates were generated by combining historical flow data 203 at USGS stream gages and the corresponding drainage basin boundaries and hydrologic units 204 boundaries.(USGS, 2014a) In a previous study, this dataset was used to calibrate the Variable 205 Infiltration Capacity (VIC) model for the conterminous US (Oubeidillah et al., 2014). The objective of our multi-site calibration of runoff was to calibrate hydrologic parameters (No. 1-14) 206 207 of the subbasins within each HUC8. For example, when we implemented calibration in terms of the HUC8-06010102, the parameters in four subbasins (1, 4, 5 and 6) in this HUC8 were 208 209 calibrated. We calculated simulated HUC8 runoff as the area-weighted-average of runoff from 210 subbasins within each HUC8:

211 
$$R_{HUC8} = \sum_{j=1}^{m} \left[ R_{sub}(j) \times Area_{sub}(j) / Area_{HUC8} \right]$$
(4)

212 
$$Area_{HUC8} = \sum_{j=1}^{m} \left[ Area_{sub}(j) \right]$$
(5)

where  $R_{HUC8}$  is the runoff in a HUC8;  $R_{sub}(j)$  and  $Area_{sub}(j)$  are the simulated runoff (mm) and the area (km<sup>2</sup>) in the *j*th subbasin; and  $Area_{HUC8}$  is the total area of the HUC8 that includes *m* subbasins. The *NSE* of monthly HUC8 runoff was defined as the objective function in hydrologic calibration.

## 217 2.6 Calibration and Validation of Monthly Nutrient Fluxes

218 Nutrient measurements are sparse in rivers of the TRB. We have attempted to collect in-situ 219 water quality monitoring data from over 6,000 USGS and EPA (Environmental Protection 220 Agency) stations within the TRB through the National Water Quality Monitoring Council





221 (NWQMC)'s online Water Quality Portal (WQP) (NWQMC, 2015). However, we did not find 222 long-term water quality data that coincided with our model simulation period (i.e., after 1980's). 223 Therefore, we used the LOADEST (LOAD ESTimator) dataset (Runkel et al., 2004) as reference to calibrate water quality parameters. The LOADEST dataset provided estimates of monthly 224 225 nutrient fluxes (1996–2006) at the Tennessee River near Paducah, KY (i.e., the outlet of TRB) 226 (USGS, 2015). We used three-year (1994–1996) of data for model spin-up and 10 years (1997– 2006) for model calibration. Another seven years (2007-2013) of data were used for model 227 228 validation. Four water quality variables were available from the LOADEST dataset: sediment, 229 total phosphorus (TP), total nitrogen (TN), and  $NO_3+NO_2$ . The hydrologic parameters (No. 1–14) 230 calibrated against runoff were fixed during the calibration of water quality parameters (No. 15-39). The NSE of monthly water quality was defined as the objective function during SWAT 231 calibration. When multiple response variables (e.g., Sediment, TP, TN, NO<sub>3</sub>+NO<sub>2</sub>) were 232 233 considered in model calibration, we used Eq. (4) to calculate the overall objective function. In 234 addition, the spatial distribution of mean annual nutrient loadings estimated by the SPARROW model (Hoos and McMahon, 2009) was employed as another dataset for model validation at the 235 236 HUC8 level. The mean annual loads (MAL) of nutrients at the HUC8 level were calculated as 237 the area-weighted average of the MALs at all subbasins within the HUC8.

### 238 2.7 Spatial Correlation Analyses

Understanding how water yield and nutrient loadings vary with watershed characteristics is important for quantifying primary drivers controlling water quantity and quality and for developing nutrient management policies (Hoos and McMahon, 2009). To this end, we implemented spatial correlation analyses between response variables and watershed attributes. For this study, two variables were considered highly correlated if the absolute value of





correlation coefficient (|r|) was greater than 0.6 and the correlation was significant (p-value < 0.05), and moderately correlated if |r| was between 0.2 and 0.6 and the correlation was significant.

Based on the 29-year (1985–2013) simulation results from the calibrated SWAT model, we 247 248 calculated the mean annual values of response variables including Runoff, RC (Runoff 249 Coefficient, i.e., the ratio of runoff to precipitation), Sediment, OrgP (organic phosphorus), SolP (soluble P), MinP (mineral P attached to sediment), TP, TN, OrgN (organic N), and NO<sub>3</sub>. We 250 251 first conducted spatial correlation analysis between these response variables at the subbasin level. 252 We implemented spatial correlation analysis between the response variables and the subbasin 253 attributes (explanatory variables): Precipitation (mm), Subbasin Slope (subbasin slope, %), 254 Elevation Drop (difference between highest and lowest elevations, m), and fractions of major 255 land-use types (Forest Fraction, Grassland Fraction, Hay\_Fraction, Crop Fraction, 256 Shrubland Fraction, Wetlands Fraction, Water Fraction, Developed Fraction, see Supplement 257 Fig. S1).

## 258 **3 Results and Discussion**

In the sections below, we describe calibration and validation of different SWAT modelresponses including runoff and water quality metrics.

261 3.1 Runoff

SWAT simulations of TRB runoff were implemented with regard to the period from 1985 to 2013. We divided the 29-year runoff dataset into three sub-datasets: (i) a 1-year spin-up period (1985), (ii) a 10-year calibration period (1986–1995), and (iii) an 18-year validation period (1996–2013). The spatial resolution was the 8-digit hydrologic units (HUC8s) throughout the TRB.





Hydrologic parameters (No. 1–14 in Table 1) were calibrated by comparing simulated monthly HUC8 runoff with the USGS dataset. As an example, Fig. S2 shows the comparison between the SWAT-simulated monthly runoff (i.e., water yield, denoted by 'Sim') and USGS runoff (denoted by 'Obs') in HUC8-06040006, which is the outlet HUC8 of TRB. The *NSE* values for this HUC8 were 0.90 and 0.70 for model calibration and validation, respectively.

272 Values of NSE across the 32 HUC8s (Fig. 3a) ranged from 0.56 to 0.93 with 50% confidence interval (CI) of 0.74–0.88 (median 0.83); the PBIAS values (Fig. 3b) were within a narrow range 273 274 (-7%-13%). The model also performed well over the validation period, although NSE was lower 275 than that during the calibration period, as one would expect. The median NSE was 0.72 with 50% 276 CI of 0.57–0.77; and the *PBIAS* values were within the satisfactory range, i.e.,  $\pm 25\%$  (Moriasi et 277 al., 2007) except for two HUC8s (06010108 and 06010204). Regarding the whole dataset for the 278 combined calibration and validation periods (1986–2013), the median NSE was 0.79 (50% CI: 279 0.69-0.84) and all of the *PBIAS* values were within  $\pm 25\%$  except for one HUC8 with a 280 marginally satisfactory PBIAS (-26%).

The SWAT-simulated mean annual runoff (MAR) in the two aforementioned HUC8s (06010108 and 06010204) might be more reasonable than the USGS-estimated MAR. We analyzed the mean annual precipitation (MAP) and MAR data from 1986–2013 and found that the runoff in these two HUC8s might be underestimated in the USGS dataset to some degree (See Fig. S3).

286 **3.2 Water Quality** 

The SWAT simulation of water quality began with the year 1996 owing to data availability. The 20-year (1994–2013) water quality dataset was also divided into three sub-datasets: (i) a 3-year





spin-up period (1994–1996), (ii) a 10-year calibration period (1997–2006), and (iii) a 7-year

290 validation period (2007–2013).

291 The water quality parameters (No. 15-39 in Table 1) were calibrated against the LOADEST dataset by taking into account multiple objectives, i.e., four response variables, including 292 293 sediment, TP, TN, and NO<sub>3</sub>+NO<sub>2</sub>. Calibration greatly improved the performance of the model, particularly for sediment (NSE = -100 and 0.06 for pre- and post-calibration, respectively), TP 294 (NSE = -2.5 and 0.44 for pre- and post-calibration, respectively), and TN (NSE = 0.02 and 0.38 m)295 296 for pre- and post-calibration, respectively) (See Supplement Table S2). The NSE values for 297 model validation were not as good as the NSE for calibration, but the PBIAS values (Table S2) 298 were satisfactory except for NO<sub>3</sub>+NO<sub>2</sub> (-157%). The squared correlation coefficients ( $r^2$ ) for TN 299 and TP during both calibration and validation periods equaled or exceeded 0.4 whereas the  $r^2$ 300 values for sediment and inorganic N were less than 0.4 (Table S2). SWAT-simulated water 301 quality responses reproduced the seasonal patterns found in LOADEST data during both 302 calibration and validation periods (See Fig. S4).

303 We further conducted the water quality simulation for a longer period of time (1985–2013) 304 than the period for model calibration and calibration (1997–2013). The spatial distributions of 305 SWAT-simulated MALs (1986-2013) of TN and TP were comparable to the SPARROW 306 estimates. The spatial patterns of SWAT-simulated TN and TP at the subbasin level are shown in 307 Fig. 4 and other variables (runoff, RC, sediment, and NO<sub>3</sub>) are shown in Fig. S5. The spatial MALs of TN and TP from SWAT were compared with the SPARROW dataset (MALs from 308 309 1975–2004) at the HUC8 level (Fig. 5). The PBIAS values (between SWAT and SPARROW) for 310 TN (Fig. 5a) at 26 out of 32 HUC8 units were within the range of  $\pm 70\%$ , and the *PBIAS* values at 311 three HUC8 were higher than 80%. The 50% CIs of MAL of TN were 2.5-6.7 kg N/ha and 4.7-





- 312 7.4 kg N/ha by SWAT and SPARROW, respectively. The SWAT-simulated MAL of TN across
- 313 the 32 HUC8 units was 5.5 kg N/ha, which was 12% lower than the TN loading (6.2 kg N/ha)
- 314 estimated by SPARROW.
- As for phosphorus (Fig. 5b), the SWAT-simulated MAL of OrgP+SolP (organic P + soluble P) was 48% lower than the SPARROW-modeled TP, while the SWAT-simulated MAL of TP (organic P + soluble P + mineral P) was 50% higher than the SPARROW-modeled TP. This was because mineral P contributed most (75.2%) to the TP yield and organic P contributed least (8.5%) to TP. The SWAT-simulated MAL of OrgP+MinP (organic P + mineral P, 0.93 kg P/ha) was comparable to the SPARROW-estimated TP (0.88 kg P/ha).
- The spatial patterns of TN from the two models (SWAT and SPARROW) were significantly correlated with each other (r = 0.54, p-value < 0.001). The spatial pattern of SPARROWestimated TP was not significantly correlated with SWAT-simulated TP, but moderately correlated with SWAT-simulated OrgP+SolP (r = 0.38, p-value = 0.03) and highly correlated with SWAT-simulated TN (r = 0.84, p-value < 0.001).

326 Different from the multi-site hydrologic calibration for each HUC8, water quality was 327 calibrated against data from one site owing to data availability. Notice that this site (Paducah, 328 KY) is located at the outlet of TRB. In addition, 10 out of 25 water quality parameters are basin-329 wide parameters (Table 1, denoted by 'basins.bsn') that are spatially identical in SWAT. 330 Therefore, current water quality calibration could represent the overall water quality regime in the watershed. In summary, the temporal comparison of water quality simulations between 331 332 SWAT and LOADEST and the spatial comparison between SWAT and SPARROW showed a 333 correspondence between process-based SWAT modeling results and those from empirically 334 modeled data in the TRB.





#### 335 **3.3 Spatial Correlation between Response Variables**

336 Functional validation seeks to compare key functional relationships found in process-based 337 models with those in data. This approach goes beyond simple 'validation' or casting a stamp of 338 approval on a model to understand the reasons for any remaining differences. We found that 339 SWAT-simulated MALs of MinP (mineral P attached to sediment) and TP were highly 340 correlated with sediment, which confirms that sediment plays an important role in watershed 341 phosphorus dynamics (Fig 6a). The TN yield was highly correlated with NO<sub>3</sub>. TN loadings were 342 dominated by NO<sub>3</sub>, i.e., the fraction of TN that was NO<sub>3</sub> ranged from 37% to 99% with an average of 80%. TP was not correlated with TN, but OrgP (organic P) was moderately correlated 343 344 with OrgN (organic N) and SolP (soluble P) was moderately correlated with NO<sub>3</sub>, which implies 345 similarity between SolP and NO<sub>3</sub> dynamics and similarity between OrgP and OrgN dynamics in 346 SWAT (Neitsch et al., 2011). The SPARROW-estimated spatial patterns of TN and TP were 347 correlated with each other: however, the SWAT-simulated spatial distributions of TN and TP 348 were decoupled because the MinP component (attached to sediment) in SWAT and TN was 349 dominated by inorganic nitrogen. Nutrient (Sediment, P and N) loadings were not significantly 350 correlated with runoff (Fig. 6a), suggesting that nutrient point-source and non-point sources and 351 other physical landscape variables (Hoos and McMahon, 2009) control variation in nutrient 352 loadings simulated by SWAT.

## 353 3.4 Correlation between SWAT Response Variables and Subbasin Attributes

The spatial correlation analyses showed that the response variables differed in their controlling factors. Runoff was highly correlated with precipitation (r = 0.68) and moderately and positively related to Forest\_Fraction (r = 0.36). The runoff coefficient (RC) was moderately and positively correlated with Elevation Drop (r = 0.32) and Subbasin Slope (r = 0.31) (Fig. 6b).





Sediment loadings were moderately and positively correlated with Elevation Drop (r = 0.47), 358 359 which verifies that the representation of topography and topology in this mountainous region 360 drives sediment dynamics (Wellen et al., 2015). We did not find any significant correlation between TP and the aforementioned subbasin attributes. However, OrgP (organic P) was highly 361 associated with Developed Fraction (r = 0.64) that represented human activities in urban area 362 363 (Hoos and McMahon, 2009); SolP (soluble P) was moderately correlated with Hay Fraction (r =0.43) indicating the influence of agricultural fertilization; and MinP (mineral P) was moderately 364 correlated with Elevation Drop (r = 0.37) that was the primary driver for sediment generation. 365

Organic N (OrgN) was moderately correlated with Wetlands\_Fraction (r = 0.27) and Shrubland\_Fraction (r = -0.30). NO<sub>3</sub> was highly correlated with Hay\_Fraction (r = 0.63) and moderately correlated with Crop\_Fraction (r = 0.48), mostly owing to the response of NO<sub>3</sub> yield to agricultural fertilization. In addition, NO<sub>3</sub> showed a moderate negative correlation with Forest\_Fraction (r = -0.54), Subbasin\_Slope (r = -0.44), and Elevation\_Drop (r = -0.34). Note that TRB subbasins with steeper slopes generally had more forest and less cropland. The primary drivers controlling TN were the same as those for NO<sub>3</sub> as TN was dominated by NO<sub>3</sub>.

#### 373 4 Summary

Model-data comparisons are always challenging, especially when working at a large spatial scale and evaluating multiple response variables. We developed three innovations to overcome hurdles associated with limited data for model testing: 1) we implemented an auto-calibration approach to allow simultaneous calibration against multiple responses, including intermediate response variables, 2) we identified empirical modeled datasets interpolated in space and time to use in our comparison, and 3) we compared functional patterns in landuse-nutrient relationships between SWAT and empirical data. Using these innovations, we were able to successfully





implement a calibrated model for the river basin and to evaluate performance. The SWATcalibration tool developed in this study can be accessed upon request via GitHub

383 (https://github.com/wanggangsheng/SWATopt.git).

In addition to quantitative performance evaluation, we also discerned what the most 384 important influences on SWAT responses were. Runoff was mainly controlled by precipitation; 385 386 runoff coefficient and sedimentation were controlled by topographic attributes; whereas NO<sub>3</sub> and 387 soluble P were highly influenced by land use types, particularly the croplands (hay and other 388 crops). This is likely because our management of these croplands included applying fertilizers 389 containing N and P. Patterns in phosphorus dynamics differed more between the empirical and 390 process-based model than patterns in nitrogen dynamics, suggesting an area for future 391 exploration.

392

## **393 AUTHOR INFORMATION**

### 394 Corresponding Author

- 395 \*Tel: +1 (865)576-6685. E-mail: <u>wangg@ornl.gov</u>.
- 396 Notes
- 397 The authors declare no competing financial interest.
- 398

## 399 Acknowledgements

This material is based upon work supported by the U.S. Department of Energy, Office of Energy
Efficiency & Renewable Energy's Bioenergy Technology Program. Oak Ridge National

- 402 Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract
- 403 No. DE-AC05-00OR22725. We thank Michelle Thornton for providing DAYMET data and





- 404 David Gorelick and Jasmine Kreig for discussion on this study. In addition, we appreciate the
- 405 insightful reviews of Drs. Shih-Chieh Kao and Sujithkumar Surendran Nair.

406

407





420

421

References

## 422 Manual, Swiss Federal Institute of Aquatic Science and Technology (Eawag) Duebendorf, 423 Switzerland, 2014. 424 Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., and 425 Srinivasan, R.: Modelling hydrology and water quality in the pre-alpine/alpine Thur 426 watershed using SWAT, Journal of Hydrology, 333, 413-430, 2007. 427 Arnold, J. G. and Fohrer, N.: SWAT2000: current capabilities and research opportunities in 428 applied watershed modelling, Hydrological Processes, 19, 563-572, 2005. 429 Arnold, J. G., Kiniry, J. R., Srinivasan, R., Williams, J. R., Haney, E. B., and Neitsch, S. L.: Soil 430 and Water Assessment Tool Input/Output Documentation Version 2012, Texas Water 431 Resources Institute, College Station, TXTR-439, 650 pp., 2012. 432 Baskaran, L., Jager, H., Schweizer, P. E., Srinivasan, R., Douglas-Mankin, K., and Arnold, J.: 433 Progress toward evaluating the sustainability of switchgrass as a bioenergy crop using the 434 SWAT model, Transactions of the Asabe, 53, 1547-1556, 2010. Bekele, E. G. and Nicklow, J. W.: Multi-objective automatic calibration of SWAT using NSGA-435

Abbaspour, K. C.: SWAT-CUP 2012: SWAT Calibration and Uncertainty Programs - A User

- 436 II, Journal of Hydrology, 341, 165-176, 2007.
- 437 Cao, W. Z., Bowden, W. B., Davie, T., and Fenemor, A.: Multi-variable and multi-site
- 438 calibration and validation of SWAT in a large mountainous catchment with high spatial
  439 variability, Hydrological Processes, 20, 1057-1073, 2006.
- 440 Chen, Y., Ale, S., Rajan, N., Morgan, C. L. S., and Park, J.: Hydrological responses of land use
- 441 change from cotton (Gossypium hirsutum L.) to cellulosic bioenergy crops in the





- 442 Southern High Plains of Texas, USA, GCB Bioenergy, doi: 10.1111/gcbb.12304, 2015.
- 443 n/a-n/a, 2015.
- 444 Dodds, W. K. and Oakes, R. M.: Headwater influences on downstream water quality,
- Environmental Management, 41, 367-377, 2008.
- 446 Duan, Q. Y., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for
- 447 conceptual rainfall-runoff models, Water Resources Research, 28, 1015-1031, 1992.
- 448 EISA: Energy Independence and Security Act of 2007. 110-140, 2007.
- 449 Gassman, P. W., Reyes, M. R., Green, C. H., and Arnold, J. G.: The soil and water assessment
- 450 tool: Historical development, applications, and future research directions, Transactions of
- 451 the ASABE, 50, 1211-1250, 2007.
- 452 Haag, W. R. and Williams, J. D.: Biodiversity on the brink: an assessment of conservation
- 453 strategies for North American freshwater mussels, Hydrobiologia, 735, 45-60, 2014.
- 454 Hoos, A. B. and McMahon, G.: Spatial analysis of instream nitrogen loads and factors
- 455 controlling nitrogen delivery to streams in the southeastern United States using spatially
- 456 referenced regression on watershed attributes (SPARROW) and regional classification
- 457 frameworks, Hydrological Processes, 23, 2275-2294, 2009.
- 458 Jager, H. I., Baskaran, L. M., Schweizer, P. E., Turhollow, A. F., Brandt, C. C., and Srinivasan,
- R.: Forecasting changes in water quality in rivers associated with growing biofuels in the
  Arkansas-White-Red river drainage, USA, Global Change Biology: Bioenergy, 7, 774-
- 461 784, 2015.
- 462 Keck, B. P., Marion, Z. H., Martin, D. J., Kaufman, J. C., Harden, C. P., Schwartz, J. S., and
- 463 Strange, R. J.: Fish functioanl traits correlated with environmental variables in a
- temperate biodiversity hotspot, PLOS ONE, 9, e93237, 2014.





- 465 Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., and Veith, T.: Model evaluation
- 466 guidelines for systematic quantification of accuracy in watershed simulations,
- 467 Transactions of the Asabe, 50, 885-900, 2007.
- 468 Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R.: Soil and Water Assessment Tool
- 469 Theoretical Documentation. Version 2009, Texas Water Resources Institute, Texas A&M
- 470 University System, College Station, TX, 647 pp., 2011.
- 471 NWQMC: Water Quality Portal. National Water Quality Monitoring Council, 2015.
- 472 Oubeidillah, A. A., Kao, S. C., Ashfaq, M., Naz, B. S., and Tootle, G.: A large-scale, high-
- 473 resolution hydrological model parameter data set for climate change impact assessment
- for the conterminous US, Hydrol. Earth Syst. Sci., 18, 67-84, 2014.
- 475 Price, K. and Leigh, D. S.: Morphological and sedimentological responses of streams to human
- 476 impact in the southern Blue Ridge Mountains, USA, Geomorphology, 78, 142-160, 2006.
- 477 Runkel, R. L., Crawford, C. G., and Cohn, T. A.: Load Estimator (LOADEST): A FORTRAN
- 478 program for estimating constituent loads in streams and rivers, U.S. Geological Survey,
- 479 Reston, Virginia2328-7055, 2004.
- 480 Saad, D. A., Schwarz, G. E., Robertson, D. M., and Booth, N. L.: A Multi-Agency Nutrient
- 481 Dataset Used to Estimate Loads, Improve Monitoring Design, and Calibrate Regional
- 482 Nutrient SPARROW Models, Journal of the American Water Resources Association, 47,
- 483 933-949, 2011.
- 484 Sagona, F. J.: Tennessee River Basin Watershed Management Plan, Tennessee River Clean
  485 Water Partnership, Rainsville, AL, 148 pp., 2003.





- 486 Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R., and Hauck, L. M.:
- 487 Validation of the SWAT model on a large river basin with point and nonpoint sources,
- 488 Journal of the American Water Resources Association, 37, 1169-1188, 2001.
- 489 Scott, M. C., Helfman, G. S., McTammany, M. E., Benfield, E. F., and Bolstad, P. V.: Multiscale
- 490 influences on physical and chemical stream conditions across Blue Ridge landscapes,
- 491 Journal of the American Water Resources Association, 38, 1379-1392, 2002.
- 492 Soil Survey Staff: U.S. General Soil Map (STATSGO), Natural Resources Conservation Service,
- 493 U.S. Department of Agriculture, Washington, D.C., 1994.
- 494 Srinivasan, R., Arnold, J., and Jones, C.: Hydrologic modelling of the United States with the soil
- 495 and water assessment tool, International Journal of Water Resources Development, 14,
- 496 315-325, 1998.
- Thornton, P. E., Running, S. W., and White, M. A.: Generating surfaces of daily meteorological
  variables over large regions of complex terrain, Journal of Hydrology, 190, 214-251,
- 499 1997.
- 500 TVA: http://www.tva.gov, last access: 8/8/2014 2014.
- 501 USDA-NASS: <u>http://www.nass.usda.gov/research/Cropland/SARS1a.htm</u>, last access: 3/16/2014,
  502 2014.
- 503 USGS: http://waterwatch.usgs.gov/index.php?id=romap3, last access: 10/1/2014 2014.
- 504 USGS: http://tn.water.usgs.gov/lten/tenn.html, last access: 4/8/2014 2014.
- 505 USGS: Streamflow and Nutrient Flux at Tennessee River at Highway 60 near Paducah,
- 506 Kentucky. U. S. Geological Survey, 2015.
- 507 Van Griensven, A.: Sensitivity, auto-calibration, uncertainty and model evaluation in
- 508 SWAT2005, UNESCO-IHE, Delft, Ntherlands, 2005.





509

510

511 and Risk Assessment, 28, 225-238, 2014. Wang, G. and Chen, S.: A review on parameterization and uncertainty in modeling greenhouse 512 513 gas emissions from soil, Geoderma, 170, 206-216, 2012. 514 Wang, G. and Xia, J.: Improvement of SWAT2000 modelling to assess the impact of dams and 515 sluices on streamflow in the Huai River basin of China, Hydrological Processes, 24, 516 1455-1471, 2010. 517 Wellen, C., Kamran-Disfani, A.-R., and Arhonditsis, G. B.: Evaluation of the Current State of 518 Distributed Watershed Nutrient Water Quality Modeling, Environmental science & 519 technology, 49, 3278-3290, 2015. Winchell, M., Srinivasan, R., Di Luzio, M., and Arnold, J. G.: ArcSWAT Interface for 520 521 SWAT2012 User's Guide, Blackland Research and Extension Center, Temple, TXTR-522 439, 464 pp., 2013. Wu, Y. and Liu, S.: Improvement of the R-SWAT-FME framework to support multiple variables 523

Wang, G., Barber, M. E., Chen, S., and Wu, J. Q.: SWAT modeling with uncertainty and cluster

analyses of tillage impacts on hydrological processes, Stochastic Environmental Research

- 524 and multi-objective functions, Science of the Total Environment, 466, 455-466, 2014.
- 525 Wu, Y. and Liu, S.: Modeling of land use and reservoir effects on nonpoint source pollution in a
- 526 highly agricultural basin, Journal of Environmental Monitoring, 14, 2350-2361, 2012.
- 527 Zhang, X., Srinivasan, R., Zhao, K., and Liew, M. V.: Evaluation of global optimization
- 528 algorithms for parameter calibration of a computationally intensive hydrologic model,
- 529 Hydrological Processes, 23, 430-441, 2009.
- 530
- 531





## Tables

## Table 1. Selected SWAT parameters for model calibration

No	Parameter <sup>a</sup>	Description	Default	Min	Max	Input file	Fortran code
1	CN2	Initial SCS curve number II	85	35	98	*.mgt	Readmgt.f
2	ESCO	Soil evaporation compensation factor	0.95	0.01	1	*.hru	Readhru.f
3	EPCO	Plant uptake compensation factor	1	0.01	1	*.hru	Readhru.f
4	OV N	Manning's n value for overland flow	0.1	0.01	0.6	*.hru	Readhru.f
5	CH N2	Manning's n value for main channel	0.014	0.01	0.5	*.rte	Readrte.f
6	CH K2	Channel effective hydraulic conductivity	0.001	0.001	150	*.rte	Readrte.f
		(mm/hr)					
7	ALPHA BF	Baseflow alpha factor (days)	0.048	0.001	1	*.gw	Readgw.f
8	GW DELAY	Ground water delay (days)	31	0.0001	500	*.gw	Readgw.f
9	RCHRG DP	Deep aquifer percolation fraction	0.05	0.0001	1	*.gw	Readgw.f
10	GW REVAP	Groundwater revap coefficient	0.02	0.02	0.2	*.gw	Readgw.f
11	GW SPYLD	Specific yield for shallow aquifer (m <sup>3</sup> /m <sup>3</sup> )	0.003	0.0001	0.4	*.gw	Readgw.f
12	SOL AWC	Available water capacity (mm H <sub>2</sub> O/mm soil)	0.2	0.01	0.4	*.sol	Readsol.f
13	SOL K	Saturated hydraulic conductivity (mm/h)	10	0.01	100	*.sol	Readsol.f
14	SURLAG	Suface runoff lag coefficient (days)	4	0.5	12	sub.lag	Readhru.f
15	SPCON	Linear re-entrainment parameter	0.0001	0.0001	0.01	basins.bsn	Readbsn.f
16	SPEXP	Exponent re-entrainment parameter	1	1	2	basins.bsn	Readbsn.f
17	PRF	Adjustment factor for sediment routing in the	1	0.001	2	basins.bsn	Readbsn.f
		main channel					
18	ADJ PKR	Adjustment factor for sediment routing in	1	0.5	2	basins.bsn	Readbsn.f
	-	tributary channels					
19	CH COV	Channel cover factor	0.001	0.001	1	*.rte	Readrte.f
20	CH EROD	Channel erodibility factor	0.001	0.001	1	*.rte	Readrte.f
21	USLE K	Soil erodability factor	0.28	0.01	0.65	*.sol	readsol.f
22	BIOMIX	Biological mixing coefficiency	0.2	0.01	1	*.mgt	Readmgt.f
23	RSDCO	Residue decomposition factor	0.05	0.02	0.1	basins.bsn	Readbsn.f
24	NPERCO	Nitrogen percolation factor	0.2	0.001	1	basins bsn	Readbsn f
25	N UPDIS	N uptake distribution parameter	20	0.001	100	basins bsn	Readbsn f
26	NSETLR	N settling rate in reservoir (m/yr). Line 7 & 8	5.5	1	15	* lwa	Readlwg f
27	SHALLST N	Concentration of $NO_3$ in groundwater (mg N/L)	0.0001	0.0001	1000	* gw	Readgy f
28	ERORGN	Organic N enrichment for sediment	0.001	0.001	5	* hru	Readhru f
29	SOL ORGN	Initial organic N concentration (mg N kg <sup><math>-1</math></sup> soil)	0.01	0.01	50	*.chm	Readchm f
30	SOL NO3	Initial NO <sub>3</sub> concentration in the soil layer (mg	0.01	0.01	50	*.chm	Readchm f
		N kg <sup>-1</sup> soil)					
31	PPERCO	Phosphorus percolation factor (10 m <sup>3</sup> Mg <sup>-1</sup> )	10	10	17.5	basins bsn	Readbsn f
32	PHOSKD	Phosphorus soil partitioning coefficient (m <sup>3</sup>	175	100	200	basins bsn	Readbsn f
52	THOULD	$Mg^{-1}$ )	170	100	200	0401110.0001	readomin
33	PSP	P sorption coefficient	0.4	0.01	0.7	basins.bsn	Readbsn.f
34	PSETLR	P settling rate in reservoir (m/yr), Line 5 & 6	10	2	20	*.lwq	Readlwq.f
35	BC4	Rate const for mineralization of organic P to	0.35	0.01	0.7	*.swq	Readswq.f
		dissolved P (1/d)					
36	RS5	Organic P settling rate (1/d)	0.05	0.001	0.1	*.swq	Readswq.f
37	ERORGP	Organic P enrichment ratio with sediment	0.001	0.001	5	*.hru	readhru.f
		loading					
38	SOL_ORGP	Initial organic P (mg P kg <sup>-1</sup> soil)	0.01	0.01	50	*.chm	Readchm.f
39	SOL_SOLP	Initial soluble P concentration in the soil layer	5	0.01	50	*.chm	Readchm.f
		(mg P kg <sup>-1</sup> soil)					

<sup>a</sup>Four groups of parameters: No. 1–14: Water quantity; No. 15–21: Sediment; No. 22–30: Nitrogen; No. 31–39:

Phosphorus.





## **Figure Captions**

Figure 1. Fifty-five subbasins and 22 reservoirs of the Tennessee River Basin (TRB) in the Soil and Water Assessment Tool (SWAT). The mainstem Tennessee River runs from east to west, exiting the basin below Kentucky Dam.

Figure 2. Integrating the Shuffled Complex Evolution (SCE) algorithm into the Soil and Water Assessment Tool (SWAT) permitted calibration against intermediate response variables.

Figure 3. Model calibration of SWAT-modeled runoff by optimizing hydrologic parameters and validation. The distribution shows values for 32 HUC8 units (8-digit Hydrologic Unit Codes). Measures of model performance are (a) Nash-Sutcliffe Efficiency (NSE), (b) Percent Bias (PBIAS).

Figure 4. Spatial distribution of SWAT-simulated mean annual values at 55 subbasins: (a) TN yield (kg N/ha), (b) TP yield (kg P/ha).

Figure 5. Comparison of spatial distribution of TN and TP yield between SWAT simulation and SPARROW dataset at 32 HUC8 units. SWAT metrics:  $OrgP\_SolP = OrgP$  (organic P) + SolP (soluble P);  $OrgP\_MinP = OrgP + MinP$  (mineral P attached to sediment);  $SolP\_MinP = SolP + MinP$ ; and TP = OrgP + SolP + MinP.

Figure 6. Spatial correlation analysis (a) between response variables (mean values from 1996 to 2013), (b) between response variables and subbasin attributes. Larger circle denotes higher





correlation coefficient and only significant correlations (p-value < 0.05) are shown. Numbers in (a) denote correlation coefficients. Response variables are: sediment yield (kg TSS/ha), organic phosphorus yield (OrgP, kg P/ha), soluble P yield (SolP, kg P/ha), mineral P yield attached to sediment (MinP, kg P/ha), total P yield (TP = OrgP + SolP + MinP, kg P/ha), total nitrogen yield (TN, kg N/ha), organic N yield (OrgN, kg N/ha), nitrate yield (NO3, kg N/ha), runoff depth (mm), and runoff coefficient (RC, ratio of runoff to precipitation).





## Figures

























# Figure 5



32







