HYDROLOGICAL PROCESSES
Hydrol. Process. (2015)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10536

Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin†

Prasad Daggupati,[1]* Haw Yen,[2,3] Michael J. White,[3] Raghavan Srinivasan,[1] Jeffrey G. Arnold,[3] Conor S. Keitzer[4] and Scott P. Sowa[5]

[1] *Spatial Science Laboratory, Texas A&M University, Suite 221, 1500 Research Plaza, College Station, TX 77843, USA*
[2] *Blackland Research and Extension Center, Texas A&M AgriLife Research, 720 East Blackland Road, Temple, TX 76504, USA*
[3] *Grassland, Soil and Water Research Laboratory, USDA-ARS, 808 East Blackland Road, Temple, Texas 76502, USA*
[4] *Aquatic Ecology Lab, Department of Evolution, Ecology and Organismal Biology, The Ohio State University, 1314 Kinnear Rd, Columbus, OH 43212, USA*
[5] *The Nature Conservancy, Michigan Field Office, 101 East Grand River Ave, Lansing, MI 48906, USA*

Abstract:

Watershed simulation models are used extensively to investigate hydrologic processes, landuse and climate change impacts, pollutant load assessments and best management practices (BMPs). Developing, calibrating and validating these models require a number of critical decisions that will influence the ability of the model to represent real world conditions. Understanding how these decisions influence model performance is crucial, especially when making science-based policy decisions. This study used the Soil and Water Assessment Tool (SWAT) model in West Lake Erie Basin (WLEB) to examine the influence of several of these decisions on hydrological processes and streamflow simulations. Specifically, this study addressed the following objectives (1) demonstrate the importance of considering intra-watershed processes during model development, (2) compare and evaluated spatial calibration *versus* calibration at outlet and (3) evaluate parameter transfers across temporal and spatial scales. A coarser resolution (HUC-12) model and a finer resolution model (NHDPlus model) were used to support the objectives. Results showed that knowledge of watershed characteristics and intra-watershed processes are critical to produced accurate and realistic hydrologic simulations. The spatial calibration strategy produced better results compared to outlet calibration strategy and provided more confidence. Transferring parameter values across spatial scales (i.e. from coarser resolution model to finer resolution model) needs additional fine tuning to produce realistic results. Transferring parameters across temporal scales (i.e. from monthly to yearly and daily time-steps) performed well with a similar spatial resolution model. Furthermore, this study shows that relying solely on quantitative statistics without considering additional information can produce good but unrealistic simulations. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS intra-watershed processes; spatial calibration; spatial and temporal scale parameter transfer; SWAT; NHD plus model

Received 6 October 2014; Accepted 2 May 2015

INTRODUCTION

Watershed simulation models are increasingly used to investigate policy relevant issues related to hydrologic topics, landuse changes, climate change impacts, pollutant load assessments and best management practices (BMPs) (Stone *et al*., 2001; Veith *et al*., 2003; Kannan *et al*., 2005; Benham *et al*., 2006; Tuppad *et al*., 2010; Daggupati *et al*., 2011; Knisel and Douglas-Mankin, 2012; Jha and Gassman, 2014). Model practitioners have the responsibility to make a number of critical decisions in model development, calibration and validation to ensure that the model accurately simulates real world conditions (Eckhardt *et al*., 2005; Zhang *et al*., 2009; Arnold *et al*., 2012a; Arnold *et al*., 2015).

A key step in model development (building a model for a watershed or study region) is that model practitioners should have a good understanding of the watershed characteristics and processes being simulated and should represent them appropriately in the model. Incorporating knowledge about watershed characteristics and processes from literature sources and expert opinion can ensure that models are spatially capturing the hydrological processes and water balance within reasonable limits and realistically simulating real world conditions (Seibert and McDonnell, 2002; Arnold *et al*., 2015). For example, Yen *et al*. (2014a,b) demonstrated that considering intra-watershed characteristics and processes produced accurate spatial and temporal results

*Correspondence to: Prasad Daggupati, Spatial Science Laboratory, Texas A&M University, Suite 221, 1500 Research Plaza, College Station, TX 77843, USA.
E-mail: pdaggupati@tamu.edu
†USDA is an equal opportunity employer and provider.

that enable the model to provide the right answer for the right reasons.

In addition to considering watershed characteristics and processes during model development, decisions about how models are calibrated further influence the ability of the model to produce relatively more realistic results (Eckhardt *et al.*, 2005; White and Chaubey, 2005; Moriasi *et al.*, 2007; Zhang *et al.*, 2009; Arnold *et al.*, 2012a). Calibration of a watershed simulation model at a single site (generally the outlet of a watershed) remains a widely used calibration strategy (Cao *et al.*, 2006; Wang *et al.*, 2012). This approach is best used in small watersheds with fairly uniform characteristics (e.g. soil, slope, vegetation, meteorology). The use of a single site to calibrate large watersheds may result in calibrated parameters which, (1) represent an average of characteristics over the entire watershed, or (2) present a combination of over- or underestimated values which result in poor intra-watershed spatial accuracy. This may be undesirable for simulations of larger watersheds that are more spatially heterogeneous. In these cases, spatial calibration with additional sites is recommended because larger watersheds may contain varied, complex physical characteristics (Qi and Grunwald, 2005; Piniewski and Okruszko, 2011; Cho *et al.*, 2013; Daggupati *et al.*, 2015). This process better accounts for spatial biophysicochemical variations and reduces the problem of non-unique solutions because fewer parameter sets would satisfy calibration criteria at all sites.

After calibration, the model has to be validated to demonstrate that a given site-specific calibrated model can make sufficiently accurate simulations in a new modeling situation. Several studies have focused on transferring parameters temporally (from one time period (e.g. 1990 to 1999) to another (e.g. 2000 to 2010) (Bingner *et al.*, 1997; Van Liew and Garbrecht, 2003; Abbaspour *et al.*, 2007; Chaubey *et al.*, 2010; Sheshukov *et al.*, 2011; Douglas-Mankin *et al.*, 2013; Seo *et al.*, 2014) and spatially (from gauged to ungauged watershed) (Vandewiele and Elias, 1995; Xu, 1999; Santhi *et al.*, 2001; Merz and Blöschl, 2004; Santhi *et al.*, 2008; Parajuli *et al.*, 2009; He *et al.*, 2011; Kumar *et al.*, 2013a,b) to validate the performance of the model. However, little is known about model performance when calibrated parameters are transferred across temporal and spatial scales. For example, how would the model perform if the parameters are transferred across a temporal scale, i.e. from one time-step (e.g. monthly) to another time-step (e.g. daily)? Or how would the model perform if parameters are transferred across spatial scales such as a coarser resolution model to a finer resolution model within the same watershed? Transferring original parameters across spatial or temporal scales might be one way to save on time without sacrificing model performance. An attempt was made by Troy *et al.* (2008) to evaluate the effects of parameter transfer across spatial and temporal scales using a global land surface model known as Variable Infiltration Capacity (VIC). VIC was used to model the entire continental United States, and the results suggested that the transfer of parameters across temporal scales performed better than transfer across spatial scales. Troy *et al.* (2008) also emphasized the need for more studies using hydrologic models to determine if transferring parameters across scales is a viable validation option for producing realistic results.

This paper focuses on understanding how decisions in model development, calibration and validation influence model realism and ensure that watershed simulations provide the information needed to support science-based policy decisions. This research was motivated by the need to provide a realistic hydrologic simulation for a large watershed at a finer spatial resolution to inform policy decisions in the West Lake Erie Basin (WLEB). The major objectives of this study were to utilize the Soil and Water Assessment Tool (SWAT) model in the WLEB to (1) demonstrate the importance of considering intra-watershed processes during model development (2) Compare and evaluate spatial calibration versus calibration at an outlet and (3) evaluate parameter transfers across temporal and spatial scales. In this study, we developed a SWAT model at the 12-digit Hydrologic Unit Code (HUC-12) resolution and another at the National Hydrography Dataset (NHDPlus) resolution. Objective 1 and 2 were examined using the HUC-12 resolution model, and objective 3 was examined using both the HUC-12 and the NHDPlus resolution model.

STUDY AREA

The WLEB watershed drains 28 330 km^2 encompassing the Maumee River, Sandusky River to the south and the Raisin River in the north (Figure 1). There are over 23 000 km of natural and man-made streams in the watershed, which covers portions of Indiana (17%), Michigan (17%) and Ohio (76%). Other major rivers include the Portage, Sandusky, Blanchard, Auglaize, St. Marys, St. Joseph and Tiffin. There is little topography, with elevation ranging from 246 m to 387 m and an average slope of 2%. Average annual precipitation ranges from 838 to 940 mm.

Prior to European settlement, the watershed primarily consisted of Beech, Maple, Ash and Elm forests (Sears, 1941). The Great Black Swamp, a large wetland (>3800 km^2) located centrally in the watershed, was a major landscape feature (Kaatz, 1955). Widespread forest clearing and wetland draining began in the mid-19th century (Kaatz, 1955). The watershed is now predominantly agricultural, with more than 70% of the land in

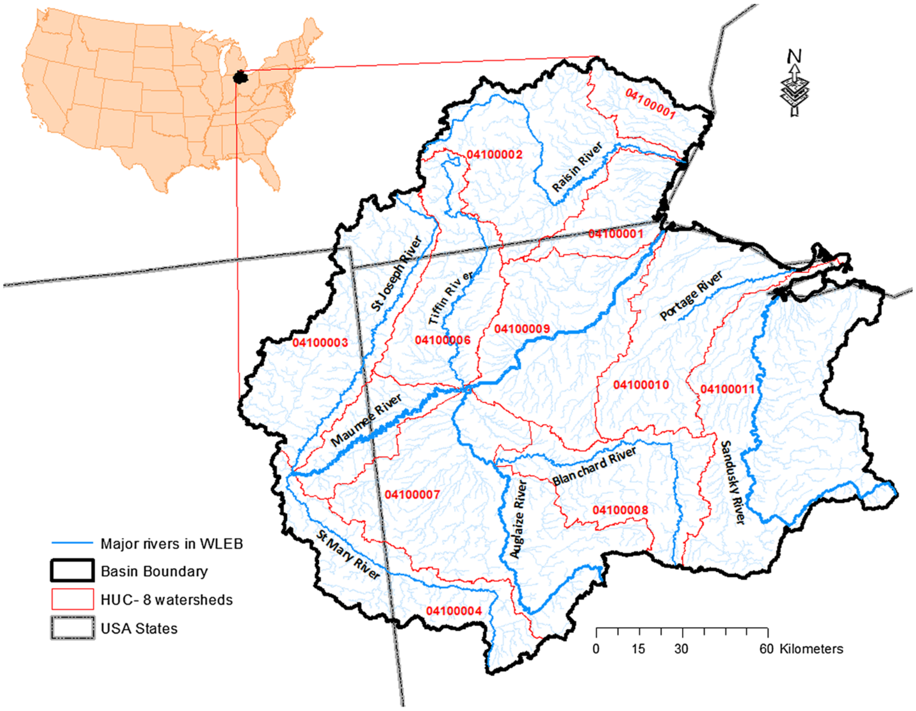

Figure 1. The study area—West Lake Erie Basin (WLEB) with major rivers and HUC-8 watershed boundaries. The insert shows continental United States with the study area shaded

cultivated cropland, the majority of which is in corn–soybean crop rotations. Tile drainage is used extensively throughout the watershed. The next most dominant land uses, forested and urban land use, each make up about 12% of the watershed.

The widespread conversion of native vegetation to agriculture and associated drainage practices (e.g. stream channelization) have degraded freshwater habitat quality and negatively affected freshwater biodiversity (Trautman, 1939; Trautman and Gartman, 1974; Karr *et al*., 1985). Additionally, the Maumee River appears to be a major contributor to eutrophication and the recent increase in harmful algal blooms in Lake Erie (Kane *et al*., 2014). These freshwater conservation and human health concerns require a finer resolution hydrologic model that realistically simulates hydrologic processes to allow policy makers to make informed decisions related to improve conditions in the WLEB.

DATA INPUTS AND MODEL SETUP

The latest version of SWAT, ArcSWAT 2012 (rev 593) for ArcGIS10.1 Geographic Information System interface, was used to set up the SWAT model. The SWAT model is a continuation of nearly 30 years of modeling efforts by the USDA Agricultural Research Service (ARS) and is widely used, watershed-scale, process-based model (Gassman *et al*., 2007; Douglas-Mankin *et al*., 2010; Arnold *et al*., 2012a). The model is supported by online documentation (Neitsch *et al*., 2011; Arnold *et al*., 2012b) which reviews all processes simulated with the model. The ArcSWAT interface allows importing predefined watershed boundaries and streams along with automatic delineation of streams and subwatersheds (Luo *et al*., 2011). This function was employed to develop two models, a HUC-12 model based on a predefined HUC-12 Watershed Boundary Dataset (WDB) and a more detailed NHDPlus model using the National Hydrography Dataset (NHD) and NHD-plus stream network. The HUC-12 WBD (1: 240 000 scale) was downloaded from http://datagateway.nrcs.usda.gov and is a coordinated effort between the United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), the United States Geological Survey (USGS) and the Environmental Protection Agency (EPA). The NHDPlus data consists of NHD Plus (Version 2) streams and catchments at scale of 1: 100 000 and was downloaded from http://www.horizon-systems.com/NHDPlus/index.php. The NHDPlus framework is a coordinated effort by the EPA Office of Water and the USGS. A 30-m Digital Elevation Model (DEM) was used to define the topographical characteristics for each model. A total of 391 and 11 128 subbasins, respectively, were characterized for HUC-12 and NHD Plus setup in the Western Lake Erie Basin (Figure 2). The average size of the subwatersheds in the HUC-12 model was 72 km^2 (range, 25 to 191) while the average watershed size in the

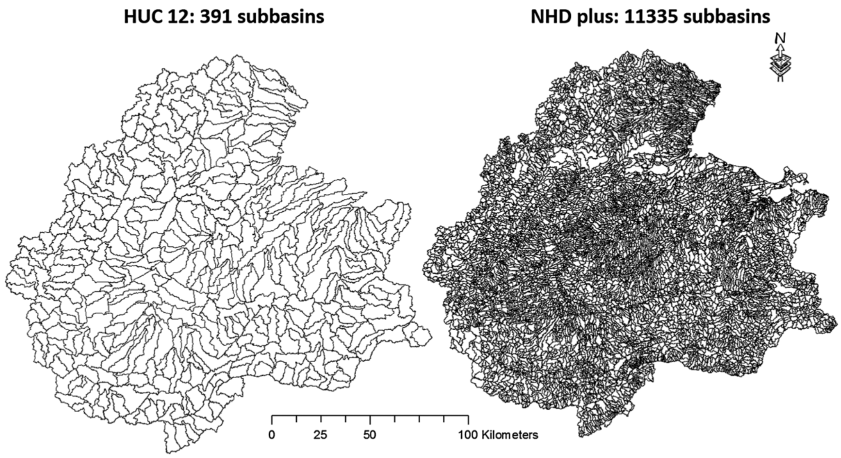

Figure 2. HUC-12 (coarse resolution) and NHDPlus (finer resolution) subwatersheds in WLEB

NHDPlus model was 2.6 km^2 (range, 0.001 to 80). Developing SWAT model at NHDPlus resolution is first of its kind, and no studies were reported in literature that used SWAT at such finer resolution.

Land use was defined using 2010 and 2011 Crop Data Layers (CDLs). The data was processed using techniques recommended by Srinivasan *et al*. (2010) to prepare a single 30-m resolution landuse/landcover layer which includes major crop rotations. Soils were derived from STATSGO (USDA-NRCS, 1995) at a scale of 1: 250 000. All soil properties needed for the SWAT model were extracted from the national STATSGO layer and processed with the ArcSWAT interface.

Land use, soils and slope (derived from DEM) were intersected within each subbasin by ArcSWAT to create unique Hydrologic Response Units (HRUs). Three slope classes (0%–2%, 2%–5% and >5%) were used with landuse, soil and slope (by class) thresholds of 50/50/50 ha. All agricultural crops were exempt from landuse thresholds such that all agricultural crops were included as HRUs. A total of 13 156 and 34 807 HRUs were derived in the WLEB using the HUC-12 and the NHDPlus models.

Both models included daily precipitation and temperature data from 1960 to 2010. This data was derived from the National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer network and Weather-Bureau-Army-Navy stations. Missing data at each station was supplied using an inverse distance weighted interpolation algorithm and observations from the nearest five stations.

Tile drain systems are designed to remove excess field water and lower water tables to reduce crop stresses and allow timely field tillage and planting. However, no clear record of tile locations was available in this basin. It was therefore assumed that tile drains occur in agricultural areas that are located in poorly drained soils and have a slope less than 1%. Poorly drained soils were identified within the basin by processing SSURGO soil using soil data viewer 6.0 program (http://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/soils/home/?cid=nrcs142p2_053620) in ArcGIS.

SWAT management operations (i.e. planting, tillage, harvest and fertilizer application) were assembled from a variety of sources. Operation scheduling was derived from management templates developed by the NRCS for the RUSLE2 model (Foster, 2005). Cropland tillage was derived from Baker (2011). SWAT plant growth-related parameters were developed from local weather statistics. Cropland fertilization was derived from NASS reported county average crop yields. Data was processed and combined into SWAT format management files using software written specifically for this purpose.

Measured streamflow from 12 gauge stations (Figure 3, Table I) was collected from 1 January 1990 to 31 December 2006. The data was used during calibration and validation to facilitate spatial calibration and validation assessments.

METHODS

Intra-watershed processes

The majority of the WLEB is comprised of agricultural land of which more than 85% have tile drainage systems implemented to facilitate artificial drainage and to improve crop yields and field trafficability. Tile drainage is an important and major intra-watershed process in the basin. Representing tile drain in the SWAT model was necessary to accurately capture the spatial hydrological processes and water balance within the watershed. This study evaluated the effects of tile drain to demonstrate the importance of considering intra-watershed processes within the watershed during model development and simulations. The default SWAT model (without

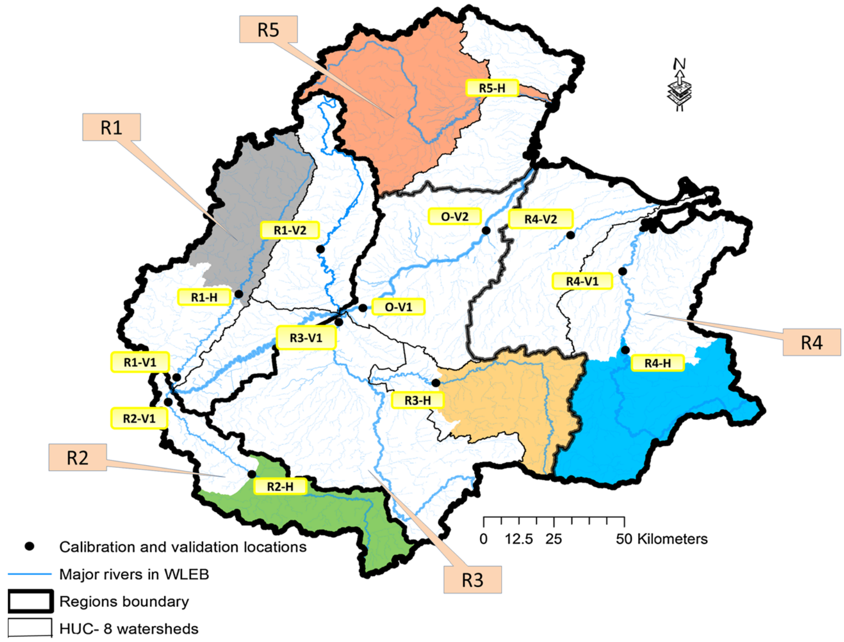

Figure 3. Five regions (separated with thick black border), five head watersheds (represented in shaded colors), and calibration and validation locations in WLEB

calibration) was used to evaluate the effectiveness of tile drains in capturing the overall hydrologic water balance in the watershed. Tile drain parameters including depth to drain (DDRAIN), time to drain (TDRAIN), drain lag time (GDRAIN) and depth to impervious layer (D_IMP) were changed to 1500 (mm), 48 (h), 24 (h) and 1200 (mm), respectively. These tile drain parameters values were based on expert opinions of watershed specialists working in the watershed. The computation of the daily CN value as a function of plant evapotranspiration (ICN = 1) was used because the default soil moisture method (ICN = 0) was predicting too much runoff in shallow soils (Yen *et al.*, 2014c). ICN = 1 (plant-based ET) and ICN COEF = 0.5 along with other tile drain parameters reduced surface runoff and transferred that water as tile flow and thereby simulated tile drains reasonably well in the watershed. The SWAT model was simulated from 1990 to 1999 and a 3-year warm-up period (1987 to 1989) was used prior to the model simulation period as recommended by Daggupati *et al.* (2015). Average annual hydrologic components (surface runoff, ground flow, lateral flow and tile flow) of the water balance along with quantitative statistics and graphical comparisons (discussed in the next section) at R4-H gauge station (Figure 3) were used to assess the performance of simulations with and without tile drains.

Spatial and outlet calibration

The default model after the inclusion of tile drain information (a major intra-watershed process) needed to be calibrated to increase the accuracy of model predictions. A regular calibration procedure where the model is calibrated using observed and simulated data at the outlet may not work well in the WLEB because of the large size of this watershed and potential spatial variability within the basin. In order to capture this spatial variability, spatial calibration is needed. We used a proxy-basin spatial calibration strategy originally proposed by Klemes (1986) and summarized by Daggupati *et al.* (2015). This strategy involves calibration of a model in a gauged watershed and transferring calibrated parameters to nearby or adjacent watersheds within the same eco-region. Further, a spatial validation is performed at various locations to evaluate the performance of model. The logic behind this method is that in a similar eco-region, the climate and watershed conditions vary smoothly over space and the parameters in the region are expected to be similar (Jin *et al.*, 2009).

The WLEB was divided into five different regions (R1, R2, R3, R4 and R5) (Figure 3, Table I). R1, R2 and R3 drain into Maumee River, R4 drains into Sandusky and Cedar-Portage rivers while R5 drains into Raisin and other adjacent tributary rivers. During separation of the regions, landuse, soil, slope and precipitation were used to visualize the spatial variability within the basin. HUC-8 watershed boundaries within the basin were used as guidelines to separate regions. A head watershed (shaded areas in Figure 3) was selected in each region and was calibrated and validated using a temporal split-sample approach in which one period (1990 to 1999) was used for calibration and another period (2000 to 2006) was used for temporal validation. After satisfactory calibration

Table I. Calibration–validation locations, their station name, stationID, source of data collected at each location and corresponding draining HUC-8 watershed with its name. Locations used for calibration and spatial validation are also shown. In the table, R represents region, H represents head watershed location, V represents validation location and O represents outlets

Calibration /validation locations	Station name	StationID	Source	HUC-8 (name)	Region	Calibration	Validation
R1-H	St. Joseph River near Newville, IN	4178000	United States Geologic Survey	04100003 (St. Joseph)	R1	√	
R1-V1	Tiffin River at Stryker, OH	4185000	United States Geologic Survey	04100006 (Tiffin)	R1		√
R1-V2	St. Joseph River	LEJ100-0003	Indiana Department of Environmental Management	04100003 (St. Joseph)	R1		√
R2-H	St. Marys River	LES040-0007	Indiana Department of Environmental Management	04100004 (St. Marys)	R2	√	
R2-V2	St. Marys River	LES060-0004	Indiana Department of Environmental Management	04100004 (St. Marys)	R2		√
O-V	Maumee River at Waterville, OH	4193500	Heidelberg College River Studies	04100009 (Lower Maumee)	R3		√
R3-H	Blanchard River upstream of Ottawa, OH at County Road 8	500100	Ohio Environmental Protection Agency	04100008 (Blanchard)	R3	√	
R3-V1	Auglaize River upstream of Defiance, OH at Harding Road	500290	Ohio Environmental Protection Agency	04100007 (Auglaize)	R3		√
O-V	Maumee River upstream of Independence Dam	P09W19	Ohio Environmental Protection Agency	04100009 (Lower Maumee)	R3		√
R4-V2	Portage River in Woodville, OH at railroad bridge	4195600	Ohio Environmental Protection Agency	04100010 (Cedar-Portage)	R4		√
R4-H	Sandusky River near Mexico, OH	4197000	Ohio Environmental Protection Agency	04100011 (Sandusky)	R4	√	
R4-V1	Sandusky River near Fremont, OH	4198000	Heidelberg College River Studies	04100011 (Sandusky)	R4		√
R5-H	River Raisin near Monroe, MI	4176500	Heidelberg College River Studies	04100002 (Raisin)	R5	√	

results in head watersheds in five regions (based on quantitative statistics and graphical comparisons as discussed later), the parameters were transferred to other watersheds in the region. Spatial validation was performed at various locations (Figure 3, Table I) within the transferred region and also at two outlet locations along the Maumee River. During the process of validating the model in regions, the temporal split-sample approach, as discussed previously, was used to complete a comprehensive evaluation and thereby accomplishing spatial calibration.

During the calibration of head watersheds, SWAT was manually calibrated to make sure that the hydrology, overall water balance and general seasonal patterns within the watershed were captured. Additional automated calibration was done using the Sequential Uncertainty Fitting version-2 (SUFI-2) routine in SWAT-CUP program. A monthly time-step was used during calibration. Quantitative statistics and criteria recommended by Moriasi *et al*. (2007) were used to evaluate the simulation performance. The quantitative statistics applied in this study were Nash–Sutcliffe simulation efficiency (NSE) and percentage bias (PBIAS). The model performance for monthly and daily streamflow can be categorized into four classes according to the threshold NSE, and PBIAS values: very good ($0.75 < \text{NSE} \leq 1.00$, $\text{PBIAS} < \pm 10$); good ($0.65 < \text{NSE} \leq 0.75$, $\pm 10 \leq \text{PBIAS} < \pm 15$); satisfactory ($0.50 < \text{NSE} \leq 0.65$, $\pm 15 \leq \text{PBIAS} < \pm 25$); and unsatisfactory ($\text{NSE} \leq 0.50$, $\text{PBIAS} \geq \pm 25$). Graphical comparisons of time-variable plots of observed and simulated flow provide important insights into model representation of hydrographs, baseflow recession and other pertinent factors often overlooked by quantitative comparisons. In this study, visual comparisons of hydrographs between observed and simulated were evaluated, and the simulation was considered satisfactory only when the shapes (peaks and base flow) of observed and predicted hydrographs were similar.

Outlet calibration strategy was tested by performing an auto calibration using SUFI-2 routine in SWAT-CUP program at a monthly time-step utilizing streamflow at the outlet (O-V2 location, Figure 3). After calibration, the quantitative statistics and graphical comparisons at various locations within the watershed were evaluated. The model performance using spatial and outlet calibration strategies were compared and analysed. In this study, graphical representation of quantitative statistics (only NSE) for spatial and outlet calibration strategies at various locations within WLEB is shown in Figure 4 to have a better view of results spatially and would ensure a more comprehensive evaluation of model performance based on recommendation by Saraswat *et al.* (2015).

Transfer of parameters across scale

Transferring parameters across scales may be desirable as a part of validation option to adapt models to address new issues beyond their original intent. Developing and calibrating a new model to address these new issues may be time consuming. This study investigated the model performance after transferring parameters across spatial and temporal scales to determine if this is a viable option to address novel issues beyond the scope of the original models intent. We transferred parameters from a spatially calibrated and validated HUC-12 model (coarser resolution model) at a monthly time-step to NHDPlus model (finer resolution model) to evaluate the impacts of transferring parameters across spatial scale. Next, in the HUC-12 model, we transferred parameters to daily and yearly time-step to evaluate the temporal scale effects where the model was calibrated on monthly time-step. Also, the temporal scale effects in the NHDPlus model were also evaluated on daily and yearly time-step after the transfer of parameters from a monthly time-step calibrated HUC-12 model. In both the cases, quantitative statistics and graphical criteria (temporal time series plots) at the outlet (O-V2) and one another location (R4-H)

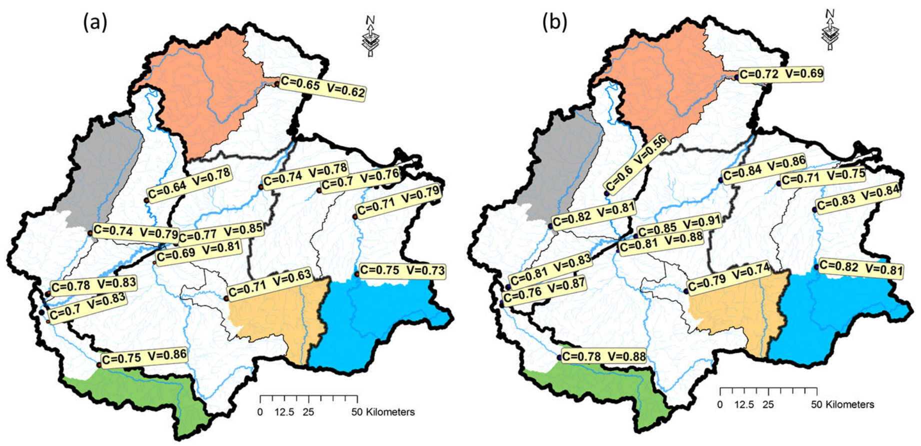

Figure 4. NSE values at various locations in the basin using a) spatial calibration strategy and b) outlet calibration strategy

Table II. Average annual hydrologic components for S1, S2 and S3 model simulations for the time period of 1990 to 1999 in WLEB

Average annual hydrologic components	S1 (mm)	S2 (mm)	S3 (mm)
% surface runoff (surface runoff/water yield)	71%	11%	28%
% tile flow tile flow/water yield	0%	60%	53%
% ground water (ground water/water yield)	28%	28%	15%
% later flow (lateral/water yield)	1%	1%	3%

Table III. Quantitative statistics for S1, S2 and S3 at R4-H location

Intra-watershed processes	R2	NS	Median simulated	Median observed	PBIAS
Without tile (S1)	0.61	0.60	20.22	14.43	−6.95
After tile (S2)	0.69	0.27	31.71	14.43	−57.29
After calibration/ with tile (S3)	0.83	0.82	13.43	14.21	−4.60

were used to evaluate the impacts of parameter transfer across scale.

RESULTS AND DISCUSSION

Impacts of considering intra-watershed processes

Scenario 1 (S1) is used to denote SWAT model simulation without tile drain and S2 for SWAT model with tile drain. Scenario 3 (S3) which is a SWAT model simulation after calibration is also used for comparison and discussion purposes. However, more discussion on calibration is given later. Average annual hydrologic components of the water balance in WLEB are shown in Table II for S1, S2 and S3 model simulations. Quantitative statistics are presented in Table III for all three scenarios using R4-H gauge location (Sandusky River near Mexico OH) which is in Sandusky watershed (HUC8, 4100011) and is heavily dominated by agricultural land (>83%) and is mostly implemented with tile drains. Graphical comparisons are shown in Figure 5. The surface runoff contribution without tile drainage (S1) was 71% (ratio of surface runoff to total water yield),

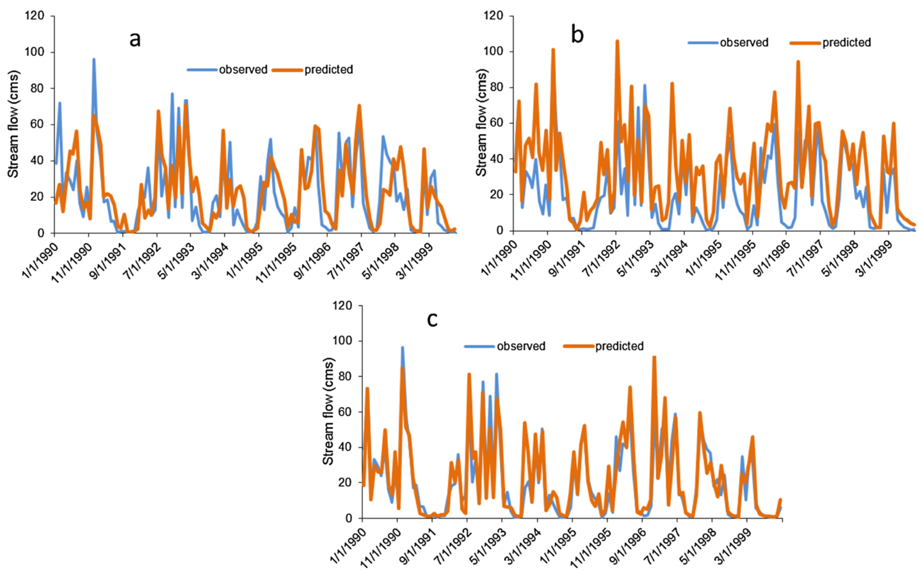

Figure 5. a) Monthly time series comparison in S1, S2 and S3 model simulations for the time period of 1990 to 1999 at R4-H gauge location

Table IV. Quantitative statistics at head watershed locations for Spatial and Outlet calibration strategies

Calibration strategy	Time period	R1-H		R2-H		R3-H		R4-H		R5-H	
		NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS
Spatial calibration	1990–1999	0.82	11.89	0.78	0.09	0.79	4.81	0.82	−4.72	0.72	0.97
	2000–2006	0.81	5.21	0.88	−3.90	0.74	26.19	0.81	−5.81	0.69	−8.82
Outlet calibration	1990–1999	0.74	−3.98	0.75	4.73	0.71	13.72	0.75	−10.72	0.65	−9.15
	2000–2006	0.79	−14.88	0.86	0.21	0.63	25.94	0.73	−15.24	0.62	−19.41

while tile flow contribution was 0% (ratio of tile flow to total water yield) (Table II). Monthly quantitative statistics (Table III) show that the model predictions were satisfactory and reliable when tile flow was not included ($R^2 = 0.61$, NSE = 0.60, PBIAS = −6.95). However, having tile flow contribute 0% to total water yield is unrealistic for this watershed. Communication with the experts that work in the watershed suggested that the tile drain contribution is generally around 50% while the surface runoff contribution is around 30%. Thus, despite being unrealistic, the scenario without tile drainage (S1) produced satisfactory performance statistics (Table III). Graphical comparisons between observed and predicted flow data showed that there were differences in timing and magnitude of peak flows and the shape of recession curves (Figure 4a). This was primarily because of soft data such as tile drain information not being included (a major intra-watershed processes in the watershed) during model development. When tile drainage was included without calibration (S2) the surface runoff contribution was 11%, and the tile flow was 60% (Table II). The predictions in S2 were more realistic and close to the opinion of watershed experts. Graphical comparisons of observed and predicted flow showed that the timing and magnitude of peak flows and the shape of recession curves aligned better; however, the predicted flows were higher than observed (Figure 4b). The quantitative statistics were poor ($R^2 = 0.69$, NSE = 0.27, PBIAS = −57.29) despite more realistic contributions of tile flow, mainly because the predicted flows were higher (median observed = 14.43, median simulated = 31. 71) compared to observed flows. This is because of the inclusion of tile drains, which altered the hydrodrological processes and needs calibration to lower predicted flows and align better to improve statistics. When tile drainage was included and the model was calibrated (S3), surface runoff and tile flow contributed 28% and 53%, respectively. These contributions were close to the opinion of watershed experts. Graphical comparison (Figure 4c) shows that the observed and predicted flows align with each other and the quantitative statistics were very good ($R^2 = 0.84$, NSE = 0.82, PBIAS = −4.60). These results showed that using only quantitative performance statistics can be misleading and should not be used alone to make absolute modeling decisions (e.g. Developing Total maximum Daily Loads (TMDLs) or assessing the impacts of best management practices). However, combining quantitative statistics along with graphical comparisons of time series plots and the incorporation of literature or expert knowledge to account for all intra-watershed processes will ensure that hydrological processes and water balance are within reasonable limits and will produce better and more reliable predictions.

Table V. Quantitative statistics at validation locations for Spatial and Outlet calibration strategies

Calibration strategy	Time period	R1-V1		R1-V2		R2-V1		R3-V1		R4-V1		R4-V2		O-V1		O-V2	
		NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS	NSE	PBIAS
Spatial calibration	1990–1999	0.81	17.29	0.60	4.33	0.76	11.96	0.81	8.65	0.83	12.51	0.71	5.34	0.85	9.99	0.84	11.88
	2000–2006	0.83	13.23	0.81	15.05	0.87	7.65	0.88	10.99	0.84	12.68	0.75	17.27	0.91	7.86	0.86	4.42
Outlet calibration	1990–1999	0.78	4.14	0.64	16.46	0.70	15.55	0.69	8.84	0.71	15.26	0.70	9.43	0.77	7.96	0.74	10.27
	2000–2006	0.83	−4.25	0.78	12.10	0.83	10.74	0.81	11.37	0.79	14.98	0.76	10.54	0.85	5.03	0.78	1.88

Influence of spatial and outlet calibration strategy

Spatial calibration strategy. In the head watershed locations, the NSE values ranged from 0.72 to 0.82 (mean, 0.80) and PBIAS values ranged from −4.72 to 11.89 (absolute [abs] mean 4.49) during the calibration period (Table IV). In the temporal validation period, the NSE and PBIAS ranged from 0.69 to 0.88 (mean, 0.78) and −8.82 to 26.19 (abs.mean, 9.98) (Table IV). In the spatial validation locations and in the calibration period, the NSE ranged from 0.60 to 0.85 (mean, 0.78) and PBIAS ranged from 4.33 to 17.29 (abs.mean, 9.17) (Table V). In the temporal validation period, the NSE ranged from 0.75 to 0.91(mean, 0.84) and PBIAS ranged from 4.42 to 17.27 (abs.mean, 9.99) (Table V). Quantitative statistics showed that spatial calibration strategy performed very good (based on NSE and PBIAS criteria) in the headwater watersheds as well as in the spatial validation locations. The performance was even better in the in the spatial validation locations, especially in the temporal validation period (mean NSE, 0.84 *vs* mean NSE, 0.78).

Outlet calibration strategy. In the headwater calibration locations and in the calibration period, the NSE and PBIAS ranged from 0.65 to 0.75 (mean, 0.73) and −10.72 to 13.72 (abs.mean, 8.46) in the headwater subbasins (Table IV). In the validation period, the NSE and PBIAS ranged from 0.62 to 0.86 (mean, 0.73) and −15.24 to 25.94 (abs.mean, 15.14) (Table IV). In the validation locations and in the calibration period, the NSE and PBIAS ranged from 0.64 to 0.78 (mean, 0.72) and 4.14 to 15.26 (abs.mean, 9.84). In the validation period, the NSE ranged from 0.76 to 0.85 (mean, 0.80) and PBIAS ranged from −4.25 to 14.98 (abs.mean, 7.96) (Table V). The quantitative statistics showed that the outlet calibration strategy rated as good based on NSE criteria and very good based on PBIAS criteria in the various locations of the watershed when it was calibrated at the outlet and verified across various locations in the watershed.

Comparing the two strategies showed that the spatial calibration strategy statistics slightly outperforms the outlet calibration statistics in all locations within the basin (based on NSE and PBIAS statistics). It should also be noted that the statistics at the outlet (O-V2) were better for the spatial calibration during calibration and validation periods mainly because the spatial variations within the basin were more realistically captured. However, the outlet calibration strategy still performed well for this watershed. This could be the result of the basin being very homogeneous with a flat topology and agriculture being the most prominent landuse. In addition, the tile drains which are major intra-watershed processes in the basin are represented in the model. Outlet calibration may perform even less well than the spatial calibration strategy for watersheds that have greater variation in topology and land use.

A time series plot (Figure 6) between observed and predicted flow using outlet calibration strategy and spatial calibration strategy at R4-H location showed that the spatial calibration strategy better represented the peaks and recession of the hydrographs. The outlet calibration strategy was either under or over predicting the peaks and recession. This again showed that using quantitative statistics alone may be misleading and unreliable. The use of graphical comparisons along with quantitative statistics resulted in better evaluation criteria. The results of this study showed that the spatial calibration strategy gave greater confidence in modeling efforts.

Effects of transfering parameters across scale

Spatial scale transfer. The monthly NSE values for NHDPlus model at OV2 and R4-H locations were 0.71 and 0.71, while the PBIAS values were 25.25 and 23.95, respectively (Table VI). Quantitative statistics showed that the NHDPlus model was good in the selected locations based on NSE criteria and satisfactory based on PBIAS criteria. Similar statistics were seen in other locations within the basin. Average annual hydrologic components of the water balance were evaluated in the NHDPlus model by comparing contributions of surface runoff and tile flow. Contribution of surface runoff was 29%, while the contribution of tile flow was 54%, which indicated that the intra-watershed processes (tile drainage)

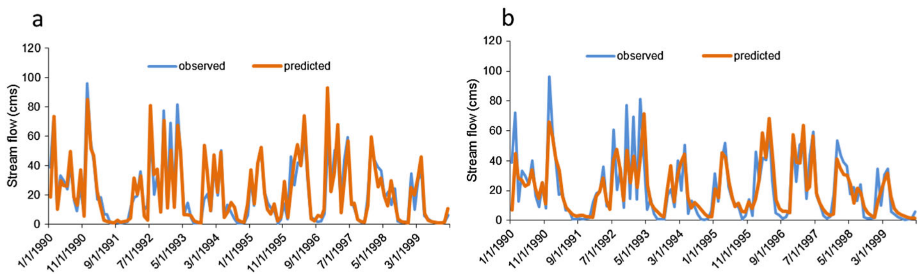

Figure 6. Time series plots at R4-H location using a) spatial calibration strategy and b) outlet calibration strategy

Table VI. NSE and PBIAS values on daily, monthly and yearly at two locations (O-V2 and R4-H) for NHDPlus and HUC12 models

Spatial scale	Location	Temporal scale	NSE	PBIAS
NHD Plus	O-V2	Daily	−0.02	25.31
		Monthly	0.71	25.25
		Yearly	0.01	25.31
	R4-H	Daily	0.16	23.99
		Monthly	0.71	23.95
		Yearly	0.28	23.99
HUC-12	O-V2	Daily	0.70	10.59
		Monthly	0.82	10.53
		Yearly	0.80	10.60
	R4-H	Daily	0.67	−4.58
		Monthly	0.82	−4.60
		Yearly	0.87	−4.58

were captured reasonably well with spatial transfer of parameters. The time series graphs at both locations showed that there were some differences in timing and magnitude of peak flows (simulated data was under predicting most of the time); however, the NHDPlus model captured the pattern reasonably well (Figure 7). The differences are likely because of the very small (36% smaller) subwatershed size and irregular shape in NHDPlus model compared to the HUC-12 model (Figure 2). In smaller sized and irregular shaped subwatersheds, the length of the reach and time of concentration are small. This would result in faster transport of flow and associated constituents within a day or sometimes within hours after a rainfall event. The resulting daily hydrograph will have higher peaks and a very quick receding curve and thereby under or early prediction at monthly level (Figure 7). In HUC-12 model, the time of concentration and length of reach are longer resulting in more days for the flow to transport and the hydrograph captures peaks and recession reasonably well. Fine-tuning of the general parameters (e.g. time of concentration) may be needed for the NHDPlus model, after spatial transfer of parameters, to accurately capture the timing and magnitude of peak flows and the shape of rising and recession

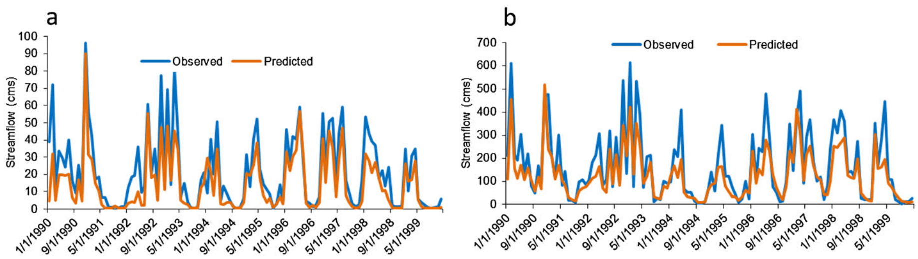

Figure 7. Monthly time series for NHDPlus model at (a) R4-H and (b) OV2 locations

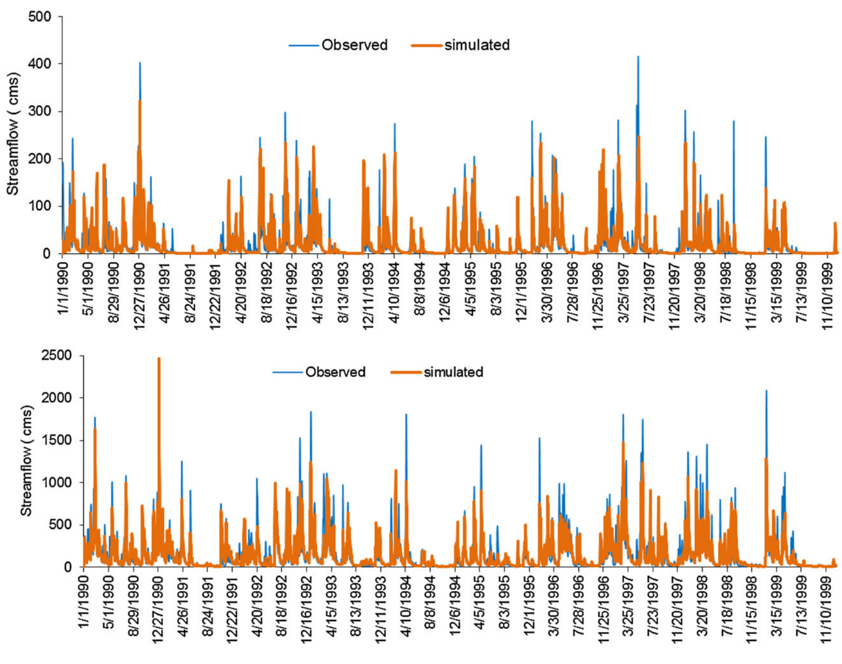

Figure 8. Daily time series plot for HUC-12 model at (a) R4-H (b) OV2 locations

curves and to improve quantitative statistics. Fine tuning of the parameters and reporting associated results is beyond the scope of this paper and will be presented elsewhere (Yen *et al.* in prep).

Temporal scale transfer. The daily NSE and PBIAS values for the HUC-12 model at O-V2 location are 0.70, 0.80 and 10.59 and 10.60, respectively (Table VI). At the R4-H location, the statistics were 0.67, 0.87 and −4.58 and −4.58, respectively. Quantitative statistics showed that the performance of temporal scale transfer was good to very good based on NSE criteria and very good based on PBIAS criteria on daily and monthly time-step when the original model was calibrated at monthly time-step. The time series plots show that the simulated data was under predicting the peaks flows; however, it captured the timing and the shape of rising and recession curves well (Figure 8). General parameters, such as time of concentration, can be adjusted to better capture the peaks on daily time-step. Overall, the performance of temporal scale transfer was good in the HUC-12 model based on quantitative statistics and temporal time series plots. Quantitative statistics were poor for the NHDPlus model for daily and yearly time-step at both OV2 and R4-H locations (Table VI). The poor quantitative statistics are because of poor performance of the NHDPlus model at the monthly time-step as seen above after the spatial transfer of parameters from the HUC-12 model. Performance on daily and yearly time-steps may be improved after fine tuning the general parameters of the NHDPlus model.

CONCLUSIONS

Coarse resolution (HUC-12) and finer resolution (NHDPlus) models were developed within WLEB to demonstrate the significance of considering intra-watershed processes during model development, compare and contrast two calibration strategies (spatial calibration vs. outlet calibration) and evaluate spatial and temporal transfer of parameters.

We found that including intra-watershed processes (i.e. tile drainage) produced accurate and realistic hydrologic simulations. However, failure to include these processes may still result in a model that performs well according to model performance statistics. Thus, considering only model performance statistics may be misleading. The spatial calibration strategy produced better results in terms of quantitative statistics and graphical comparisons. The outlet calibration strategy also produced decent results in various locations within the watershed. This was likely because the WLEB is a fairly homogenous watershed. However, we believe that the spatial calibration strategy results in greater confidence for modeling efforts that support science-based decisions. Transferring parameters across temporal scales worked well with a similar spatial resolution model; however, additional fine tuning is required when transferring parameters across spatial scales to produce realistic results. This study showed that quantitative statistics should be used in conjunction with graphical comparisons and knowledge of the watershed (e.g. literature sources or expert knowledge) to ensure that hydrological processes and water balance are within reasonable limits and will produce better and more reliable predictions.

ACKNOWLEDGEMENTS

We would like to thank the USDA Natural Resource Conservation Service for funding this work through the Wildlife and Cropland components of the Conservation Effects Assessment Project (CEAP). We particularly want to thank Charlie Rewa for his unwavering support of this project through the Wildlife Component CEAP. This study represents a component of the much broader Western Lakes Erie Basin (WLEB) CEAP effort, and we would like to thank the project team for their contributions: Gust Annis, Kim Hall, Matt Herbert, August Froehlich, Stu Ludsin, Lee Norfleet, Maura O'Brien, Jeff Reuter, Anthony Sasson, Bill Stanley, Mari-Vaughn Johnson, Carrie Vollmer-Sanders and Chris Winslow. Finally, we would like to thank our WLEB CEAP Advisory Panel members that have provided critical input, review and data needed to successfully complete this project: Vicki Anderson, Jeff Ankney, Larry Antosch, Dave Baker, Steve Davis, Jane Frankenberger, Gail Hesse, Kevin King, Mary Knapp, Melinda Koslow, Ken Krieger, Greg LaBarge, Kip Miller, Dale Robertson and Steve Shine.

REFERENCES

Abbaspour KC, Yang J, Maximov I, Siber R, Bogner K, Mieleitner J, Zobrist J, Srinivasan R. 2007. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* **333**: 413–430. DOI: 10.1016/j.jhydrol.2006.09.014

Arnold JG, Moriasi DN, Gassman PW, Abbaspour KC, White MJ, Srinivasan R, Santhi C, Harmel RD, van Griensven A, Van Liew MW, Kannan N, Jha MK. 2012a. Swat: model use, calibration, and validation. *Transactions of the ASABE* **55**: 1491–1508.

Arnold JG, Kiniry JR, Srinivasan R, Williams JR, Haney EB, Neitsch SL. 2012b. Soil and water assessment tool input/output file documentation: version 2012. Texas Water Resources Institute TR 436. http://swatmodel.tamu.edu/documentation/ (Accessed on May 2014).

Arnold JG, Youssef MA, Yen H, White MJ, Sheshukov AY, Sadeghi AM, Moriasi DN, Steiner JL, Amatya DM, Skaggs RW, Haney EB, Jeong J, Arabi M, Gowda PH. 2015. Hydrological processes and model representation: impact of soft data on calibration. *Transactions of the ASABE* (in Press).

Baker N. 2011. Tillage practices in the conterminous United States, 1989–2004—Datasets Aggregated by Watershed, U.S. Geological Survey Data Series 573.

Benham BL, Baffaut C, Zeckoski RW, Mankin KR, Pachepsky YA, Sadeghi AM, Brannan KM, Soupir ML, Habersack MJ. 2006. Modeling bacteria fate and transport in watersheds to support TMDLs. *Transactions of the ASABE* **49**: 987–1002.

Bingner RL, Garbrecht J, Arnold JG, Srinivasan R. 1997. Effect of watershed subdivision on simulation runoff and fine sediment yield. *Transactions of the American Society of Agricultural Engineers* **40**: 1329–1335.

Cao W, Bowden WB, Davie T, Fenemor A. 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes* **20**: 1057–1073. DOI: 10.1002/hyp.5933

Chaubey I, Chiang L, Gitau MW, Mohamed S. 2010. Effectiveness of best management practices in improving water quality in a pasture-dominated watershed. *Journal of Soil and Water Conservation* **65**: 424–437. DOI: 10.2489/jswc.65.6.424

Cho J, Bosch D, Vellidis G, Lowrance R, Strickland T. 2013. Multi-site evaluation of hydrology component of SWAT in the coastal plain of southwest Georgia. *Hydrological Processes* **27**: 1691–1700.

Daggupati P, Douglas-Mankin K, Sheshukov A, Barnes P, Devlin D. 2011. Field-level targeting using SWAT: mapping output from HRUs to fields and assessing limitations of GIS input data. *Transactions of the ASABE* **54**: 501–514.

Daggupati P, Pai N, Ale S, Douglas-Mankin KR, Zeckoski R, Jeong J, Parajuli PB, Saraswat D, Youssef MA. 2015. Recommended calibration and validation strategy for hydrologic and water quality models. *Transactions of the ASABE*, (in press).

Douglas-Mankin K, Daggupati P, Sheshukov A, Barnes P. 2013. Paying for sediment: field-scale conservation practice targeting, funding, and assessment using the Soil and Water Assessment Tool. *Journal of Soil and Water Conservation* **68**: 41–51.

Douglas-Mankin K, Srinivasan R, Arnold JG. 2010. Soil and Water Assessment Tool (SWAT) model: current developments and applications. *Transactions of the ASABE* **53**: 1423–1431.

Eckhardt K, Fohrer N, Frede H. 2005. Automatic model calibration. *Hydrological Processes* **19**: 651–658.

Foster GR. 2005. *Revised Universal Soil Loss Equation, Version 2 (RUSLE2)* Science Documentation (In Draft). USDA-ARS: Washington, DC.

Gassman PW, Reyes MR, Green CH, Arnold JG. 2007. The soil and water assessment tool: historical development, applications, and future research directions.

He Y, Bárdossy A, Zehe E. 2011. A review of regionalisation for continuous streamflow simulation. *Hydrology and Earth System Sciences* **15**: 3539–3553.

Jha MK, Gassman PW. 2014. Changes in hydrology and streamflow as predicted by a modelling experiment forced with climate models. *Hydrological Processes* **28**: 2772–2781.

Jin X, Xu C, Zhang Q, Chen YD. 2009. Regionalization study of a conceptual hydrological model in Dongjiang basin, south China. *Quaternary International* **208**: 129–137. DOI: 10.1016/j.quaint.2008.08.006

Kaatz MR. 1955. The Black Swamp, a study in historical geography. *Annals of the Association of American Geographers* **45**: 1–35.

Kane DK, Conroy JD, Richards RP, Baker DB, Culver DB. 2014. Re-eutrophication of Lake Erie, correlations between tributary nutrient loads and phytoplankton biomass. *Journal of Great Lakes Research* **40**: 496–501.

Kannan N, Santhi C, Di Luzio M, Potter S, Arnold JG. 2005. Measuring environmental benefits of conservation practices: the conservation effects assessment project (CEAP)-A model calibration approach at the national level. 2005 ASAE Annual International Meeting.

Karr JR, Toth LA, Dudley DR. 1985. Fish communities of midwestern rivers, a history of degradation. *BioScience* **35**: 90–95.

Klemes V. 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal* **31**: 13–24.

Knisel WG, Douglas-Mankin KR. 2012. Creams/gleams: model use, calibration, and validation. *Transactions of the ASABE* **55**: 1291–1302.

Kumar R, Samaniego L, Attinger S. 2013a. Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research* **49**. DOI: 10.1029/2012WR012195

Kumar R, Livneh B, Samaniego L. 2013b. Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme. *Water Resources Research* **49**: 5700–5714. DOI: 10.1002/wrcr.20431

Luo Y, Su B, Yuan J, Li H, Zhang Q. 2011. GIS techniques for watershed delineation of SWAT model in plain polders. *Procedia Environmental Sciences* **10**: 2050–2057. DOI: 10.1016/j.proenv.2011.09.321

Merz R, Blöschl G. 2004. Regionalisation of catchment model parameters. *Journal of Hydrology* **287**: 95–123.

Moriasi D, Arnold J, Van Liew M, Bingner R, Harmel R, Veith T. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE* **50**: 885–900.

Neitsch SL, Arnold JG, Kiniry JR, Williams JR. 2011. Soil and water assessment tool theoretical documentation: version 2009. USDA–ARS, Grassland, Soil and Water Research Laboratory, Temple, TX; and Blackland Research and Extension Center, Texas AgriLife Research, Temple, TX. Texas Water Resources Institute Technical Rep. 406, Texas A&M University System, College Station, TX. http://swatmodel.tamu.edu/documentation/ (Accessed on may 2014).

Parajuli PB, Nelson NO, Frees LD, Mankin KR. 2009. Comparison of AnnAGNPS and SWAT model simulation results in USDA-CEAP agricultural watersheds in south-central Kansas. *Hydrological Processes* **23**: 748–763.

Piniewski M, Okruszko T. 2011. Multi-site calibration and validation of the hydrological component of SWAT in a large lowland catchment. In *Modelling of Hydrological Processes in the Narew Catchment*, Świątek D, Okruszko T (eds). Springer-Verlag: Berlin Heidelberg; 15–41. DOI: 10.1007/978-3-642-19059-9_2

Qi C, Grunwald S. 2005. GIS-based hydrologic modeling in the Sandusky watershed using SWAT. *Transactions of the American Society of Agricultural Engineers* **48**: 169–180.

Santhi C, Arnold JG, Williams JR, Hauck LM, Dugas WA. 2001. Application of a watershed model to evaluate management effects on point and nonpoint source pollution. *Transactions of the American Society of Agricultural Engineers* **44**: 1559–1570.

Santhi C, Kannan N, Arnold JG, Di Luzio M. 2008. Spatial calibration and temporal validation of flow for regional scale hydrologic modeling. *Journal of the American Water Resources Association* **44**: 829–846.

Saraswat D, Frankenberger JR, Pai N, Ale S, Daggupati P, Douglas-Mankin KR, Youssef MA. 2015. Hydrologic and water quality models, documentation and reporting procedures for calibration, validation, and use. *Transactions of the ASABE*, (in Press).

Sears PB. 1941. Postglacial vegetation in the Erie–Ohio Area. *The Ohio Journal of Science* **41**: 225–234.

Seibert J, McDonnell J. 2002. On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration. *Water Resources Research* **38**: 23-1–23-14. DOI: 10.1029/2001WR000978

Seo M, Yen H, Kim MK, Jeong J. 2014. Transferability of SWAT Models between SWAT2009 and SWAT2012. *Journal of Environmental Quality* **43**(3): 869–880.

Sheshukov AY, Daggupati P, Douglas-Mankin K, Lee MM. 2011. High spatial resolution soil data for watershed modeling: 1. Development of a SSURGO-ArcSWAT utility. *Journal of Natural and Environmental Sciences* **2**: 15–24.

Srinivasan R, Zhang X, Arnold J. 2010. SWAT ungauged, hydrological budget and crop yield predictions in the upper Mississippi River basin. *Transactions of the ASABE* **53**: 1533–1546.

Stone MC, Hotchkiss RH, Hubbard CM, Fontaine TA, Mearns LO, Arnold JG. 2001. Impacts of climate change on Missouri River Basin water yield. *Journal of the American Water Resources Association* **37**: 1119–1129.

Trautman MB. 1939. The effects of man-made modifications on the fish fauna in Lost and Gordon Creeks, Ohio, between 1887–1938. *Ohio Journal of Science* **39**: 275–288.

Trautman MB, Gartman DK. 1974. Re-evaluation of the effects of man-made modifications on Gordon Creek between 1887 and 1973 and especially as regards its fish fauna. *The Ohio Journal of Science* **74**: 162–173.

Troy TJ, Wood EF, Sheffield J. 2008. An efficient calibration method for continental-scale land surface modeling. *Water Resources Research* **44**. DOI: 10.1029/2007WR006513

Tuppad P, Santhi C, Srinivasan R. 2010. Assessing BMP effectiveness: multiprocedure analysis of observed water quality data. *Environmental Monitoring and Assessment* **170**: 315–329. DOI: 10.1007/s10661-009-1235-8

Van Liew MW, Garbrecht J. 2003. Hydrologic simulation of the Little Washita River experimental watershed using SWAT. *Journal of the American Water Resources Association* **39**: 413–426.

Vandewiele GL, Elias A. 1995. Monthly water balance of ungauged catchments obtained by geographical regionalization. *Journal of Hydrology* **170**: 277–291.

Veith T, Wolfe M, Heatwole C. 2003. Optimization procedure for cost effective bmp placement at a watershed scale. *Journal of the American Water Resources Association* **39**: 1331–1343. DOI: 10.1111/j.1752-1688.2003.tb04421.x

Wang S, Zhang Z, Sun G, Strauss P, Guo J, Tang Y, Yao A. 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**: 4621–4632. DOI: 10.5194/hess-16-4621-2012

White KL, Chaubey I. 2005. Sensitivity analysis, calibration, and validations for a multisite and multivariable SWAT model. *Journal of the American Water Resources Association* **41**: 1077–1089. DOI: 10.1111/j.1752-1688.2005.tb03786.x

Xu C. 1999. Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resources Management* **13**: 353–368.

Yen H, Bailey RT, Arabi M, Ahmadi M, White MJ, Arnold JG. 2014a. The role of interior watershed processes in improving parameter estimation and performance of watershed models. *Journal of Environmental Quality* **43**: 1601. DOI: 10.2134/jeq2013.03.0110

Yen H, Wang X, Fontane DG, Harmel RD, Arabi M. 2014b. A framework for propagation of uncertainty contributed by parameterization, input data, model structure, and calibration/validation data in watershed modeling. *Environmental Modelling and Software* **54**: 211–221. DOI: 10.1016/j.envsoft.2014.01.004

Yen H, White MJ, Arabi M, Arnold JG. 2014c. Evaluation of alternative surface runoff accounting procedures using the SWAT model. SWAT Special Issue. *International Journal of Agricultural and Biological Engineering* **8**.

Zhang X, Srinivasan R, Bosch D. 2009. Calibration and uncertainty analysis of the SWAT model using Genetic Algorithms and Bayesian Model Averaging. *Journal of Hydrology* **374**: 307–317. DOI: 10.1016/j.jhydrol.2009.06.023

