JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

Vol. 55, No. 5 AMERICAN WATER RESOURCES ASSOCIATION October 2019

Modeling Flow, Nutrient, and Sediment Delivery from a Large International Watershed Using a Field-Scale SWAT Model

Awoke Dagnew, Donald Scavia, Yu-Chen Wang, Rebecca Muenich, Colleen Long, and Margaret Kalcic

Research Impact Statement: A well-calibrated and validated field-scale flow and water quality model was used to assess nutrient load, concentration, yield, and distribution for a large international watershed.

ABSTRACT: A large international watershed, the St. Clair-Detroit River System, containing both extensive urban and agricultural areas, was modeled using the Soil and Water Assessment Tool (SWAT) model. The watershed, located in southeastern Michigan, United States, and southwestern Ontario, Canada, encompasses the St. Clair, Clinton, Detroit (DT), Sydenham (SY), Upper, and Lower Thames subwatersheds. The SWAT input data and model resolution (i.e., hydrologic response units, HRUs), were established to mimic farm boundaries, the first time this has been done for a watershed of this size. The model was calibrated (2007–2015) and validated (2001–2006) with a mix of manual and automatic methods at six locations for flow and water quality at various time scales. The model was evaluated using Nash–Sutcliffe efficiency and percent bias and was used to explore major water quality issues. We showed the importance of allowing key parameters to vary among subwatersheds to improve goodness of fit, and the resulting parameters were consistent with subwatershed characteristics. Agricultural sources in the Thames and SY subwatersheds and point sources from DT subwatershed were major contributors of phosphorus. Spatial distribution of phosphorus yields at HRU and subbasin levels identified locations for potential management targeting for both point and nonpoint sources and revealed that in some subwatersheds nonpoint sources are dominated by urban sources.

(KEYWORDS: SWAT; watershed modeling; international watershed; field-scale; flow and water quality.)

INTRODUCTION

Watersheds are widely accepted units of analysis for water resources planning and management (McKinney et al. 1999; IJC 2009; Sheelanere et al. 2013), and have been the focus for guiding water resource and management decisions for decades. However, their natural and anthropogenic processes and activities are often too complex and variable, both spatially and temporally, to be captured thoroughly through monitoring alone (Mirchi et al. 2009). Therefore, watershed modeling tools, especially flow and water quality models, have been used increasingly to simulate watershed processes and human use to help in guiding those decisions at local, national, and international scales (Singh and Frevert 2006; Madani and Mariño 2009; Daniel 2011). These modeling tools are particularly valuable for developing a common understanding and framework for setting goals among nations with shared watersheds (IJC 2009).

One of the most widely used watershed models is the Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998); a semi-distributed, physically based flow, and water quality model that has been used in watersheds around the world with widely varying

Paper No. JAWRA-18-0142-P of the *Journal of the American Water Resources Association* (JAWRA). Received October 10, 2018; accepted May 21, 2019. **Discussions are open until six months from issue publication**.

Water Resources (Dagnew), Environmental Consulting and Technology, Inc., Ann Arbor, Michigan, USA; School for Environment and Sustainability (Scavia), and Graham Sustainability Institute (Wang, Long), University of Michigan, Ann Arbor, Michigan, USA; Sustainable Engineering and the Built Environment (Muenich), Arizona State University, Tempe, Arizona, USA; and Food, Agriculture and Biological Engineering (Kalcic), Ohio State University, Columbus, Ohio, USA (Correspondence to Dagnew: adagnew@ectinc.com).

Citation: Dagnew, A., D. Scavia, Y.-C. Wang, R. Muenich, C. Long, and M. Kalcic. 2019. “Modeling Flow, Nutrient, and Sediment Delivery from a Large International Watershed Using a Field-Scale SWAT Model.” *Journal of the American Water Resources Association* 55 (5): 1288–1305. https://doi.org/10.1111/1752-1688.12779.

characteristics in size and composition (Gassman et al. 2007, 2014). It is designed to capture information ranging from very coarse to fine spatial scales by dividing the watershed into subbasins based on topography, and then dividing the subbasins into smaller hydrologic response units (HRUs) based on unique land use (LU), soil type, slope, and/or management combinations. While these HRUs can be at very fine scales, this increased resolution and complexity improves results only when there is an equivalent level of input information (Jakeman et al. 2006; Johnston and Smakhtin 2014). Fortunately, in recent years, extensive datasets, such as land-use data generated from remote sensing and tile drainage systems characteristics collected by government and nongovernment organizations, enable relatively detailed watershed models.

However, even with detailed input data, SWAT still has a large number of parameters that cannot be measured directly and therefore need to be estimated through model calibration (Li, Weller, et al. 2010; Li, Shao, et al. 2010). The most frequently used calibration practice is to evaluate simulation performance at a single downstream location (Shi et al. 2013), which ignores spatial heterogeneity. This is particularly problematic for large systems where parameters estimated for some parts of the watershed may be unrealistic for other parts. For example, Leta et al. (2017) assessed the impact of calibrating at a single site, at multiple sites with constant parameter values, and at multiple sites with varying parameter values for a 1,162 km^2 watershed in Belgium. Their results indicated using different parameter values among different regions improved calibration results. In their study for a 239 km^2 watershed in Idaho, Zhang et al. (2008) also showed the importance of calibrating at multiple monitoring sites for better representations of regional conditions and goodness of fit. Hence, for large and/or spatially heterogeneous watersheds, calibration/validation processes at multiple locations is crucial to ensure accurate representations of local and regional flow, sediment, and nutrient simulations (Zhang et al. 2008; Wang et al. 2012; Bai et al. 2017; Leta et al. 2017).

A water quality agreement between the United States (U.S.) and Canada (GLWQA 2016), crafted in response to Lake Erie's re-eutrophication (Scavia et al. 2014), has led to new phosphorous loading targets. Attention has logically been placed on loads from the Detroit (DT) and the Maumee rivers because they contribute about 90% of total phosphorus (TP) load to the western basin of the lake (Scavia et al. 2016). While there have been several assessments for the Maumee watershed (e.g., Kalcic et al. 2016; Muenich et al. 2017; Scavia et al. 2017), there has been no similar assessment for the nearly 20,000 km^2 international watershed that drains into Lake Erie from the Detroit River. This study was designed to begin filling that gap with a robust watershed model to allow assessing potential nutrient load reduction strategies.

The goal of this study was to calibrate the SWAT model for this very large, complex international watershed at multiple locations and investigate the spatial distribution of nutrient sources and loads. In pursuit of this goal, we first assembled and harmonized into seamless model input U.S. and Canadian data that have their own characteristics, developed with different methodologies and interpretations, and with their own formatting and naming conventions (IJC 2015).

STUDY AREA

The St. Clair-Detroit River system drains a 19,040 km^2 watershed area from parts of southeastern Michigan in the U.S. (40% of watershed area) and southwestern Ontario in Canada (60% of watershed area) and contributes its load to Lake Erie through the Detroit River (Figure 1). It is composed of about 50% cropland, 20% urban area, 12% forest, 8% grassland, and 7% water bodies. The U.S. portion of the watershed is dominated by the DT Metropolitan area, whereas the Canadian portion is dominated by tile-drained croplands growing corn, soybeans, and winter wheat. Over the 15-year study period (2001–2015), total annual precipitation and annual average temperatures vary between 740 and 1,200 mm, and 7.5 and 11.0°C, respectively, averaging at 908 mm and 9.3°C. Elevation ranges from 422 m above sea level at the watershed boundary to 145 m at the outlet, with mostly flat slopes.

The U.S. portion drains three HUC8 watersheds (St. Clair [SC], Clinton [CL], and DT subwatersheds) drained primarily by the Black River (BR), Clinton River (CR), and Rouge River (RR), respectively. The Canadian portion drains three tertiary watersheds (Upper Thames [UT], Lower Thames [LT], and Sydenham [SY] subwatersheds) through the Thames River (TR) and Sydenham River (SR). For this study, the TR includes both Upper Thames River (UTR) and Lower Thames River (LTR) segments. The watershed includes two smaller subwatersheds, Essex in Canada and Lake SC in the U.S. While calibration and validation were performed at the outlet of the six major rivers (BR, CR, RR, SR, UTR, and LTR), most load assessments were made for the entirety of each subwatershed (SC, CL, DT, SY, UT, and LT) that the major rivers drain. Hence, it is important to note the difference in names between the subwatershed and

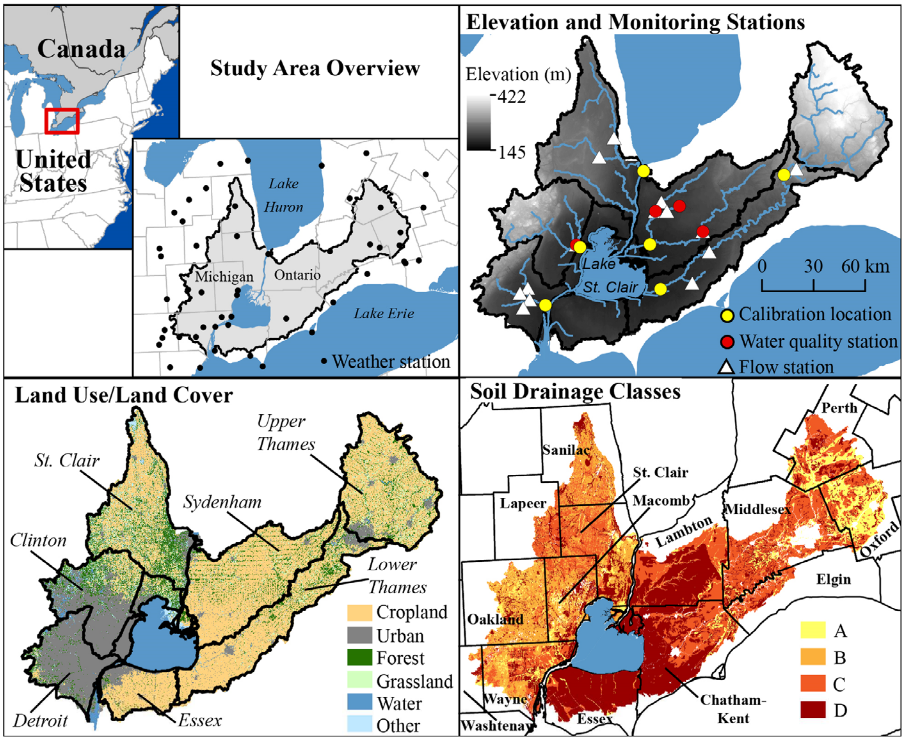

FIGURE 1. Study area with geographic location and weather stations (top-left), land use/land cover and subwatershed boundaries (bottom-left), soil and county boundaries (bottom-right) and digital elevation model and calibration locations (top-right) information. The channel which connects Lake Huron to Lake St. Clair (SC) is SC River, and Lake SC to Lake Erie is Detroit River. Water flows from Lake Huron to Lake Erie through Lake SC.

river, especially for the DT and SC subwatersheds that are drained through the Rouge and BRs.

Overall, 79% of the watershed's agricultural land is in Canada and 83% of the urban land is in the U.S. The CL and DT subwatersheds are heavily urbanized (about 56% and 89% of each as urban, respectively), and the SC, SY, UT, and LT subwatersheds are dominated by agriculture (63%, 89%, and 87% agricultural, respectively). This spatial variation in LU/land cover (LULC) provides both challenges and opportunities for investigating model performance. Moreover, five of the six HUC8 (tertiary) subwatersheds drain into the 1,100 km^2 Lake SC (Figure 1) that retained an average 13% of its TP input over the 1998–2016, and 21% over the 2013–2015 time period (Bocaniov and Scavia 2018; Scavia et al. 2019).

DATA

Basic Inputs

With the exception of data on elevation and weather, all model input was obtained separately for the U.S. and Canada and then merged. Digital elevation model (DEM) data with 30 × 30 m resolution from the U.S. Geological Survey–The National Map (USGS 2016) were used for the entire watershed for elevation, slope, and subbasin delineation. Daily precipitation and maximum and minimum temperatures were obtained from the National Oceanic and Atmospheric Administration's Global Historical Climatology Network (NOAA-GHCN 2016) for 16 U.S. stations and 15 Canadian stations for 1999–2015 (Figure 1). LULC layers for 2011–2015 with 30 × 30 m grid cells were from the U.S. Department of Agriculture National Agricultural Statistics Service (USDA-NASS 2016) Cropland Data Layer and the Agriculture and Agri-Food Canada Annual Crop Inventory (AAFC 2016). The 2015 LULC data layer was used to setup the SWAT model and the five-year dataset was used to generate crop rotations. Soil data layers were from the USDA Natural Resources Conservation Service Soil Survey Geographic Database (USDA-NRCS 2017) and from the AAFC's Soil Landscapes of Canada (version 3.2) (AAFC 2016). Road network data were from U.S. Census Bureau (2016) and Ontario Ministry of Natural Resources and Forestry (OMAFRA 2016. OMNRF. Accessed November 2016, https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/en/main.home?uuxml:id=290bfd40-0c8b-46d0-9a6c-0c648d096515).

Flow and Water Quality

The USGS National Water Information System (USGS-NWIS 2016) and the Canadian National Water Data Archive hydrometric data (HYDAT 2016) were used to obtain daily flow data for the most downstream gauging stations in each subwatershed (Figure 1, Table S2). Any data gap of 60 days or more was filled using either the stage discharge relationship, if stage data were available, or with the unit area method using data from a nearby station along the same or adjacent stream. If a gap was <60 days, it was filled using structural time series (Ryberg and Vecchia 2017).

Total suspended sediment (TSS), total nitrogen (TN), nitrate (NO3), TP, and dissolved reactive phosphorus (DRP) concentration data for the U.S. were obtained from the Water Quality Portal (WQP 2016). Canadian data were from the Provincial Stream Water Quality Monitoring Network (PWQMN 2016) and Environment and Climate Change Canada (Debbie Burniston, Alice Dove, 2017, personal communication). Average sampling frequency ranged from 3 to 17 samples per year for the U.S. and 7 to 21 for Canada.

Because flow and water quality data were often measured at different locations (Figure 1), calibration points were generally at the most downstream water quality stations to avoid extensive interpolation of water quality concentrations and to account for most of the subwatershed areas. Daily flow data at the calibration locations were estimated using the drainage-area method (Hirsch 1979) from the upstream flow stations. Monthly and annual nutrient load estimates for calibration at these locations were made using the weighted regression on time, discharge, and season (WRTDS) method (Hirsch et al. 2010) based on sample concentration values and daily flow.

Management Data Layers

Management data layers include cropping systems, fertilizer and manure application rates and placement, tillage practices, and tile drainage. County level fertilizer sales data were from the International Plant Nutrition Institute (IPNI 2016) for the U.S. and provincial level fertilizer sale data were from Statistics Canada (STATCAN 2016). Unique application rates for individual crops were based on regional N and P fertilizer application rate information from USDA Economic Research Service (USDA-ERS 2016) and Canadian Field Print Initiative (2017). Manure amounts were based on livestock (dairy, beef, swine, sheep, goat, chicken, and turkey) counts in each county from USDA-NASS (USDA-NASS 2016) and from the Ontario Ministry of Agriculture, Food and Rural Affairs (OMAFRA 2016). Spatial distribution of manure application in Canada was provided by OMAFRA (Kevin McKague, 2017, personal communication) as locations (points) of animal farms and field areas that receive manure from each animal farm without explicit indication of which field (s).

Tillage practices for subwatersheds in the U.S. and county/subcounty level for Canada were obtained from USGS and STATCAN, respectively. The latest U.S. tillage data were from 2004, but it detailed practices for each crop type. Canadian data were from 2011, but they did not distinguish among crop types. Data on the distribution of subsurface (tile) drainage systems in Canada were from OMAFRA (2016). Tile drainage information is not available for the U.S., so we assumed all cropland with poorly drained soils employed tiles (Kalcic et al. 2015). Tile drainage installation depth and spacing specification for the Canadian side of the watershed were recommended to vary by soil type (Kevin McKague, 2017, personal communication). As such, tile depths were set at 650, 750, and 950 mm for clayey, silty, and sandy soils, respectively, with corresponding spacing at 8, 12, and 15 m, respectively. For the U.S. side, a uniform 1,000 mm depth and 20 m spacing were used.

Three reservoirs in the UT region (Fanshawe, Wildwood, and Pittock) with surface-area (ha)/volume (ha-m) controls of 262/1,235, 192/796, and 142/266, respectively, were included in the model. Information about the physical features of the reservoirs, daily outflow data, and water quality samples were obtained from the UTR Conservation Authority website (UTCA 2017) and Mark Helsten (2017, personal communication). Monthly industrial and municipal point source (Figure 2) data were collected from EPA Enforcement and Compliance History (U.S. Environmental Protection Agency 2017) and the Great Lakes Water Authority — Water Resources Recovery Facility (GLWA-WRRF) (Majid Khan, Catherine Willey, personal communication, 2018) for the U.S., and from Ontario Ministry of Environment and Climate Change Effluent Monitoring and Effluent Limits Regulations (https://www.ontario.ca/data/industrial-waste water-discharges) for Canada.

METHODOLOGY

Data Assimilation

Because this was a binational watershed study, it was essential to ensure data from the two countries were harmonized. The U.S. and Canadian LULC data

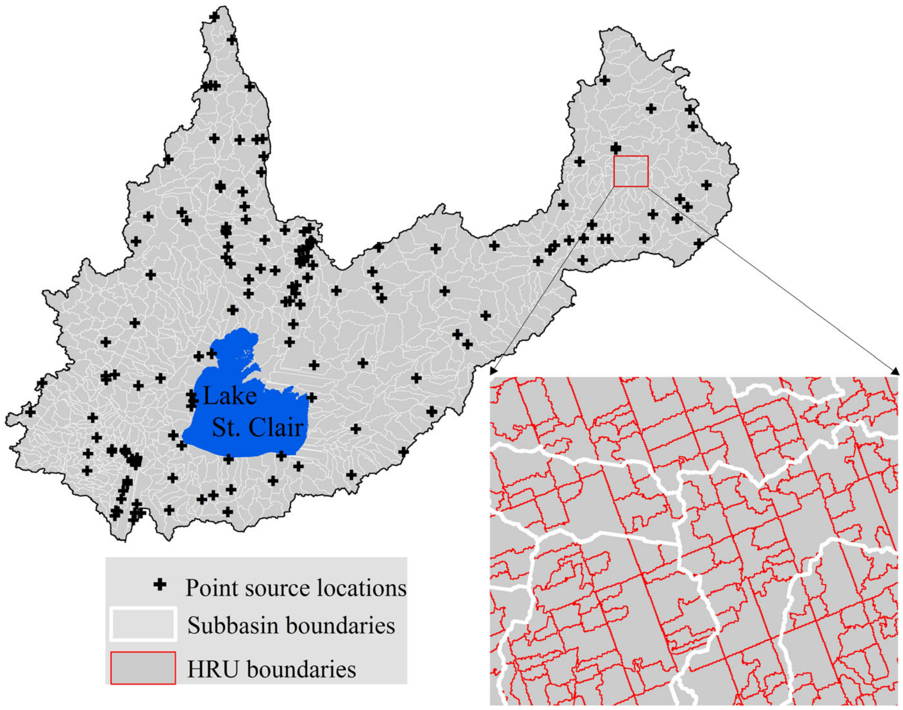

FIGURE 2. Subbasins and hydrologic response units (HRUs) along with point source locations in the watershed.

have the same resolution but different land-use type names and identification codes. Because SWAT is based on U.S. data types, Canadian LULC type names and identification codes were converted to the U.S. format (Figure 1). Canadian soil data required additional calculations and unit conversions to conform to U.S.-based SWAT parameters (Table 1). Although there is some anecdotal evidence Canadian manure production per animal may be different from the U.S., we used U.S. values for both.

Model Setup

Using an area threshold based on the DEM and identification of additional outlet locations to accommodate future comparison and/or spatial verification from smaller subwatersheds models and/or evolving monitoring efforts, the watershed was divided into 800 subbasins (Figure 2) with an average area of 24 km^2. Smaller subbasins were created in predominantly urban areas to capture their higher variation in drainage and land-use types, and to potentially test urban management scenarios in future work at finer spatial scales. Each subbasin was further divided into HRUs using predefined field boundaries as discussed below. The ArcGIS interface, ArcSWAT, version 2012.10_3.18 was used for setup and SWAT2012 rev635, as modified by Kalcic et al. (2016), was used for simulations.

Field Boundaries and Data Processing

LULC, road network, and subbasins were used to define field boundaries using a combination of the methods described by Kalcic et al. (2015) and Teshager et al. (2016). Following Teshager et al. (2016), LULC and road network data were used as the primary sources to identify field boundaries. As such, the watershed was divided into 27,751 "fields" with an average area of about 69 ha, of which 15,219 (54.8%) are cropland. These fields were assigned unique soil type identifiers (Kalcic et al. 2015), and an ArcGIS shapefile that contains the soil identifiers and LULC for each field was created. The shapefile was then used to define HRUs in the ArcSWAT model setup with 0% thresholds for LULC, soil, and slope, and the 27,751 fields thus became the SWAT HRUs (Figure 2).

A key advantage of using field boundaries to generate HRUs is that management practices can be assigned at a more detailed spatial scale than in more traditional SWAT models. Crop rotations for each HRU were estimated by overlaying the 2011–2015 LULC data layers and extracting the major cropping systems in each cropland fields. The most dominant crop rotations involved corn, soybeans, and winter wheat. In order to maintain a manageable number of rotations, crop rotations were limited to a maximum of three years. Tile drainage data and field boundaries were overlaid to determine fields with tile

TABLE 1. Relationship between Canadian vs. Soil and Water Assessment Tool (SWAT) major soil parameter names and units, and the changes made.

SWAT soil		Canadian soil			
Parameter	Unit	Parameter	Unit	Comments	Equations
SOL_ZMX	mm	max (LDEPTH)	cm	*Converted*	Unit conversions
SOL_Z	mm	LDEPTH	cm	*Converted*	
SOL_AWC	mmH_2O/mm soil	NA	NA	*Calculated*	SOL_AWC = KP1500-KP33
SOL_K	mm/h	KSAT	cm/h	*Converted*	Unit conversions
ROCK	% total weight	COFRAG	% by volume	*Converted*	
SOL_ALB	fraction	NA	NA	*Calculated*	SOL_ALB = 0.4/(0.688*SOL_CBN)
USLE_K	0.013 ($t.m^2.h$)/(m^3.t.cm)	NA	NA	*Calculated*	Equation from SWAT I/O documentation (Arnold et al. 2012 Page 307)

Notes: NA, parameter not available; SOL_ZMX, max(LDEPTH) = maximum rooting depth of soil; SOL_Z, LDEPTH = depth from soil surface; SOL_AWC, available water capacity of soil; SOL_K, KSAT = saturated hydraulic conductivity; ROCK, COFRAG = rock fragment content; SOL_ALB, moist soil albedo; USLE_K, soil equation erodibility factor; SOL_CBN, organic carbon content of soil; KP1500, water retention at 1500 kP; KP33, water retention at 33 kP.

drainage systems. If the majority of a field was covered by the tile drainage layer, the field was considered to have tiles. Canadian fields (HRUs) that receive manure were determined based on proximity to animal farm location and total field area receiving manure from the animal farm.

The field boundaries were also used to distribute the county level conventional (Cv), conservation (Cs), and no-till (NT) tillage practices. The type of tillage practices assigned for a crop field in a county depended on the proportions of practices (Cv:Cs:NT) in that county and the cropping system (crop rotation) in the field. Cv tillage practices were assigned more in fields with intensive corn, single crop, or nonalternate rotations (e.g., continuous corn). On the other hand, more conservative tillage practices (Cs and NT) were assigned more in fields with alternate rotations (e.g., corn–soybeans–winter wheat). Given this information on field-scale crop rotations and regional application rates of mineral N and P for different crops, a similar approach was used to allocate county/provincial level fertilizer applications across agricultural HRUs. Corn fields generally received N and P fertilizer at higher application rates than winter wheat or soybeans. Corn in continuous-corn rotation received more mineral fertilizer than corn in any other alternate rotations (Table S1).

The field boundaries were also designed for analysis and display of input and output information (e.g., distribution of fertilizer/manure application, flow, phosphorus load, etc.), and to model infield best management practices (e.g., filter strips, grassed waterways, drainage management, etc.) at finer scales.

Calibration and Validation

Calibration and validation were performed at the outlets of the three U.S. subwatersheds and the three Canadian subwatersheds (Figure 1). The model simulated 1999–2015, using the first two years as the warm-up period. Flow was calibrated for 2007–2015 and validated for 2001–2006 at daily, monthly, and annual time scales. Upon successful flow calibration, the model was calibrated for TSS loads, followed by nutrients (TN, NO3, TP, and DRP) at daily time steps. Since monthly and annual scales were more relevant for management application and policy advice, water quality parameters were further adjusted to also match WRTDS's monthly and annual water quality loads.

The significant variation in LULC and land management across such large watershed was expected to result in different controlling dynamics, especially physical drivers. Therefore, during calibration, certain subbasin and HRU parameters were allowed to vary across the six major subwatersheds (Tables S3 and S4). We used both manual calibration and SWATCUP's SUFI2 (Abbaspour 2015) auto-calibration procedures. Watershed level parameters were initially adjusted manually based on experience and information about local conditions. For example, parameters that control snow cover were estimated based on comparisons of observed and simulated snowfall frequency and snow depth values for the area. Then, SUFI2 was used to estimate HRU and subbasin parameter values and to understand their general direction of change in each major subwatershed. Finally, manual calibration was used for all parameters to improve fit.

Model performance was evaluated by comparing observed and simulated values using three commonly used statistics for watershed modeling: coefficient of determination (R^2), Nash–Sutcliffe efficiency coefficient (NSe), and percent bias (PBs).

The NSe is used to assess how good simulated values fit observations. The NSe values range from 1 to $-\infty$ with 1 being a perfect 1:1 fit between simulated and observed values. PBs provides insights on the tendency of simulations in under- or over-estimating values, and ranges from $-\infty$ to $+\infty$. A PBs value of 0.0% indicates a perfect match between average simulated and observed values, and negative and positive values show under- and over-estimation, respectively. The R^2 values examine how well simulated values are correlated with observations, that is, follow similar trends; 0.0 indicates no correlation and 1.0 a perfect correlation. According to Moriasi et al. (2007), monthly simulations with NSe > 0.75 are considered "very good," >0.65 and ≤0.75 are "good," >0.50 and ≤0.65 are "satisfactory," and values ≤0.50 are "unsatisfactory" for watershed models. Similarly, values of |PBs| < 10%, 10%–15%, 15%–25%, and ≥25% fall into those same categories for flow simulations. The same categories apply for sediment if |PBs| < 15%, 15%–30%, 30%–55%, and ≥55% and for nutrients |PBs| < 25%, 25%–40%, 40%–70%, and ≥70%.

Finally, to evaluate the significance of allowing parameters to vary among subwatersheds, the final calibrated flow parameter set for each subwatershed was assigned uniformly across the entire watershed and NSe and PBs were compared to those for the varying parameter case. As a result, six sets of statistics for each subwatershed were compared.

RESULTS AND DISCUSSION

Input Characterization

Using the spatial allocation scheme (HRU boundaries), we distributed crop rotations, fertilizer/manure applications, tile drainage, and tillage practices for each HRU explicitly (Figure 3) to better represent actual conditions. With respect to cropping systems, three-year rotations involving corn (C), soybeans (S), and winter wheat (W) covered about 43% of the cropland area. Distribution of crop rotation types was similar within each country, with CSW dominating, followed by CS and then SS (Table 2). However, corn-only or soybeans-only cropping systems were more abundant in Canada than the U.S. (Figure 3), and 40% of the Canadian soybean intensive fields were in the Essex region. Crop rotations for each county and HUC8/tertiary subwatershed are detailed in Figures S1 and S2.

Allocation of Cv, Cs, and NT tillage practices (Figure 3) resulted in about 70% of cropland receiving alternating practices with either two or three tillage types (Figure 4). The most dominant tillage practice was Cs-NT (39.4%) and was mainly in Canada. U.S. croplands were dominated by Cv-Cs tillage. While cropping systems that alternate corn–soybeans–winter wheat in a three-year rotation received all three tillage practices, most of the continuous Cv tillage practices were assigned for single crop rotations (Figure 5).

Tile drainage was denser in Essex region, lower parts of SY and LT, and upper parts of SC and UT subwatersheds (Figure 3). About 67% of Canadian and 55% of U.S. agricultural areas were considered tiled (Table 3). Most of the UT and upper parts of SY agricultural fields receive manure generated in their respective counties, whereas few fields in LT and Essex area received manure. In the U.S., manure was assumed to be distributed across all agricultural fields, and because of this and fewer livestock, solid manure application rates in the U.S. were lower (85–670 kg/ha for dairy, 8–50 kg/ha for beef, and 1–35 kg/ha for swine) than in Canada (345–1,082 kg/ha for dairy, 261–695 kg/ha for beef, and 667–1,556 kg/ha for swine).

Calibration and Validation

Flow. The model reproduced observed flow hydrographs fairly well (Figure 6). Using Moriasi et al. (2007) performance criteria, the monthly flow calibration NSe (Table 4) were judged "very good" for the ULT, LTR, and SR subwatersheds; "good" for BR and RR; and "satisfactory" for CR. PBs during calibration and both NSe and PBs during validation for all six locations were rated as "very good." The model also performed well at daily (NSe > 0.5 except BR, and |PBs| < 10%) and annual (NSe > 0.65 and |PBs| < 10%) time scales (Table S5).

As expected, allowing parameters to vary among subwatersheds provided a better representation of regional conditions and improved model performance (Tables S2 and S3). During calibration, some flow parameter values varied substantially across the watershed, especially between agricultural- and urban-dominated subwatersheds (Table S4). Flow was particularly affected by changes in parameters for main channel average width (CH_W2) and/or depth (CH_D) and average slope (CH_S2) in both of the highly urbanized streams (CR and RR). This adjustment for urban streams is consistent with the

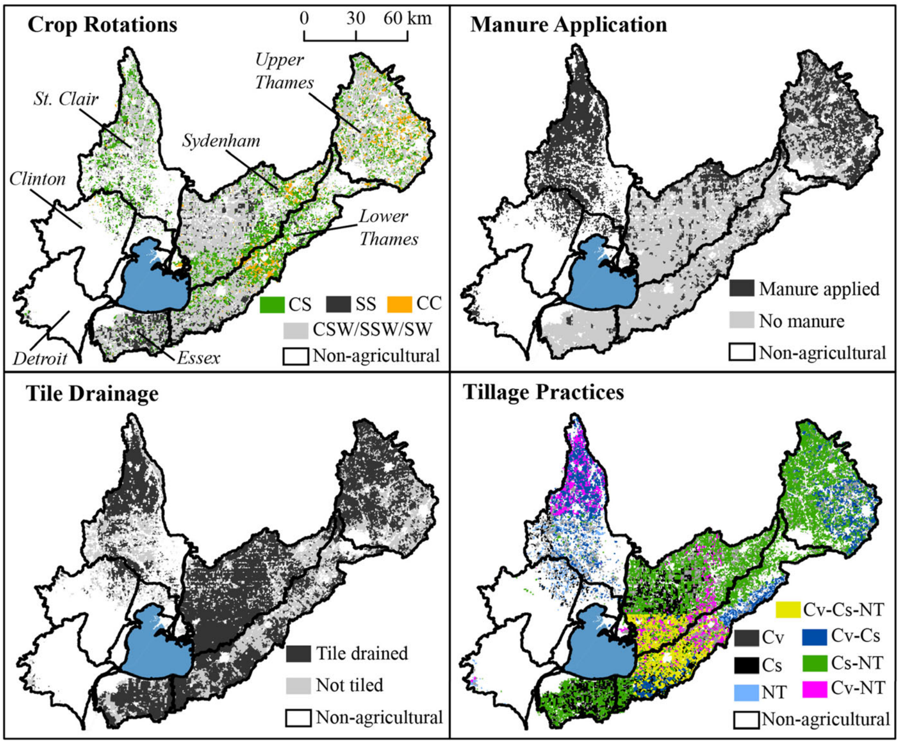

FIGURE 3. HRU-level agricultural management practice model inputs (C, corn; S, soybeans; W, winter wheat; Cv, conventional tillage; Cs, conservation tillage; NT, no-till).

TABLE 2. Percentages of cropland area covered with the different types of crop rotations divided between United States (U.S.) and Canada.

	% Cropland area		
Crop rotation	**Canada**	**U.S.**	**Overall**
CC	8.4	1.6	7.1
CS[1]	25.4	35.5	27.3
SS	13.5	13.1	13.4
CSW[2]	42.8	45.4	43.3
SW	0.4	0.3	0.4
SSW	9.5	4.1	8.5
Total	100.0	100.0	100.0

[1]Includes both CS and SC rotations.
[2]Includes CSW or SWC or WCS rotations.

fact urbanization not only increases runoff but also alters routing of flow downstream through changes in channel dimensions (Booth 1990; Baker et al. 2014).

The calibration also resulted in substantially lower soil water capacity parameter values (SOL_AWC) in urbanized areas, consistent with the fact urbanization reduces soil permeability, infiltration, and water holding capacity through soil disturbance, displacement, pore space reduction, low organic matter, and high surface traffic (Craul 1985; Jim 1998; Yang and Zhang 2015; Wiesner et al. 2016). For example, the European Commission Bio Intelligence Serve (2014) reported changing forest land to urban land could decrease the maximum soil water content by up to 25%.

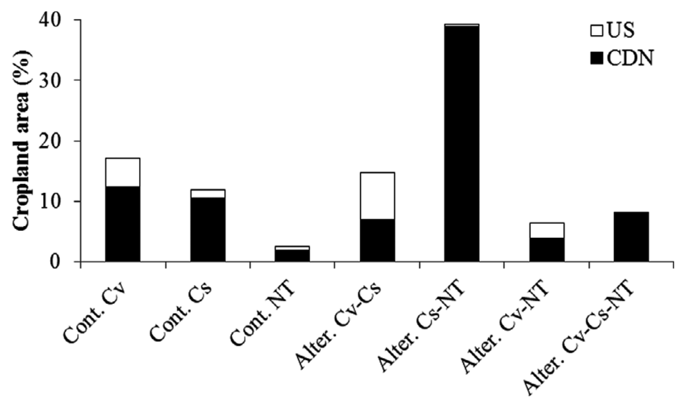

FIGURE 4. Estimated distribution of tillage practices in U.S. and Canadian parts of the St. Clair-Detroit River system watershed (Cont. Cv, continuous Cv; Cont. Cs, continuous Cs; Cont. NT, continuous NT; Alter., alternating; CDN, Canada).

Differences in other parameter values, such as increasing the runoff curve number from the SWAT default value for moisture condition II (CNII) for the UT by 10% and the LT by 4% reflected the differences in slopes between the two regions (~0.12% and ~0.03%, respectively, along the main stream course).

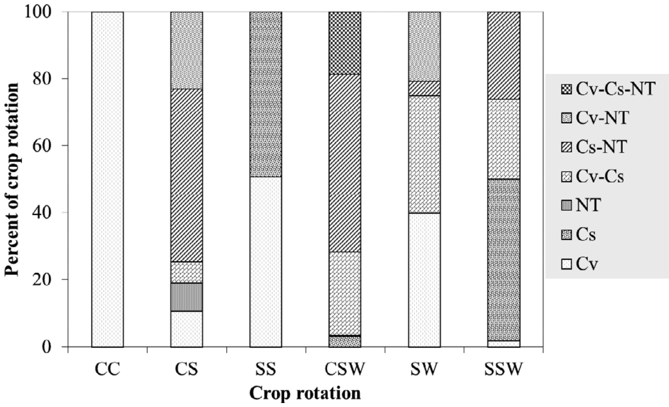

FIGURE 5. Estimated relationship between tillage practices and crop rotations.

TABLE 3. Percentages of agricultural area with tile drainage systems divided between U.S. and Canada at subwatershed level.

HUC8/tertiary name	Tiled area % Total area	Tiled area % Agricultural area
SC	37	59
Clinton (CL)	8	46
Detroit (DT)	1	16
Lake SC	5	29
U.S. total	18	55
Upper Thames (UT)	54	62
Lower Thames (LTR)	49	55
Thames total	51	59
Sydenham (SY)	69	77
Essex	58	72
Canada total	58	67
Watershed total	42	64

These two regions also have different soil drainage class distributions. While the UT has more well drained soils, the LT is dominated by poorly drained soils. As such, SOL_AWC was increased by 10% above the default value and the soil evaporation compensation factor (ESCO) was set at 0.90 for the LT, compared to an ESCO value of 0.30, and the default value for SOL_AWC for the UT. The increase in SOL_AWC for the LT reflected the higher water holding capacity of the poorly drained soils. Moreover, the higher ESCO value for the UT was consistent with its higher water holding capacity of the soil that compensated for evaporation.

Overall, comparison of the final flow calibration statistics (Table 4) against statistics from uniform parameters across the entire watershed (Table S6) showed the strength of varying parameter values. If, for example, parameters which were best for UTR flow conditions were used across the watershed, the NSe values for CR, BR, and RR would have dropped by 62%, 11%, and 6%, respectively, and the |PBs| values for CR, BR, and SR would have increased by 34.3%, 29.2%, and 12.7%, respectively. Similarly, if best parameter sets for CR flow conditions were used across the watershed, |PBs| values would have increased by 25.4%, 19.6%, 13.6%, 12.5%, and 11.9%, for RR, BR, LTR, UTR, and SR, respectively, and the NSe values for RR and BR would have dropped by 34% and 14%.

A closer look at the effects of parameter values from one subwatershed applied to another indicated that even exchanging parameter sets between urbanized subwatershed (CR, RR) reduced fit. For example, using the CR optimal parameter values for the RR reduced its NSe and increased its PBs values by 34.3% and 25.4%, respectively. The RR parameter values had similar effects for the CR. Interestingly, while parameter values from the agricultural subwatershed (SY) reduced fit for the urbanized river (CR), the urbanized subwatershed (CL) parameters had less impact on the agricultural one (SR).

Water Quality. Measured nutrients and sediment dynamics were also replicated sufficiently (Figure 7, Table 5, Figures S4–S7). Monthly water quality calibration and validation statistics were better for TP than DRP and better for TN than NO3. All calibrations and validations were rated as "good" or better for PBs. Most calibration and validation NSe values were rated as "good" or "satisfactory." However, the phosphorus-related NSe values for UTR calibration were unsatisfactory, as was the RR validation, and both calibration and validation for the BR. Similar to flow, ratings for the major rivers in agricultural subwatersheds (SR, LTR, and UTR) were better than river in urbanized subwatersheds (CR and RR).

Similar to flow, some water quality parameters vary considerably across subwatersheds (Table S4). For example, values of initial NO3 concentration in the soil layer (SOL_NO3) were set to 100 mg N/kg-soil for UT and SY, whereas values for CL and DT were 25 and 0 mg N/kg-soil, respectively, perhaps reflecting differences in soil fertility. The rate constant for in-stream mineralization of organic phosphorus to dissolved phosphorus (BC4) was higher for Canadian rivers (0.28, 0.25, and 0.16 day^{-1} for SR, UTR, and LTR, respectively) than for U.S. rivers (0.018 day^{-1} for all BR, CR, RR), suggesting potentially higher concentrations of DRP in Canadian streams. There are also distinct differences in parameter values between UT and LT subwatersheds. Almost all nutrient parameter values were higher for UT than LT, implying higher initial soil nutrient content and increased nutrient yields in the UT compared to LT.

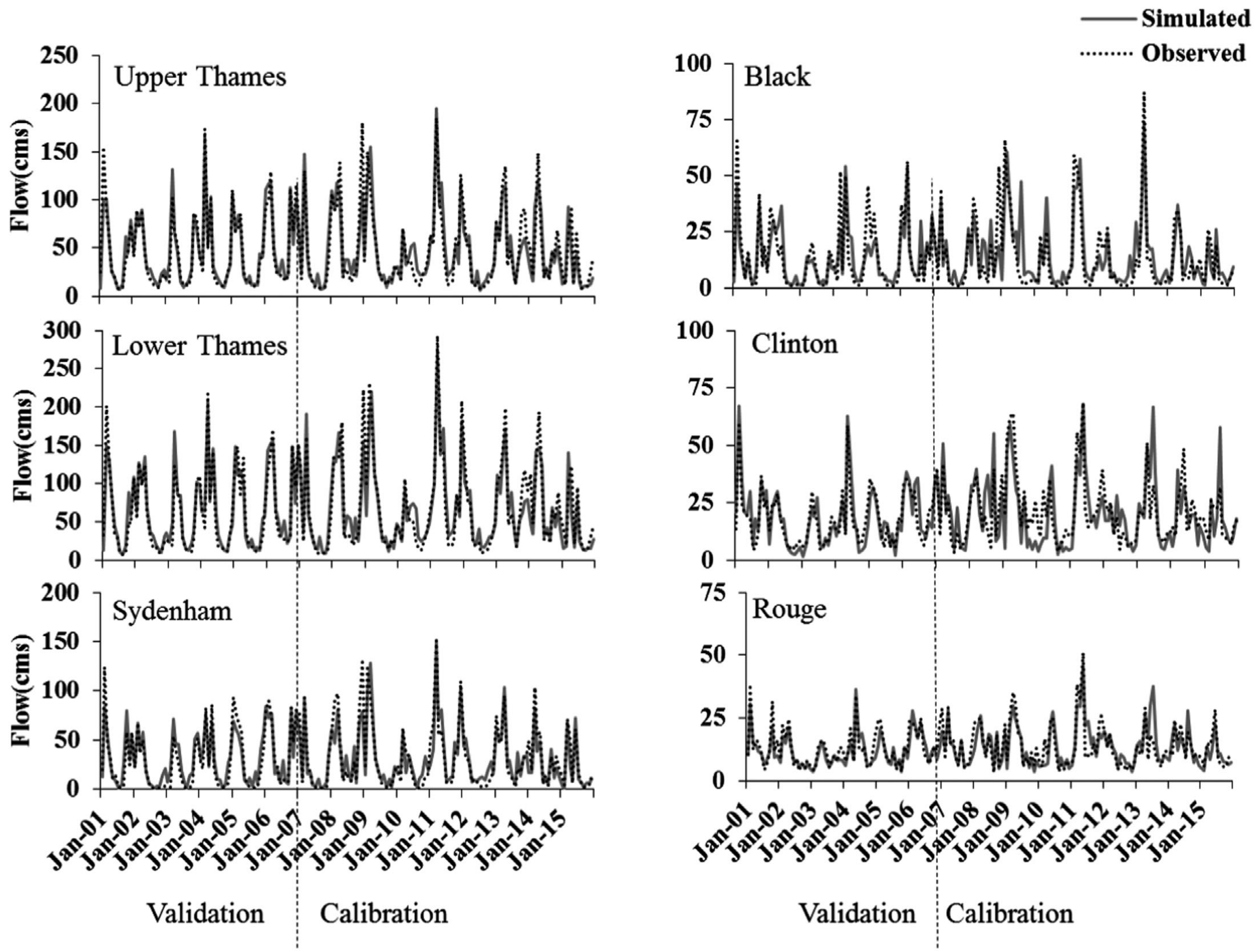

FIGURE 6. Monthly observed and estimated flow time series at each major subwatershed outlet locations for both calibration (2007–2015) and validation years (2001–2006).

TABLE 4. Monthly flow estimation performance statistics for calibration (2007–2015) and validation (2001–2006) years (R^2, coefficient of determination; NSe, Nash–Sutcliffe efficiency; PBs, percent bias).

	Monthly statistics for flow calibration (validation) period					
Statistics	**Upper Thames River**	**Black River**	**Sydenham River**	**Clinton River**	**Lower Thames River**	**Rouge River**
R^2	0.84 (0.93)	0.72 (0.76)	0.85 (0.87)	0.63 (0.80)	0.87 (0.92)	0.71 (0.78)
NSe	0.84 (0.93)	0.72 (0.76)	0.85 (0.86)	0.53 (0.75)	0.87 (0.91)	0.70 (0.75)
PBs	0.1 (3.2)	9.2 (−2.9)	−1.2 (8.4)	−2.7 (1.9)	−2.7 (5.4)	−1.1 (−8.5)

Nutrient Load Assessments

Because phosphorus is the primary driver of interest in Lake Erie (Scavia et al. 2014, 2016), we focus primarily on phosphorus loading.

Annual Average Loads. The DT and the Thames (UT and LT) subwatershed loads were similar and together contribute >60% of the TP and >70% of the DRP loads on an average annual basis (Table 6). However, about 90% of TP and DRP load from the DT subwatershed came from point sources, mainly one waste water treatment plant, whereas about 90% of the load from the Thames comes from agriculture. Despite being mainly urban, the CL subwatershed load came primarily from nonpoint source

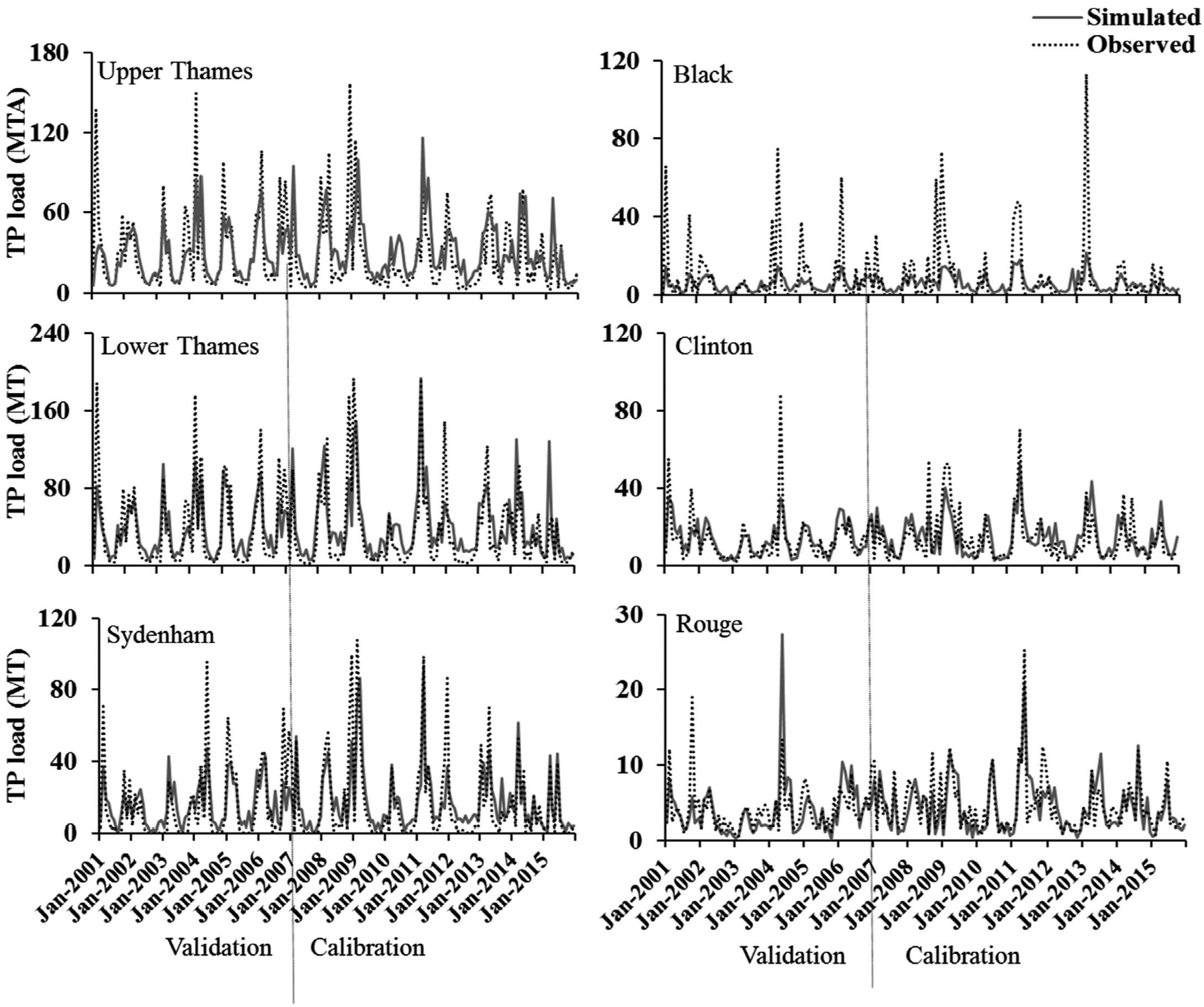

FIGURE 7. Monthly observed and estimated total phosphorus (TP) time series at the six major subwatershed outlet locations for both calibration (2007–2015) and validation (2001–2006) periods. MTA, metric ton per annum.

runoff, with combined urban and agricultural nonpoint sources accounting for 83% and 68% of CL's TP and DRP loads, respectively. Moreover, urban nonpoint source accounts for about 68% and 75% of CL's total nonpoint source TP and DRP loads, respectively. Phosphorus loads from the SY, the most agriculturally intense subwatershed, accounted for 13% of the overall watershed's TP and DRP loads. Among the six subwatersheds, the SC delivered the lowest loads (10% and 5% of TP and DRP, respectively). The smaller subwatersheds (Essex and Lake SC; Figure 1) contributed 4.4% and 0.8% of TP, and 2.5% and 0.5% of DRP loads, respectively. Even though the Essex region subwatershed area was about twice that of the Lake SC subwatershed, it delivered about five times the phosphorus load due to extensive agriculture and densely tiled soils.

DRP represented 42% of the TP load overall; however, it was 52% of the point sources and 37% of the nonpoint source TP load. While this variation in the DRP/TP ratio did not seem to be correlated with the composition of LULC, there were clear differences among different sources. The DRP fraction from U.S. nonpoint sources was much lower than from Canadian nonpoint sources, likely due to extensive tile drainage in the Canadian portion. In contrast, U.S. point sources had higher DRP fractions.

Our annual average TP load estimates were similar to the WRTDS-based averages reported by Scavia

TABLE 5. Monthly water quality model performance statistics for calibration (2007–2015) and validation (2001–2006) years.CL

	Statistics	Monthly statistics for water quality calibration(validation)					
		UT	Black	SY	CL	LT	Rouge
TP	R^2	0.54 (0.63)	0.54 (0.59)	0.75 (0.68)	0.64 (0.55)	0.62 (0.75)	0.73 (0.42)
	NSe	**0.48** (0.59)	**0.29** (**0.25**)	0.73 (0.62)	0.64 ((0.54)	0.59 (0.70)	0.71 (**0.10**)
	PBs	22.6 (9.7)	−25.6 (−29.1)	5.9 (6.3)	5.6 (4.8)	18.0 (9.6)	−5.0 (−4.8)
DRP	R^2	0.44 ((0.59)	0.48 (0.50)	0.64 (0.57)	0.57 (0.51)	0.55 (0.65)	0.71 (0.49)
	NSe	**0.42** (0.52)	**0.26** (**0.21**)	0.53 (0.52)	0.51 (**0.46**)	0.52 (0.58)	0.70 (**0.05**)
	PBs	27.8 (12.1)	−28.7 (−35.2)	−6.3 (−8.2)	9.6 (7.8)	21.5 (10.9)	25.1 (14.8)
TN	R^2	0.61 (0.65)	0.52 (0.55)	0.72 (0.65)	0.55 (0.54)	0.59 (0.66)	0.64 (0.53)
	NSe	0.54 (0.57)	**0.27** (**0.32**)	0.70 (0.61)	0.54 (0.52)	0.57 (0.62)	0.61 (**0.40**)
	PBs	7.8 (13.9)	36.4 (42.9)	17.9 (23.4)	−15.8 (−14.6)	−8.0 (8.6)	−5.2 (−11.4)
NO3	R^2	0.55 (0.52)	0.49 (0.47)	0.56 (0.52)	0.48 (0.48)	0.58 (0.66)	0.63 (0.42)
	NSe	0.53 (**0.49**)	**0.25** (**0.27**)	0.54 (**0.47**)	**0.44** (**0.42**)	0.53 (0.55)	**0.44** (**0.21**)
	PBs	15.6 (14.2)	−24.7 (−31.1)	5.9 (6.3)	−27.3 (−23.4)	−3.0 (13.6)	−15.1 (−24.8)
TSS	R^2	0.66 (0.77)	0.61 (0.62)	0.73 (0.67)	0.57 (0.63)	0.67 (0.70)	0.61 (0.68)
	NSe	0.59 (0.62)	**0.49** (0.52)	0.57 (0.55)	**0.47** (0.57)	0.60 (0.65)	0.58 (0.60)
	PBs	−7.5 (−2.9)	−15.6 (−9.9)	14.3 (11.6)	−16.5 (−12.4)	−12.0 (−7.9)	−14.0 (−18.4)

Notes: PBs and NSe ratings: bold = “unsatisfactory.”
TN, total nitrogen; NO3, nitrate; TSS, total suspended sediment.

TABLE 6. Average TP and DRP loads in MTA from both point sources (PS) and nonpoint sources (NPS) for each subwatershed.

	Total PS		Total NPS		Total Load		
HUC8/tertiary watershed name	TP	DRP	TP	DRP	TP	DRP	Drainage Area (km^2)
SC	28	15	150	21	177	36	3,025
CL	33	18	158	39	191	57	1,969
DT	492	257	55	30	547	287	1,594
Lake SC	5	3	9	1	14	4	575
U.S. total	558	293	372	91	929	384	7,163
SY	26	12	201	83	227	95	3,508
Thames	51	24	472	224	523	248	5,827
Essex	6	3	71	16	77	19	1,098
Canada total	83	39	744	323	827	362	10,433
Watershed total[1]	641	332	1,116	414	1,756	746	17,596

[1]This does not include Lake SC and other small unaccounted areas along SC and DT connecting channels.

et al. (2019) because our model was calibrated to WRTDS estimates (Figure 8). Our estimates were also similar to Maccoux et al. (2016) for the CR and BR, somewhat higher for the SR and TR, but considerably lower for the RR. Maccoux et al. (2016) and we used the same water quality monitoring station for the RR (Figure 1), but Maccoux et al. considered the drainage area for the station to be 565 km^2, whereas the actual drainage area for the station was 1,200 km^2 (USGS, https://waterdata.usgs.gov/nwis/nwismap/?site_no=04168550&agency_cd=USGS). Hence Maccoux et al.’s TP estimations for RR were overestimated because they overestimated unmonitored loads. Our annual average DRP load estimates showed similar discrepancies with Maccoux et al. (2016). Our estimate was much lower for the RR and much higher for the TR (Figure 11). Other discrepancies among the three studies could be due to the lack of more frequent water quality sample data, inherent differences in structure and assumptions of different estimation techniques, and span of years considered for the studies. For example, Maccoux et al. (2016) estimates for 2003–2013 used the Stratified Beale’s Ratio Estimator (Beale 1962; Dolan et al. 1981), Scavia et al (2019) estimates for 1998–2016 used WRTDS, and our estimates for 2001–2015 used SWAT.

In our analysis, annual TP loads increased slightly for all but CR between 2001 and 2009 and then decreased through 2015, with the trends more obvious for rivers in the agriculture dominated areas: SR, TR, and BR (Figures S3). On average between 2001 and 2009, TP increased by 24.7 metric ton per annum (MTA), 14.8 MTA, 4.1 MTA, and 1.6 MTA for TR, SR,

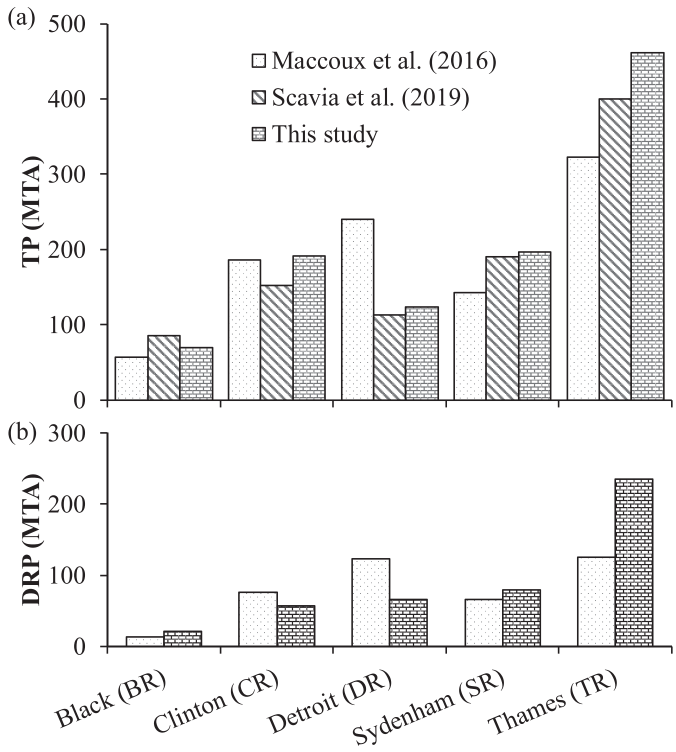

FIGURE 8. Comparisons of average annual phosphorus load estimations of TP (top), and dissolved reactive phosphorus (DRP, bottom), for each major subwatershed. The DT subwatershed loads in this figure do not include the Great Lakes Water Authority's waste water treatment point source loads.

Black, and RR, respectively. The decreases in TP between 2010 and 2015 were of 42.2 MTA, 23.7 MTA, 8.9 MTA, and 4.0 MTA, respectively. DRP followed similar trends, especially for the three rivers in agricultural subwatersheds, but to a lesser degree than TP, with DRP increases of 8.6 MTA, 4.4 MTA, 1.1 MTA, and 0.8 MTA, and decreases of 20.0 MTA, 9.7 MTA, 2.5 MTA, and 1.1 MTA for the same time intervals and river orders. Similar trends have been reported for the Maumee River (Baker et al. 2014), another major P contributor to Lake Erie. In most cases, these trends were reflecting changes in flow (Figures S3) but flow alone could not explain the trend for the TR and SR where flow was relatively constant between 2001 and 2005. It appears, in those cases, agricultural practices that provide access to more nutrient (e.g., high fertilizer applications) and facilitate nutrient movement into streams (e.g., tile drainage systems) are also responsible for these trends.

Spatial Distribution of Yields — Subwatershed Scale. Examining subwatershed and HRU yields provide information potentially useful for targeting management actions to the highest source areas. While the average annual TP loads from the DT and Thames subwatersheds were similar (Table 6), TP yields (3.43 and 0.90 kg/ha, respectively), and DRP yields (1.80 and 0.43 kg/ha, respectively) differ considerably due to the difference in drainage areas. In addition, the Thames delivered much more phosphorus from nonpoint sources (0.81 kg TP/ha and 0.38 kg DRP/ha) than the DT subwatershed (0.35 kg TP/ha and 0.19 kg DRP/ha) (Figure 9). The Thames and CL subwatersheds had similar overall TP yields; however, DRP yield was higher for the Thames. The SY and SC subwatersheds had comparable TP yields but the SY produces much higher DRP per hectare. Overall, the TP yield from the U.S. was about 60% higher than that from

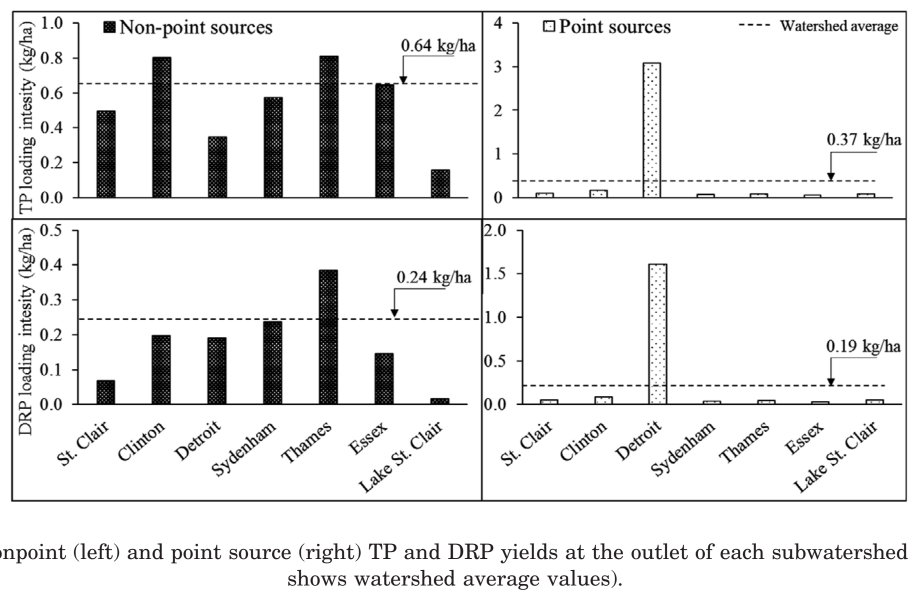

FIGURE 9. Average nonpoint (left) and point source (right) TP and DRP yields at the outlet of each subwatershed (dashed horizontal line shows watershed average values).

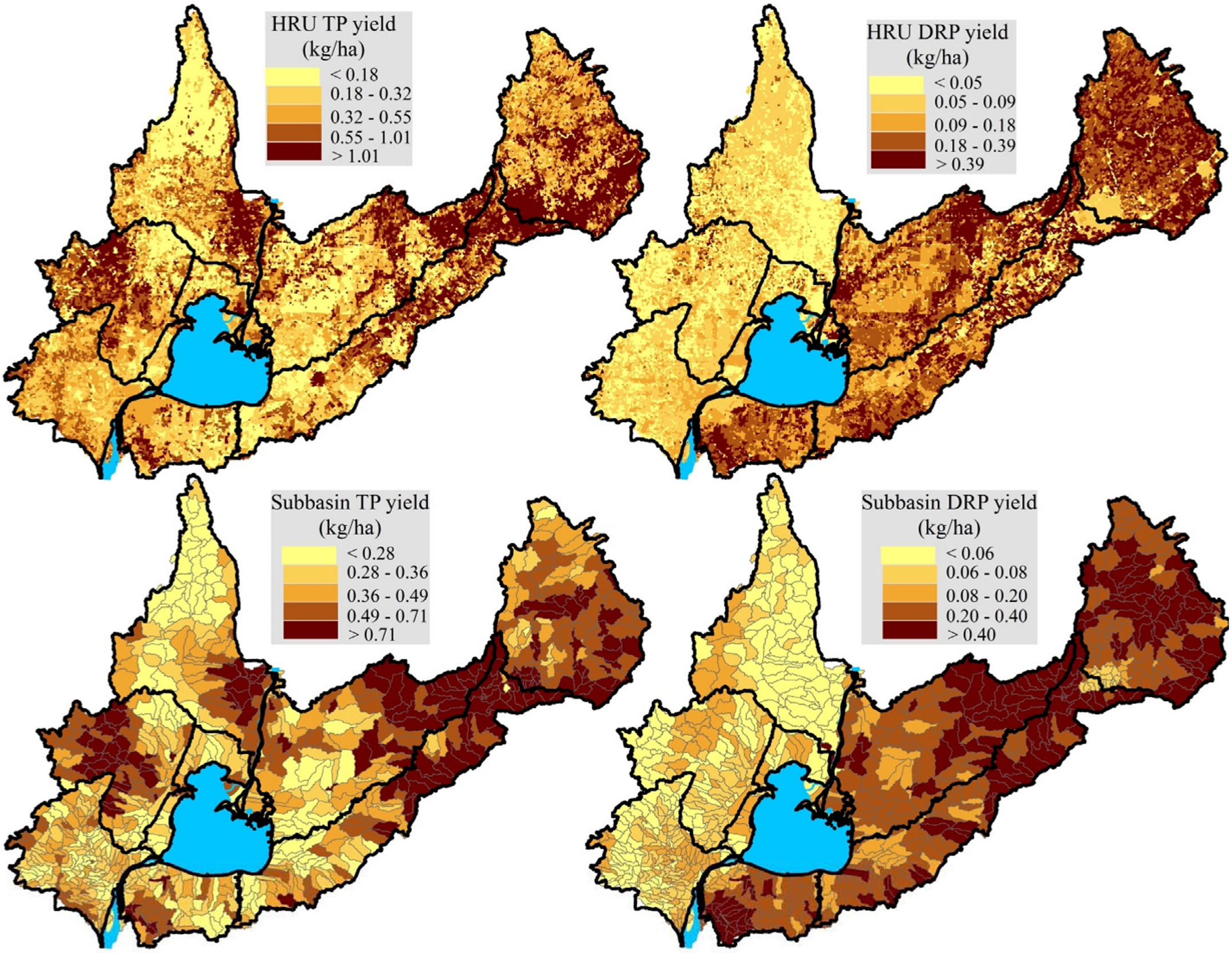

FIGURE 10. Distributions of HRU-level nonpoint source TP and DRP yields for each subwatershed. *Dots* indicate average yield values.

Canada. However, Canadian nonpoint source TP and DRP yields were 40% and 140% higher than the U.S., and the U.S. point source yields were nine times and ten times higher than Canada for TP and DRP, respectively.

These subwatershed-specific yields of total, point, and nonpoint sources (Figure 9) can be useful for developing load reduction strategies. For example, while the overall TP yield from DT subwatershed was about four times that of Thames; most of the yield from the DT subwatershed was from point sources. Comparing nonpoint source yields, on the other hand, showed the Thames subwatershed yield was about twice that of the DT. Thus, in exploring management options at this scale, more attention should be placed on point sources in the DT subwatershed and nonpoint source for agricultural areas of Thames subwatershed.

Spatial Distribution of Nonpoint Source Yields — Subbasin and HRU Scales. While evaluating yields at the subwatershed scale was useful for higher level strategies, assessments at subbasin (24 km^2) and HRU (field) scales enabled the potential targeting of management practices. Average HRU-level TP yields were 1.38, 1.10, 0.78, 0.53, 0.96, and 0.63 kg/ha for UT, LT, SY, DT, CL, and SC subwatersheds respectively. Average DRP yields are 0.69, 0.50, 0.33, 0.36, 0.32, and 0.12 kg/ha, respectively. The median HRU-level yields for TP and DRP were lower than the average values (Figure 10). This indicated regional average values were skewed by very high yielding areas across the watershed which in turn implied the presence of a good opportunity to focus management practices on certain areas to reduce the majority of nutrient loading from the watershed.

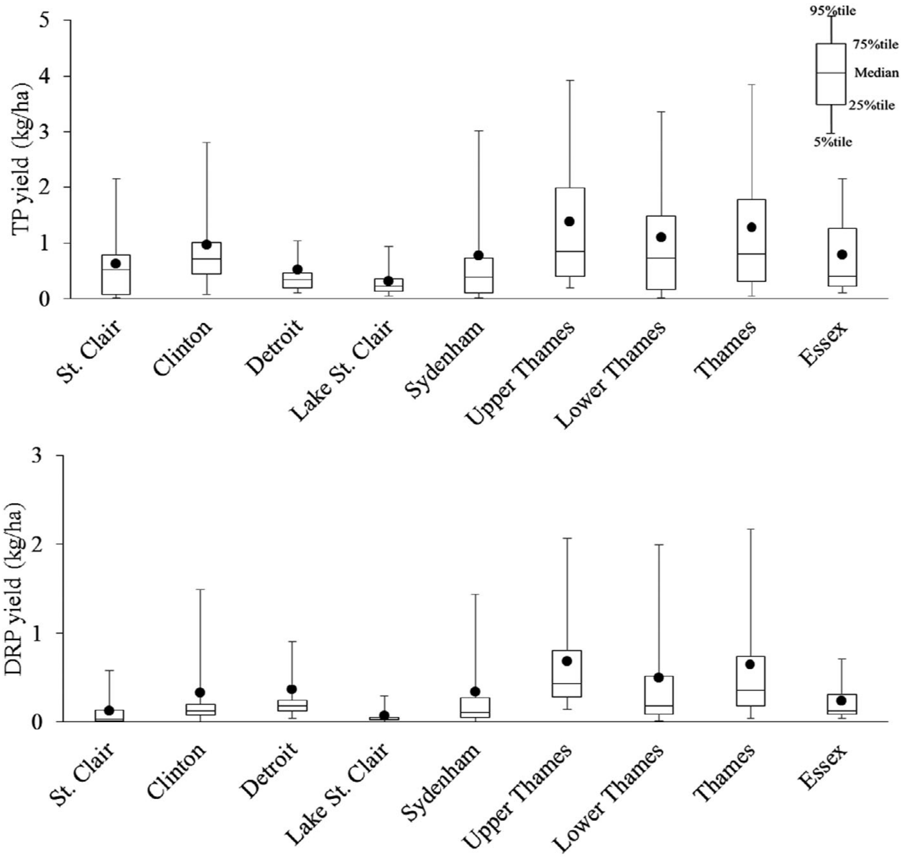

FIGURE 11. (a) HRU-level (top) and subbasin-level (bottom) distributions of nonpoint source TP (left) and DRP (right) yields.

Spatial patterns of nonpoint P yields at the HRU (field) and subbasin levels (Figure 11) provided further insight into potential areas of focus for nonpoint source reduction. High nonpoint source DRP yields spread relatively evenly across the Canadian watershed; whereas some of the highest TP yields were found in the upper parts of SY and Thames subwatersheds. DRP yields from the U.S. subwatersheds were distinctly lower than the Canadian counterparts; however, certain nonagricultural areas in the U.S. (lower parts of SC, upper parts of CL, and some places in DT subwatershed) appeared to have high yields as well. The higher DRP yields from Canadian subwatersheds could be attributed to higher tile drainage density, higher proportion of cropland, and higher fertilizer application rates. For example, inorganic P application rates ranged from 22.8 to 44.8 kg/ha, 7.8 to 24.4 kg/ha, and 7.4 to 13.7 kg/ha for corn, winter wheat, and soybeans, respectively, in Canada. These values were 5.9–10.9 kg/ha, 5.7–10.1 kg/ha, and 4.8–7.8 kg/ha in the U.S. Similarly, manure application rates were higher in Canadian agricultural areas (see “Input Characterization” section). The Canadian tile drainage system was also about twice as dense as in the U.S. (see “Management Data Layers” section). As a result, Canadian portions of the watershed had higher sources of DRP (inorganic fertilizer or manure) and a system that facilitates its movement (denser drainage tile system).

The distribution of P yields suggested U.S. agricultural areas had relatively low TP and DRP yields. For example, while the northern part of the CL subwatershed was agricultural, the higher P yields from that subwatershed were actually from nonagricultural areas in the central and west portions of the subwatershed. Similarly, yields from the agricultural areas in the northern part of the SC subwatershed were smaller than those from the nonagricultural areas. Most of the high phosphorus yielding areas in CL, for example, were urban areas located in a relatively higher slope region of the subwatershed. Moreover, the major point source contribution of the watershed came from the DT subwatershed (Table 6). These underscored the need to focus on Canadian agricultural runoff reduction strategies and both U.S. point source management and urban runoff reduction strategies.

CONCLUSION

We integrated and harmonized U.S. and Canadian datasets, including crop rotations, fertilizer/manure applications, tillage practices, and tile drainage systems; structured a SWAT model at finer resolution (field-scale) than ever done before for a 19,000 km^2 watershed; and calibrated and validated it at daily, monthly, and yearly time scales at six locations. While some input data (e.g., crop rotations) were constructed from a 30 × 30 m grid cell data, others (e.g., fertilizer application, tillage practice, manure generated, etc.) were available at county or provincial level. Hence, a great deal effort was invested in allocating model inputs from the lower spatial resolution to the field scale. Such distribution of model inputs not only improved model estimates at stream mouths but also provided more confidence in assessing flow and nutrient estimates at field level.

In most cases, a very good fit to flow measurements and good fit to water quality load estimates were achieved using manual and automatic calibration techniques at monthly time scales. It was evident from the calibration and validation processes that allowing some key parameters to vary across subwatersheds improved model performance and the variations were consistent with different subwatershed characteristics.

Annual phosphorus loads increased between 2001 and 2009 and decreased afterward, with the trend strongest in agricultural areas. Phosphorus yields were highest in Canadian agricultural areas and the U.S. watershed was dominated by point sources, primarily from Great Lakes Water Authority treatment facility (Table 6 and Figure 8). Field-scale analysis used to identify areas within the Canadian agricultural and U.S. urban landscapes with relatively high P yield from nonpoint sources point to where agricultural and urban management practices should be focused.

The main limitations of this study are the lack of some input data at the modeled scale and the relatively low number of water quality observations for calibration and validation. These limitations increased uncertainties in water quality calibration and validation results, and outputs at the field scale. More spatially explicit input data for nutrient inputs (fertilizer and manure application rates, soil nutrient content, etc.), agricultural practices (tillage, tile drainage, cover crop, filter strip in agricultural fields), and water quality observations would increase confidence of representations of nutrient and sediment estimates at both the field scale and stream mouths.

SUPPORTING INFORMATION

Additional supporting information may be found online under the Supporting Information tab for this article:

ACKNOWLEDGMENTS

This work was funded by the Fred A and Barbara M Erb Family Foundation grant number 903. We appreciate the insights and advice offered by Lynn Vaccaro and Jennifer Read from the University of Michigan, and Kevin McKague from Ontario Ministry of Agriculture, Food and Rural Affairs. We also gratefully acknowledge Debbie Burniston and Alice Dove from Environment and Climate Change Canada, Kevin McKague from Ontario Ministry of Agriculture, Food and Rural Affairs, Katie Stammler from the Essex Region Conservation Authority, Majid Khan and Catherine Willey from the Great Lakes Water Authority, Wanhong Yang from University of Guelph, Pamela Joose from Agriculture and Agri-Food Canada, and Karen Maaskant from Upper Thames River Conservation Authority for providing data.

LITERATURE CITED

AAFC (Agriculture and Agri-Food Canada). 2016. “Annual Crop Inventory.” https://open.canada.ca/data/en/dataset/ba2645d5-4458-414d-b196-6303ac06c1c9.

Abbaspour, K.C. 2015. *SWATCalibration and Uncertainty Programs. A User Manual*. Dubendorf, Switzerland: Eawag.

Arnold, J.G., J.R. Kiniry, R. Srinivasan, J.R. Williams, S.L. Haney, and S.L. Neitsch. 2012. *Soil and Water Assessment Tool Input/Output Documentation, Version 2012*. TR-439. Temple, TX: Texas Water Resources Institute.

Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams. 1998. “Large Area Hydrologic Modeling and Assessment Part I: Model Development.” *Journal of the American Water Resources Association* 34: 73–89. https://doi.org/10.1111/j.1752-1688.1998.tb05961.x.

Bai, J., Z. Shen, and T. Yan. 2017. “A Comparison of Single- and Multi-Site Calibration and Validation: A Case Study of SWAT in the Miyun Reservoir Watershed, China.” *Frontiers of Earth Science* 11: 592–600. https://doi.org/10.1007/s11707-017-0656-x.

Baker, D.B., R. Confesor, D.E. Ewing, L.T. Johnson, J.W. Kramer, and B.J. Merryfield. 2014. “Phosphorus Loading to Lake Erie from the Maumee, Sandusky and Cuyahoga Rivers: The Importance Of Bioavailability.” *Journal of Great Lakes Research* 40: 502–17. https://doi.org/10.1016/j.jglr.2014.05.001.

Beale, E. 1962. “Some Uses of Computers in Operational Research.” *Industrielle Organisation* 31 (1): 27–28.

Bocaniov, S.A., and D. Scavia. 2018. “Nutrient Loss Rates in Relation to Transport Time Scales in a Large Shallow Lake (Lake St. Clair, USA — Canada): Insights from a Three-Dimensional Model.” *Water Resources Research* 54: 3825–40. https://doi.org/10.1029/2017WR021876.

Booth, D.B. 1990. "Stream-Channel Incision Following Drainage-Basin Urbanization." *Journal of the American Water Resources Association* 26: 407–17. https://doi.org/10.1111/j.1752-1688.1990.tb01380.x.

Canadian Field Print Initiative. 2017. http://fieldprint.ca/fertilizer-use-survey/.

Craul, P.J. 1985. "A Description of Urban Soils and Their Desired Characteristics." *Arboric* 11: 330–9.

Daniel, E.B. 2011. “Watershed Modeling and its Applications: A State-of-the-Art Review.” *The Open Hydrology Journal* 5: 26–50. https://doi.org/10.2174/1874378101105010026.

Dolan, D.M., A.K. Yui, and R.D. Geist. 1981. “Evaluation of River Load Estimation Methods for Total Phosphorus.” *Journal of Great Lakes Research* 7: 207–14. https://doi.org/10.1016/S0380-1330(81)72047-1.

Gassman, P.W., C. Balmer, M. Siemers, and R. Srinivasan. 2014. “The SWAT Literature Database: Overview of Database Structure and Key SWAT Literature Trends.” Proceedings of the 2014 International SWAT Conference, July 28–1 August, Pernambuco, Brazil, Texas Water Resources Institute Technical Report — TR472.

Gassman, P.W., M.R. Reyes, C.H. Green, and J.G. Arnold. 2007. “The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions.” *Transactions of the American Society of Agricultural Engineers* 50: 1211–50. https://doi.org/10.13031/2013.23637.

GLWQA (Great Lakes Water Quality Agreement). 2016. “The United States and Canada Adopt Phosphorus Load Reduction Targets to Combat Lake Erie Algal Blooms.” https://binational.net/2016/02/22/finalptargets-ciblesfinalesdep/.

Hirsch, R.M. 1979. “An Evaluation of Some Record Reconstruction Techniques.” *Water Resources Research* 15: 1781–90. https://doi.org/10.1029/WR015i006p01781.

Hirsch, R.M., D.L. Moyer, and S.A. Archfield. 2010. “Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs.” *Journal of the American Water Resources Association* 46: 857–80. https://doi.org/10.1111/j.1752-1688.2010.00482.x.

HYDAT. 2016. “Hydrometric Data, National Water Data Archive, Water Survey of Canada.” https://www.canada.ca/en/environment-climate-change/services/water-overview/quantity/monitoring/survey/data-products-services/national-archive-hydat.html.

IJC (International Join Commission). 2009. “The International Watershed Initiative. Implementing a New Paradigm for Transboundary Basins.” http://www.ijc.org/files/publications/ID1627.pdf.

IJC (International Join Commission). 2015. “The International Watershed Initiative. From Concept to Cornerstone of the International Joint Commission. A Watershed Approach for Coordinated Stewardship of Shared Canada-U.S. Waters.” http://ijc.org/files/tinymce/uploaded/IWI/IJC-IWI-EN-WEB.pdf.

IPNI (International Pant Nutrition Institute). 2016. “Nutrient Use Geographic Information System.”http://nugis.ipni.net/About%20NuGIS/.

Jakeman, A.J., R.A. Letcher, and J.P. Norton. 2006. “Ten Iterative Steps in Development and Evaluation of Environmental Models.” *Environmental Modelling & Software* 21: 602–14. https://doi.org/10.1016/j.envsoft.2006.01.004.

Jim, C.Y. 1998. "Physical and Chemical Properties of a Hong Kong Roadside Soil in Relation to Urban Tree Growth." *Urban Ecosystems* 2: 171–81.https://doi.org/10.1023/A:1009585700191.

Johnston, R., and V. Smakhtin. 2014. “Hydrological Modeling of Large River Basins: How Much is Enough?” *Water Resources Management* 28: 2695–730. https://doi.org/10.1007/s11269-014-0637-8.

Kalcic, M.M., I. Chaubey, and J. Frankenberger. 2015. “Defining Soil and Water Assessment Tool (SWAT) Hydrologic Response Units (HRUs) by Field Boundaries.” *Biological Engineering* 8: 12.

Kalcic, M.M., C. Kirchhoff, N. Bosch, R.L. Muenich, M. Murray, J. Griffith Gardner, and D. Scavia. 2016. “Engaging Stakeholders to Define Feasible and Desirable Agricultural Conservation in Western Lake Erie Watersheds.” *Environmental Science & Technology* 50: 8135–45. https://doi.org/10.1021/acs.est.6b01420.

Leta, O.T., A. van Griensven, and W. Bauwens. 2017. “Effect of Single and Multisite Calibration Techniques on the Parameter Estimation, Performance, and Output of a SWAT Model of a Spatially Heterogeneous Catchment.” *Journal of Hydrologic Engineering* 22: 05016036. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001471.

Li, X., D.E. Weller, and T.E. Jordan. 2010. “Watershed model calibration using multi-objective optimization and multi-site averaging.” *Journal of Hydrology* 380: 277–88. https://doi.org/10.1016/j.jhydrol.2009.11.003.

Li, Z., Q. Shao, Z. Xu, and X. Cai. 2010. “Analysis of Parameter Uncertainty in Semi-Distributed Hydrological Models Using Bootstrap Method: A Case Study of SWAT Model Applied to Yingluoxia Watershed in Northwest China.” *Journal of Hydrology* 385: 76–83. https://doi.org/10.1016/j.jhydrol.2010.01.025.

Maccoux, M.J., A. Dove, S.M. Backus, and D.M. Dolan. 2016. “Total and Soluble Reactive Phosphorus Loadings to Lake Erie: A Detailed Accounting by Year, Basin, Country, and Tributary.” *Journal of Great Lakes Research* 42: 1151–65. https://doi.org/10.1016/j.jglr.2016.08.005.

Madani, K., and M.A. Mariño. 2009. “System Dynamics Analysis for Managing Iran’s Zayandeh-Rud River Basin.” *Water Resources Management* 23: 2163–87. https://doi.org/10.1007/s11269-008-9376-z.

Mckinney, D.C., X. Cai, L.S. Lasdon, U.S. Agency, I. Development, D.C. Mckinney, X. Cai, and L.S. Lasdon. 1999. “Integrated Water Resources Management Model for the SYR DARYA Basin.” Prepared by Central Asia Mission, U. S. Agency for International Development.

Mirchi, A., D. Watkins, and K. Madani. 2009. “Modeling for Watershed Planning, Management, and Decision Making.” Watersheds: Management, Restoration and Environmental Impact, Chapter 6. ISBN: 978-1-61668-667-3.

Moriasi, D.N., J.G. Arnold, M.W. Van Liew, R.L. Bingner, R.D. Harmel, and T.L. Veith. 2007. “Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations.” *Transactions of the American Society of Agricultural Engineers* 50: 885–900. https://doi.org/10.13031/2013.23153.

Muenich, R.L., M.M. Kalcic, J. Winsten, K. Fisher, M. Day, G. O’Neil, Y.-C. Wang, and D. Scavia. 2017. “Pay-for-Performance Conservation Using SWAT Highlights Need for Field-Level Agricultural Conservation.” *Transactions of the ASABE* 60: 1925–37,https://doi.org/10.13031/trans.12379.

NOAA-GHCN (National Oceanic and Atmospheric Administration — Global Historical Climatology Network). 2016. ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/.

OMAFRA (Ontario Ministry of Agriculture, Food and Rural Affairs). 2016. http://www.omafra.gov.on.ca/english/stats/welcome.html.

PWQMN (Provincial (Stream) Water Quality Monitoring Network). 2016. Ontario. https://www.ontario.ca/data/provincial-stream-water-quality-monitoring-network.

Ryberg, K.R., and Vecchia, A.V.2017. “waterData: Retrieval, Analysis, and Anomaly Calculation of Daily Hydrologic Time Series Data. R package version 1.0.8. https://CRAN.R-project.org/package=waterData.

Scavia, D., S. Baconiov, A. Dagnew, C. Long, and Y. Wang. 2019. “Interaction of Climate Change and Monitoring Protocols Can Influence Approaches to Nutrient Load Management: The St. Clair — Detroit River System and Lake Erie.” *Journal of Great Lakes Research* 45: 40–49. https://doi.org/10.1016/j.jglr.2018.11.008.

Scavia, D., J. David Allan, K.K. Arend, S. Bartell, D. Beletsky, N.S. Bosch, S.B. Brandt, et al. 2014. “Assessing and Addressing the Re-Eutrophication of Lake Erie: Central Basin Hypoxia.” *Journal of Great Lakes Research* 40: 226–46. https://doi.org/10.1016/j.jglr.2014.02.004.

Scavia, D., J.V. DePinto, and I. Bertani. 2016. “A Multi-Model Approach to Evaluating Target Phosphorus Loads for Lake Erie.” *Journal of Great Lakes Research* 42: 1139–50. https://doi.org/10.1016/j.jglr.2016.09.007.

Scavia, D., M. Kalcic, R. Muenich, J. Read, N. Aloysius, I. Bertani, C. Boles, et al. 2017. "Multiple Models Guide Strategies for Agricultural Nutrient Reductions." *Frontiers in Ecology and the Environment* 15: 126-32. https://doi.org/10.1002/fee.1472.

Sheelanere, P., B.F. Noble, and R.J. Patrick. 2013. “Institutional Requirements for Watershed Cumulative Effects Assessment and Management: Lessons from a Canadian Trans-Boundary Watershed.” *Land Use Policy* 30: 67–75. https://doi.org/10.1016/j.landusepol.2012.03.001.

Shi, P., X. Ma, Y. Hou, Q. Li, Z. Zhang, S. Qu, C. Chen, T. Cai, and X. Fang. 2013. “Effects of Land-Use and Climate Change on Hydrological Processes in the Upstream of Huai River, China.” *Water Resources Management* 27: 1263–78. https://doi.org/10.1007/s11269-012-0237-4.

Singh, V.P., and D.K. Frevert. 2006. “*Watershed Models*.” Boca Raton: Taylor & Francis; CRC Press.

STATCAN (Statistics Canada). 2016. “Canadian Fertilizer Shipments Survey.” https://www150.statcan.gc.ca/n1/en/type/data?text=001-0066.001-0069.

Teshager, A.D., P.W. Gassman, S. Secchi, J.T. Schoof, and G. Misgna. 2016. “Modeling Agricultural Watersheds with the Soil and Water Assessment Tool (SWAT): Calibration and Validation with a Novel Procedure for Spatially Explicit HRUs.” *Environmental Management* 57: 894–911. https://doi.org/10.1007/s00267-015-0636-4.

USDA-ERS (U.S. Department of Agriculture — Economic Research Service). 2016. https://data.ers.usda.gov/reports.aspx?ID=17882.

USDA-NASS (U.S. Department of Agriculture — National agricultural Statistics Services). 2016. “CropScape — Cropland Data Layer.” https://nassgeodata.gmu.edu/CropScape/.

USDA-NRCS (U.S. Department of Agriculture — Natural Resources Conservation Service). 2017. “Web Soil Survey.” https://websoilsurvey.sc.egov.usda.gov/App/WebSoilSurvey.aspx.

U.S. Census Bureau. 2016. "TIGER/Line". https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2015&layergroup=Roads.

U.S. Environmental Protection Agency. 2017. ECHO. https://echo.epa.gov/resources/general-info/loading-tool-modernization.

USGS (U.S. Geological Survey). 2016. “The National Map.” https://viewer.nationalmap.gov/basic/#startUp.

USGS-NWIS (U.S. Geological Survey — National Water Information System). 2016. “Water Data for the Nation.” https://waterdata.usgs.gov/nwis.

UTCA (Upper Thames River Conservation Authority). 2017. “Surface Water and Groundwater Studies.” http://thamesriver.on.ca/watershed-health/surfacewater-groundwater-studies/.

Wang, S., Z. Zhang, G. Sun, P. Strauss, J. Guo, Y. Tang, and A. Yao. 2012. “Multi-Site Calibration, Validation, and Sensitivity Analysis of the MIKE SHE Model for a Large Watershed in Northern China.” *Hydrology and Earth System Sciences* 16: 4621–32. https://doi.org/10.5194/hess-16-4621-2012.

Wiesner, S., A. Gröngröft, and F. Ament. 2016. "Spatial and Temporal Variability of Urban Soil Water Dynamics Observed by a Soil Monitoring Network." *Journal of Soils and Sediments* 16: 2523–37. https://doi.org/10.1007/s11368-016-1385-6.

WQP (National Water Quality Monitoring Council – Water Quality Portal). 2016. https://www.waterqualitydata.us/portal/.

Yang, J.-L., and G.-L. Zhang. 2015. “Formation, Characteristics and Eco-Environmental Implications of Urban Soils — A review.” *Soil Science and Plant Nutrition* 61: 30–46. https://doi.org/10.1080/00380768.2015.1035622.

Zhang, X., R. Srinivasan, and M. Van Liew. 2008. “Multi-Site Calibration of the SWAT Model for Hydrologic Modeling.” *Transactions of the American Society of Agricultural Engineers* 51: 2039–49. https://doi.org/10.13031/2013.25407.