



## Development of a distributed hydrological model to facilitate watershed management

Sisi Li , Margaret Gitau, Bernard A. Engel, Liang Zhang, Yun Du, Carlington Wallace & Dennis C. Flanagan

To cite this article: Sisi Li , Margaret Gitau, Bernard A. Engel, Liang Zhang, Yun Du, Carlington Wallace & Dennis C. Flanagan (2017): Development of a distributed hydrological model to facilitate watershed management, Hydrological Sciences Journal, DOI: [10.1080/02626667.2017.1351029](https://doi.org/10.1080/02626667.2017.1351029)

To link to this article: <http://dx.doi.org/10.1080/02626667.2017.1351029>



Accepted author version posted online: 04 Jul 2017.  
Published online: 19 Jul 2017.



Submit your article to this journal [↗](#)



Article views: 38



View related articles [↗](#)




View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

## Development of a distributed hydrological model to facilitate watershed management

Sisi Li <sup>a,b,c</sup>, Margaret Gitau<sup>b</sup>, Bernard A. Engel<sup>b</sup>, Liang Zhang<sup>a</sup>, Yun Du<sup>a</sup>, Carlington Wallace<sup>b</sup> and Dennis C. Flanagan<sup>b,d</sup>

<sup>a</sup>Key Laboratory of Environment and Disaster Monitoring and Evaluation of Hubei, Institute of Geodesy and Geophysics, Chinese Academy of Sciences, Wuhan, China; <sup>b</sup>Department of Agricultural & Biological Engineering, Purdue University, West Lafayette, Indiana, USA; <sup>c</sup>University of Chinese Academy of Sciences, Beijing, China; <sup>d</sup>National Soil Erosion Research Laboratory, USDA-Agricultural Research Service, West Lafayette, Indiana, USA

### ABSTRACT

To facilitate precise and cost-effective watershed management, a simple yet spatially and temporally distributed hydrological model (DHM-WM) was developed. The DHM-WM is based on the Mishra-Singh version of the curve number method, with several modifications: The spatial distribution of soil moisture was considered in moisture updating; the travel time of surface runoff was calculated on a grid cell basis for routing; a simple tile flow module was included as an option. The DHM-WM was tested on a tile-drained agricultural watershed in Indiana, USA. The model with the tile flow module performed well in the study area, providing a balanced water budget and reasonable flow partitioning. The daily coefficient of determination and Nash-Sutcliffe coefficient were 0.58 and 0.56, for the calibration period, and 0.63 and 0.62 for the validation period. The DHM-WM also provides detailed information about the source areas of flow components, the travel time and pathways of surface runoff.

### ARTICLE HISTORY

Received 29 May 2016  
Accepted 14 March 2017

### EDITOR

A. Castellarin

### ASSOCIATE EDITOR

F.-J. Chang

### KEYWORDS

distributed hydrological model; watershed management; variable source area; travel time; tile flow; non-point source pollution

## 1 Introduction

While many streams, lakes, and estuaries around the world are at risk of eutrophication due to excess input of nutrients (Smith 2003, Davis and Koop 2006, Kling *et al.* 2014, Scavia *et al.* 2014), reducing non-point source pollutant export from landscapes to water bodies with cost-effective practices remains challenging for water resources and environmental managers. In order to place the most effective management practices in the most critical areas, detailed information on critical source areas (CSAs) and pollutant pathways is needed. Nutrient mobility greatly varies with flow pathways. Nitrogen can percolate to the groundwater aquifer in the form of nitrate, but phosphorus is mainly exported with surface runoff in particulate form. For surface runoff specifically, the travel routes make a big difference. As research has shown, forests, grasslands and wetlands can reduce the loads of sediments and nutrients by sedimentation, adsorption, plant uptake and biological degradation (Chescheir *et al.* 1991, Uusi-Kamppa *et al.* 2000). The same amount of pollutants mobilized from the source area will result in different loads to the target stream reach if the runoff flows over different land covers with different travel

times. The effectiveness of best management practices (BMPs) is also influenced by the flow pathways. BMPs, such as vegetated filter strips or artificial wetlands, are ineffective if surface runoff does not flow across them or less effective if the travel time across them is not long enough. Hence, to place BMPs in the ideal areas to maximize their effectiveness, hydrological models that can track surface runoff flow paths across the landscape are needed.

Since the hydrological cycle is the driving force of sediment and pollutant transport and delivery, hydrological models are the basis of water quality models and are commonly used to assist with watershed management. For non-point source (NPS) pollution modelling research, the Soil and Water Assessment Tool or SWAT (Arnold *et al.* 1998) and the Annualized Agricultural Non-Point Source Pollution Model or AnnAGNPS (Bingner and Theurer 2001) are current widely used models (Li *et al.* 2014). Both SWAT and AnnAGNPS use the Natural Resources Conservation Service (NRCS), or previously named Soil Conservation Service (SCS), curve number method (USDA-SCS 1972) as the basis for hydrological modelling because of its simplicity and effectiveness in many watersheds. The surface runoff

simulated by the NRCS CN method accounts for some spatial heterogeneity, such as land covers, soils, management practices and antecedent moisture conditions. Where the NRCS CN method has been incorporated into long-term continuous hydrological models, the variation of soil moisture has usually been updated based on the water balance of the soil volume (Mishra and Singh 2004, Soulis and Dercas 2007, 2010). Still, topographic position or connectivity to water bodies, which significantly controls the spatial distribution of soil moisture and surface runoff in many regions, has been neglected (Dunne and Black 1970, Lyon *et al.* 2004). These spatial details should also be included to facilitate accurate and cost-effective watershed management, since CSAs for NPS pollution are regarded as the areas where high soil nutrient content co-occurs with the source areas of high flows (Gburek and Sharpley 1998).

In this study, a simple yet distributed hydrological model for watershed management that can be used to identify non-point source CSAs and track flow pathways was developed. This model calculates different flow components (surface runoff, groundwater flow and tile flow) and tracks the flow paths of surface runoff with detailed spatial representation of land cover, soil properties and topography using a simple model structure and a limited number of model parameters.

The development of the distributed hydrological model for watershed management (DHM-WM) is described first. The model is then applied in a small agricultural watershed to test its performance and applicability. Finally, modelling results are compared with observed data and evaluated based on existing

according to the topographic index. To facilitate tracking of flow paths, travel time of surface runoff was calculated on a grid cell basis for routing. Also, for watersheds where tile or subsurface drainage systems play a critical role, a simple tile flow module was included as an option. The model basis of the DHM-WM and these modifications are further described in this section.

## 2.1 Model basis: Mishra-Singh modified curve number method

The NRCS CN method is a simple empirical model that is widely used either individually or as a component of long-term models to simulate excess rainfall or direct runoff (Hawkins *et al.* 2009). In long-term hydrological models, other modules are included to simulate evapotranspiration, plant uptake, percolation, lateral flow, tile flow and baseflow, which requires numerous parameters and increases model complexity. To make the curve number (CN) method applicable for long-term simulation and also to maintain its simplicity, Mishra and Singh (2004) incorporated soil moisture content explicitly in the runoff calculation equation and separated infiltration into two parts: constant gravitational infiltration and dynamic infiltration. A brief introduction to the Mishra-Singh modified version of the CN method is described here and further details can be found in Mishra and Singh (2003, 2004).

The modified CN method calculates direct runoff for each day as:

$$Q_{(t,t+\Delta t)} = \frac{(P_{(t,t+\Delta t)} - I_{a(t)} - F_{c(t,t+\Delta t)})(P_{(t,t+\Delta t)} - I_{a(t)} - F_{c(t,t+\Delta t)} + M_{(t)})}{P_{(t,t+\Delta t)} - I_{a(t)} - F_{c(t,t+\Delta t)} + M_{(t)} + S_{(t)}} \quad (1)$$

model performance criteria. A discussion of model strengths and weaknesses is also presented.

## 2 Methods

The spatially and temporally distributed hydrological model for watershed management (DHM-WM) developed in this study is an integration and modification of current hydrological models. The basis of the DHM-WM is the Mishra-Singh modified version of the NRCS CN method (Mishra and Singh 2004). To improve the model's spatial representation of soil moisture distribution, the concept in TOPMODEL (Beven 1995) was included to distribute soil moisture

where  $Q_{(t,t+\Delta t)}$  is the direct runoff generated in time interval  $\Delta t$  in mm;  $P_{(t,t+\Delta t)}$  is the precipitation in mm;  $F_{c(t,t+\Delta t)}$  is the constant gravitational infiltration during the time interval  $\Delta t$  in mm;  $I_{a(t)}$  is the initial abstraction; and  $M_{(t)}$  is the antecedent moisture content of time step  $t$  in mm. Like the classical NRCS CN method,  $I_{a(t)} = \lambda \cdot S_{(t)}$ , where  $\lambda$  is the initial abstraction coefficient and  $S_{(t)}$  is the maximum retention parameter of time step  $t$  in mm. Equation (1) is valid when  $P_{(t,t+\Delta t)} \geq (I_{a(t)} + F_{c(t,t+\Delta t)})$ , and  $Q_{(t,t+\Delta t)}$  equals 0 otherwise. Variable  $F_{c(t,t+\Delta t)}$  equals 0 when  $P_{(t,t+\Delta t)} \leq I_{a(t)}$ , and  $F_{c(t,t+\Delta t)} = (P_{(t,t+\Delta t)} - I_{a(t)})$  when  $I_{a(t)} \leq P_{(t,t+\Delta t)} \leq (I_{a(t)} + F_{c(t,t+\Delta t)})$ , assuming precipitation will first satisfy initial abstraction, then satisfy gravitational infiltration before generating runoff.

Gravitational infiltration  $F_c$  is applied as a constant parameter that was calibrated in the study of Mishra and Singh (2004), but it can also be represented as  $F_c = f_c \cdot t_e$ , where  $f_c$  is the gravitational infiltration rate in  $\text{mm h}^{-1}$ , and  $t_e$  is the time duration of effective precipitation in hours (Mishra and Singh 2003). To account for spatial variation of gravitational infiltration, the DHM-WM uses the equation  $F_c = f_c \cdot t_e$ , and  $f_c$  is assumed to be linearly related to saturated hydraulic conductivity ( $K_S$ ) as:  $f_c = f_{c0} \cdot K_S$ , where  $f_{c0}$  is a constant parameter that requires calibration.

The dynamic infiltration component of infiltration ( $F_d$ ) during the  $\Delta t$  interval is computed according to the water balance as:

$$F_{d(t,t+\Delta t)} = P_{(t,t+\Delta t)} - I_{a(t)} - F_{c(t,t+\Delta t)} - Q_{(t,t+\Delta t)} \quad (2)$$

$$\begin{aligned} \text{ET}_{\text{act}(t,t+\Delta t)} &= \text{PET}_{(t,t+\Delta t)}, & \text{if } 0.5 \cdot I_{a(t)} > (\text{PET}_{(t,t+\Delta t)} - \text{ET}_{(t,t+\Delta t)}) \\ \text{ET}_{\text{act}(t,t+\Delta t)} &= \text{ET}_{(t,t+\Delta t)} + 0.5 \cdot I_{a(t)}, & \text{if } 0.5 \cdot I_{a(t)} \leq (\text{PET}_{(t,t+\Delta t)} - \text{ET}_{(t,t+\Delta t)}) \end{aligned} \quad (4)$$

which is valid when  $P_{(t,t+\Delta t)} \geq (I_{a(t)} + F_{c(t,t+\Delta t)})$ , otherwise  $F_{d(t,t+\Delta t)}$  equals 0. For soil moisture budgeting, the  $M_t$  value will be updated in each time step by adding  $F_{d(t,t+\Delta t)}$  and subtracting evapotranspiration ( $\text{ET}_{(t,t+\Delta t)}$ ) as:  $M_{(t+\Delta t)} = M_{(t)} + F_{d(t,t+\Delta t)} - \text{ET}_{(t,t+\Delta t)}$ . The Mishra-Singh modified version of the CN method assumes a simple two-layer soil storage profile and calculates  $\text{ET}_{(t,t+\Delta t)}$  as a function of potential evapotranspiration during the  $\Delta t$  period ( $\text{PET}_{(t,t+\Delta t)}$ ) and the soil moisture condition:

$$\text{ET}_{(t,t+\Delta t)} = \text{PET}_{(t,t+\Delta t)} [1 - (S_{(t)} / S_{\text{abs}})^2] \quad (3)$$

The initial abstraction coefficient,  $\lambda$ , is also equal to  $(S_{(t)} / S_{\text{abs}})$ , which varies from 0 to 1. However, many studies around the world (Bosznay 1989, Mishra and Singh 2003) have shown a lower  $\lambda$  value from 0 to 0.3, and recent research recommended a value of 0.05 (Jiang 2001, Hawkins *et al.* 2009). Thus, the DHM-WM limits the maximum value of  $\lambda$  by introducing a parameter  $\lambda_{\text{MAX}}$ , so that  $\lambda = \lambda_{\text{MAX}} \cdot (S_{(t)} / S_{\text{abs}})$ .

In the Mishra-Singh modified CN model, the initial abstraction was excluded from the water budget. By definition, the initial abstraction generally includes interception, surface depressional storage and the initial infiltration before runoff occurs. The interception is approximately 25% of the annual precipitation and varies with season (Chow 1964); the depressional storage on gentle or moderate slopes can be 6–12 mm (Horton 1935); the amount of initial infiltration is vague, since few

studies have measured it separately from the total infiltration. The water stored in the depressions either evaporates or infiltrates, while the intercepted water evaporates to the air eventually (Tian *et al.* 2012). From this information, the initial abstraction is part of the water budget and will either evaporate or infiltrate, but it is difficult to determine the portions. In the DHM-WM, it is assumed that half of the  $I_{a(t)}$  will first satisfy the  $\text{PET}_{(t,t+\Delta t)}$  deficit, then the rest of  $I_{a(t)}$  will infiltrate to add to soil moisture. This assumption may not represent reality in some cases, but its influence could be reduced with adjustment of the initial abstraction coefficient parameter. With this assumption, the actual evapotranspiration  $\text{ET}_{\text{act}(t,t+\Delta t)}$  is adjusted as:

and the soil moisture is updated daily as:

$$\begin{aligned} M_{(t+\Delta t)} &= M_{(t)} + F_{d(t,t+\Delta t)} + I_{a(t)} \\ &\quad - \text{ET}_{\text{act}(t,t+\Delta t)} \end{aligned} \quad (5)$$

Here, the sum of  $S$  and  $M$  represents the absolute potential maximum retention,  $S_{\text{abs}}$ , which corresponds to a completely dry soil. In the Mishra-Singh modified CN method,  $S_{\text{abs}}$  is a parameter that requires calibration. In the DHM-WM, to make the parameter distributed and representative of the spatial distribution of land covers and soils, the default values of  $S_{\text{abs}}$  for each land cover-soil combination are calculated as  $S_{\text{abs,default}} = 2.281 \cdot (25 \cdot 400 / \text{CN} - 254)$ , where CN is the curve number that can be determined from NRCS National Engineering Handbook 630 tables. The  $S_{\text{abs,default}}$  values are then multiplied by an adjustment factor to obtain the calibrated  $S_{\text{abs}}$  values. In this way, the calibration of  $S_{\text{abs}}$  is done by changing the adjustment factor, and the relative values of  $S_{\text{abs}}$  among different land cover-soil combinations are maintained during calibration.

The Mishra-Singh modified CN method calculates the baseflow at the outlet of a watershed with the assumption that  $F_c$  percolates down to supply the groundwater aquifer and the surface watershed boundary coincides with the aquifer boundary. The baseflow at the watershed outlet ( $O_b$ ) is calculated using a single linear reservoir routing:

$$O_{b(t+\Delta t)} = g_0 \cdot F_{c(t)} + g_1 \cdot F_{c(t)} + g_2 \cdot O_{b(t)} \quad (6)$$

where  $g_0$ ,  $g_1$ ,  $g_2$  are the routing coefficients, which can be calculated as:

$$g_0 = \frac{\Delta t/K_b}{2 + \Delta t/K_b}, g_1 = g_0, g_2 = \frac{2 - \Delta t/K_b}{2 + \Delta t/K_b} \quad (7)$$

where  $K_b$  is the storage coefficient of baseflow. The Mishra-Singh modified CN method uses the same routing method for surface flow. To track flow paths of surface runoff generated on each grid cell of the landscape, the DHM-WM calculates the travel time of surface runoff on a grid basis for routing, which will be described later.

## 2.2 Modification 1: improve spatial representation of soil moisture distribution using concepts in TOPMODEL

Antecedent moisture content in the CN method significantly impacts water allocation between direct surface runoff and infiltration, so better representation of soil moisture distribution is needed to improve the spatial details of hydrological models. Usually, soil moisture varies greatly across a watershed due to subsurface lateral water movement. Nachabe (2005) proposed an equivalence between the NRCS CN method and TOPMODEL in predicting variable runoff source areas, which makes it reasonable to bring the concepts in TOPMODEL into the CN method to distribute  $S$  or  $M$  parameters according to watershed soil and topographic characteristics.

TOPMODEL represents the depth to the subsurface water table as a function of the soil topographic index (STI) (Beven and Kirkby 1979, Beven 1995):

$$d_i = \bar{d} - \frac{1}{f} (STI_i - \overline{STI}) \quad (8)$$

where  $d_i$  is the local depth to water table in grid  $i$  and  $\bar{d}$  is the watershed average depth to the water table in mm;  $f$  is a coefficient in the exponential decline of conductivity with depth in  $\text{mm}^{-1}$ ;  $STI_i$  and  $\overline{STI}$  are local and watershed average soil topographic index, respectively, and calculated as:

$$STI_i = \ln \left( \frac{\alpha_i}{Z_i \cdot K_i \cdot \tan(\beta_i)} \right) \quad (9)$$

where  $\alpha_i$  is the area of the upslope watershed per unit contour length at grid  $i$  in m;  $\tan(\beta_i)$  is the local surface slope in  $\text{mm}^{-1}$ ;  $Z_i$  is the depth of the soil to an impervious layer in m; and  $K_i$  is the mean saturated hydraulic conductivity in  $\text{m d}^{-1}$ .

The DHM-WM converts the variance of the water table depth to the variance of the soil moisture deficit by multiplying the ratio of the absolute maximum retention ( $S_{\text{abs}}$ ) and the total soil depth ( $Z$ ). Also, a parameter ( $m$ )

representing subsurface water velocity is included to allow adjustment for different watershed characteristics. Therefore, in each time step, the soil moisture content ( $M$ ) is redistributed according to STI as:

$$M_{\text{dis},i(t+\Delta t)} = M_{i(t+\Delta t)} + \frac{m}{f_i} (STI_i - \overline{STI}) \cdot \frac{S_{\text{abs},i}}{Z_i} \quad (10)$$

where  $M_{\text{dis},i(t+\Delta t)}$  is the distributed soil moisture content in grid cell  $i$  that will be used as the antecedent moisture content input for the next time step.

## 2.3 Modification 2: calculate the travel time of surface runoff on a grid basis for routing

Flow paths and travel times of surface runoff have great impacts on the transport and delivery of sediment and nutrients. In the DHM-WM, a flow direction map is derived from a digital elevation model (DEM) and used to track runoff flow paths and calculate runoff travel time on a grid cell basis. The travel time of surface runoff is estimated with two parts: overland travel time and in-channel travel time (Buchanan *et al.* 2013). Before surface runoff flows into channels, the overland flow travel time over grid cell  $i$  is calculated by combining the steady-state kinematic wave approximation with Manning's equation (Chow *et al.* 1988) as follows:

$$To_{i(t,t+\Delta t)} = q_{i(t,t+\Delta t)}^{-0.4} \left( \frac{L^{0.6} n^{0.6}}{\tan(\beta_i)^{0.3}} \right) \quad (11)$$

where  $To_{i(t,t+\Delta t)}$  represents the overland flow travel time across grid cell  $i$  during the time interval  $\Delta t$  in s;  $L$  is the hillslope length on grid cell  $i$  in m;  $\tan(\beta_i)$  is the slope of grid cell  $i$  in  $\text{mm}^{-1}$ ;  $n$  is Manning's roughness coefficient in grid cell  $i$ ;  $q_{i(t,t+\Delta t)}$  denotes the surface flow over grid cell  $i$  during the time interval  $\Delta t$  in  $\text{ms}^{-1}$ , which is calculated as follows:

$$q_{i(t,t+\Delta t)} = \frac{Q_{i(t,t+\Delta t)}}{D_{i(t,t+\Delta t)} \cdot 3600 \cdot 1000} \quad (12)$$

where  $D_{i(t,t+\Delta t)}$  is the duration over which surface runoff is generated in h, which approximates the duration of effective precipitation,  $t_e$ ;  $Q_{i(t,t+\Delta t)}$  is the direct surface runoff calculated with the Mishra-Singh modified CN method in Equation (1).

When surface runoff flows into channels, in-channel travel time within grid cell  $i$  is estimated by combining Manning's equation and the continuity equation for a wide channel (Muzik 1996, Melesse and Graham 2004) as follows:

$$Tc_{i(t,t+\Delta t)} = \frac{L \cdot n^{0.6} \cdot W^{0.4}}{\tan(\beta_i)^{0.3} \cdot q_{\text{acc},i(t,t+\Delta t)}^{0.4}} \quad (13)$$



where  $T_{c_{i(t,t+\Delta t)}}$  represents the in-channel travel time in s;  $L$  is the channel length within grid cell  $i$  in m;  $\tan(\beta_i)$  and  $n$  are the same as in the overland flow, Equation (11);  $q_{acc\_i(t,t+\Delta t)}$  denotes the instantaneous concentrated storm discharge on day  $t$  through grid cell  $i$  in  $\text{m}^3 \text{s}^{-1}$ , calculated as the weighted flow accumulation of  $Q_{i(t,t+\Delta t)}$  divided by the duration of flow ( $D_{i(t,t+\Delta t)}$ ); and  $W$  is the width of channels, in m.

The cumulative travel time of surface runoff generated on grid cell  $i$  during the time interval  $\Delta t$  along its flow path ( $T_{i(t,t+\Delta t)}$ ) is obtained by summing up its downstream travel time. Theoretically, the surface runoff reaching the watershed outlet at time  $t$  ( $O_{s(t)}$ ) is calculated by adding surface runoff at time  $t - T_{i(t,t+\Delta t)}$  from the grid cells of land where the cumulative travel time is equal to  $T_{i(t,t+\Delta t)}$ :

$$O_{s(t)} = \sum Q_{i(t-T_{i(t,t+\Delta t)})} \quad (14)$$

The DHM-WM uses a daily simulation time step, which is common for operational hydrological models (Arnold *et al.* 1998, Mishra and Singh 2004, Soulis and Dercas 2007). The daily time step is chosen because the general goal of the DHM-WM is watershed management that needs information about critical areas that generate runoff and critical days that have high flows. It may be problematic if the objective is flood control or others that need hourly or finer scale temporal accuracy. To save computing time, the hydrograph is reported only at target points selected by users. The surface flow at each selected target point on day  $t$  is estimated by summing up surface runoff generated on grid cells in the following manner: (i) surface runoff generated on day  $t$  of all the grid cells with cumulative travel time less than 6 h, (ii) surface runoff on day  $t - 1$  of all the grid cells with cumulative travel time between 6 and 30 h, ..., (iii) ending with the sum of surface runoff generated on day  $t - 6$  of all the grid cells with cumulative travel time over 102 h. This summation scheme is an approximation of the actual hydrograph that allocates runoff with travel time in hours to the daily simulation time step. The 6-h travel time for surface runoff reaching the target points in the same day is based on an assumption that most areas start generating runoff at the 18th hour of the day, which is suitable for regions where precipitation usually occurs in the afternoon or evening and also allows enough time for initial abstraction. This value may vary with different climate regions as well as different storms. The maximum number of days that will be used to account for surface flow also varies with watershed size.

## 2.4 Modification 3: a simple tile flow module as an option

Watersheds with tile drainage systems create additional challenges in hydrological models since tiles change subsurface water movement and flow partitioning. To expand the applicability of the DHM-WM in tile-drained watersheds without adding a large quantity of parameters, a simple tile flow module was included as an option. Based on the tile drainage routine used in SWAT2000 and SWAT2005 (Arnold *et al.* 1999, Du *et al.* 2005), the DHM-WM developed a simple tile flow module, assuming that tile flow occurs only when the water table height exceeds the height of the tile drains, or the depth of the water table ( $D_{\text{wtr}}$ ) is smaller than the depth of the tile drains ( $D_{\text{tile}}$ ):

$$q_{\text{Tile}_{i(t,t+\Delta t)}} = (D_{\text{tile},i} - D_{\text{wtr},i}) \cdot \text{Pe}_i \cdot \left(1 - \exp\left(\frac{-24}{t_{\text{drain},i}}\right)\right) \quad (15)$$

where  $q_{\text{Tile}_{i(t,t+\Delta t)}}$  is the water amount flowing into the tiles at grid cell  $i$  during the time interval  $\Delta t$  in mm;  $\text{Pe}_i$  is the effective porosity at grid cell  $i$ ;  $t_{\text{drain},i}$  is the time required to drain the soil to field capacity in h;  $D_{\text{tile},i}$  is the depth of tile drains at grid cell  $i$ , an input parameter of the model in mm;  $D_{\text{wtr},i}$  is the depth of the water table at grid cell  $i$  in mm, which is calculated as:

$$D_{\text{wtr},i} = \frac{S_i}{S_{\text{abs},i}} \cdot Z_i \cdot 1000 \quad (16)$$

where  $S$ ,  $S_{\text{abs}}$  and  $Z$  are the same as previously defined; 1000 is used to convert m to mm. To avoid introducing more parameters, Equation (15) assumes that the ratio of the depth of water table to the depth of the impervious layer is equal to the ratio of the actual retention potential of the soil to the absolute maximum retention potential. The amount of water flowing into the tiles as calculated by this kind of routine may be larger than the maximum amount of water that the drain tiles can hold (Moriassi *et al.* 2007a), thus  $q_{\text{Tile}_{i(t,t+\Delta t)}}$  is adjusted to be the smaller value between the value calculated with Equation (14) and a pre-defined parameter, drainage coefficient (DC, in  $\text{mm d}^{-1}$ ). In the DHM-WM, the tile flow reaching the watershed outlet tile ( $O_{\text{tile}}$ ) is calculated using the same routing method in SWAT:

$$O_{\text{tile}(t+\Delta t)} = \left( \sum_i q_{\text{Tile}_{i(t,t+\Delta t)}} + R_{\text{tile}(t)} \right) \cdot \left(1 - \exp\left(\frac{-24}{g_{\text{drain}}}\right)\right) \quad (17)$$

where  $g_{\text{drain}}$  is the lag coefficient of tile flow in h;  $R_{\text{tile}(t)}$  is the tile flow that remained in the tile drainage systems from the previous time step  $t$  in mm.

Since tiles drain excess water before it can move laterally to downslope areas, soil moisture variation is mainly controlled by local water budgeting rather than topographic position. Hence, Modification 1 is bypassed and the soil moisture content used in the CN method is updated daily as:

$$M_{(t+\Delta t)} = M_{(t)} + F_{d(t,t+\Delta t)} + I_{a(t)} - q_{\text{Tile}(t,t+\Delta t)} - \text{ET}_{\text{act}(t,t+\Delta t)} \quad (18)$$

Equation (18) replaces Equation (5), and  $M_{(t+\Delta t)}$  is directly used as the antecedent moisture content for the next time step.

## 2.5 Workflow of DHM-WM

The DHM-WM was compiled in the Python programming language, which allowed use of the spatial analysis tools in ArcGIS. The workflow of DHM-WM is illustrated in Figure 1. The time step of DHM-WM is daily, and for each day of a modelling period, the soil moisture needs to be updated. Since soil moisture variation will be different in areas with or without tile drainage systems, users need to determine whether tile drainage is present and significant in their watershed. If there is no or limited tile drainage, soil moisture distribution is generally controlled by soil and topographic characteristics. Hence, Modification 1 is used to distribute soil moisture according to STI as described in Section 2.2. If tile drainage is determined to be significant, Modification 3 is used to calculate tile flow and soil moisture is updated only by water budgeting without spatial distribution. For the soil moisture update, Modification 1 and Modification 3 are either/or options in the current version of DHM-

WM, and the determination of which method should be used is generally a pre-defined issue based on field survey or consultation with local experts. For the case study below, both options are utilized for the initial application to show how they work. The updated soil moisture content is used as input to the Mishra-Singh modified CN method to calculate surface runoff depth for the next time step. The travel time of surface runoff is then calculated on a grid basis for routing. Meanwhile, baseflow is estimated using the Mishra-Singh modified CN method. Finally, the streamflow at the watershed outlet is calculated by summing direct flow, baseflow and tile flow (if the tile flow module is used). The model will record and output the time series of total stream flow and all the flow components on a daily scale.

The DHM-WM model uses a limited number of parameters to provide distributed results. All parameters that may be used in the DHM-WM are listed in Table 1. If the tile flow module is turned off, there are five parameters recommended for calibration: CN,  $\lambda_{\text{MAX}}$ ,  $f_{c0}$ ,  $K_b$  and  $m$ . If the tile flow module is included, five parameters (CN<sub>adj</sub>,  $t_{\text{drain}}$ ,  $g_{\text{drain}}$ ,  $D_{\text{tile}}$ , DC) are added and one parameter ( $m$ ) is omitted. Parameter CN<sub>adj</sub> adjusts the CN values for tile-drained fields. The CN values for these fields are calculated by multiplying the CN<sub>adj</sub> and the normal CN values of the same field without tiles. This adjustment allows larger maximum potential storage in the tile-drained soils because the tiles can drain water away. The  $t_{\text{drain}}$  is a distributed parameter influenced by soil properties. In the current version, DHM-WM users must set the  $t_{\text{drain}}$  value for NRCS hydrologic soil group (HSG) D soils, and the value for any HSG

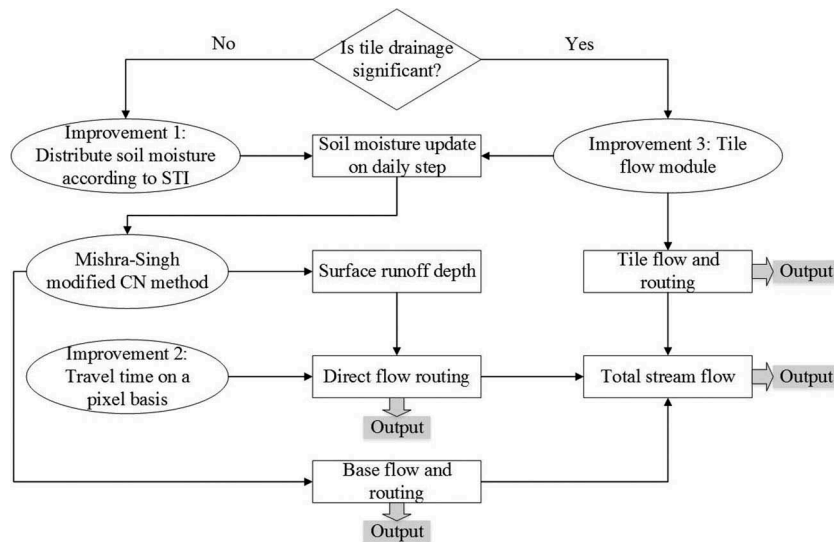


Figure 1. Work flow chart of the DHM-WM.

**Table 1.** Summary of DHM-WM parameters.

	Description	Distributed/lumped <sup>a</sup>	Range	Determination
$S_{abs}$	Absolute retention parameter in curve number method	Distributed	>0 mm	Default values from NRCS look-up table, better with calibration
$\lambda_{MAX}$	Maximum value of initial abstraction coefficient	Lumped	0–1.0 (0–0.3) <sup>b</sup>	Soft data or calibration
$f_{c,0}$	Constant gravitational infiltration rate coefficient	Lumped	0–1.0	Calibration
$K_b$	Storage coefficient of baseflow	Lumped	>0.5 d	Soft data or calibration
$m$	Watershed parameter that determines subsurface water velocity, only needed when tile flow module is turned off	Lumped	0–1.0	Calibration
$CN_{adj}$	Factor that adjusts CN values of tile drained fields, only needed when tile flow module is turned on	Lumped	0–1.0	Calibration
$t_{drain}$	Time required to drain soil to field capacity, only needed when tile flow module is turned on	Distributed	>24 h	Soft data or calibration
$g_{drain}$	Lag coefficient of tile flow, only needed when tile flow module is turned on	Lumped	>24 h	Soft data or calibration
$D_{tile}$	Depth to tile drains, only needed when tile flow module is turned on	Lumped	0–Z, mm	Field survey or soft data, calibration not recommended
DC	Drainage coefficient that limits maximum amount of water flow into tiles, only needed when tile flow module is turned on	Lumped	(0–20 mm d <sup>-1</sup> ) <sup>b</sup>	Field survey or soft data, calibration not recommended
$Z$ , $K_s$ , $f_r$	Soil properties: depth to impervious layer (Z), average saturated hydraulic conductivity ( $K_s$ ), exponential decline coefficient of soil conductivity ( $f_r$ ), effective porosity (Pe)	Distributed	-	Soil database, field survey or soft data
$n$	Manning's roughness coefficient	Distributed	0–1.0	Soft data

<sup>a</sup>Distributed parameters are set for each grid cell, variant for land covers or/and soils; lumped parameters are the same for the whole watershed.<sup>b</sup>The values in the parenthesis are recommended values by the DHM-WM.

C soils is calculated by assuming two times the water movement velocity of HSG D. Since soils with HSG A or HSG B are well drained, tile drainage is generally not installed in these soils. A total of seven parameters are recommended to be obtained through calibration, since  $D_{tile}$  and DC may be easier and better to determine with field surveys or soft data (literature values, knowledge of local experts, etc.). It should be noted that all parameter values obtained through calibration should be within expected bounds and carefully checked to ensure their representativeness. The DHM-WM also needs four soil-related parameters and one land cover-related parameter. These can be determined using available soil and land cover information and literature values, although they can also be adjusted manually in the model input files.

## 2.6 Evaluation of model performance

To test the performance of the DHM-WM, the time series of simulated streamflow was compared with observed data. Both graphical (time series, cumulative charts) and statistical (percent bias ( $P_{bias}$ ), coefficient of determination ( $R^2$ ), and Nash-Sutcliffe model efficiency coefficient ( $E_{NS}$ ) (Nash and Sutcliffe 1970)) measures were used to illuminate model performance in simulating streamflow. In addition, water budget and water partitioning among flow components were estimated to ensure different hydrological processes were being properly represented.

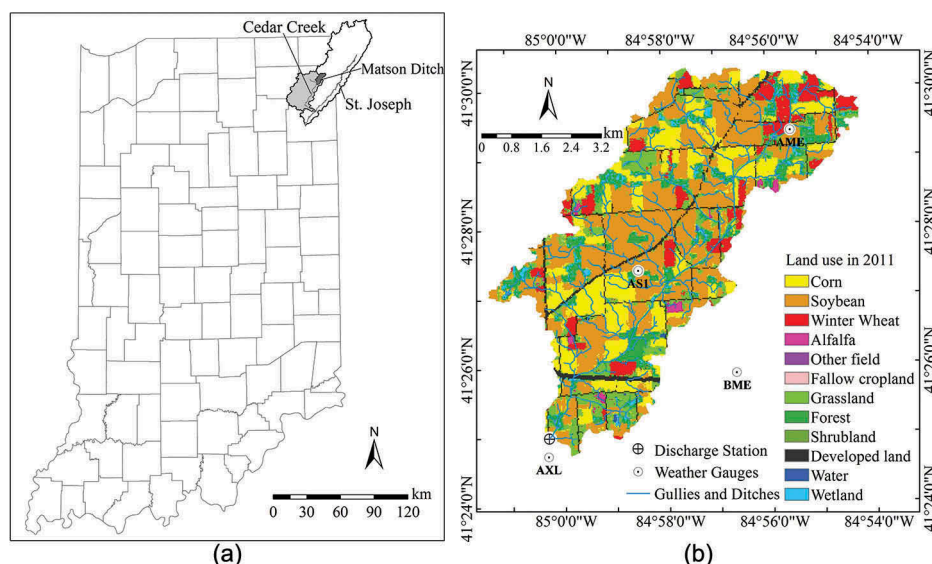
## 3 Case study

The DHM-WM was applied to the Matson Ditch watershed, a benchmark research watershed monitored by the Agricultural Research Service of USDA in north-eastern Indiana, USA. Initially, the DHM-WM was applied for a calibration period (2006–2009) with and without the tile flow module to test whether tile flow was significant and should be included. The modelling was then expanded to the validation period (2010–2012), and further analysis was conducted to illustrate model performance.

### 3.1 Site description

The Matson Ditch watershed (Fig. 2, HUC041000030603) covers an area of 43.13 km<sup>2</sup> and is a sub-watershed of the Cedar Creek watershed which drains into the St Joseph River. These areas have attracted the attention of researchers and managers since the agricultural fields in these areas are regarded as an important source of nutrient export to Lake Erie (Scavia *et al.* 2014). The Matson Ditch watershed is dominated by agricultural fields: 61% of the area is in a





**Figure 2.** (a) Location of the study watershed; and (b) the land cover, monitoring gauges and water courses of the Matson Ditch watershed.

corn–soybeans rotation, 8% is in winter wheat with a 3-year rotation as winter wheat–corn–soybeans, and 1% is in alfalfa hay production. The other land uses include 12% grasslands or pastures, 9% forests, 5% developed lands that are mainly roads, and 3% wetlands. The watershed is generally flat, with altitudes ranging from 231 m to 328 m above sea level, and the average slope is 1.0%. The two dominant soil types in the watershed are Blount (a silt loam, HSG D, accounting for 28% of the watershed) and Pewamo (a silt clay, HSG C, accounting for 20% of the watershed). Most of the soils have low infiltration rates, with 82% of the watershed soils classified as HSG C or D according to NRCS (NRCS 2009).

### 3.2 Data preparation

The data input for the DHM-WM includes daily precipitation, daily temperature, a digital elevation model (DEM) of the land surface, a digital map of the watercourses, a land cover map, and a soil map along with soil properties. The DEM data were at a resolution of 10 m from the USGS National Elevation Dataset (<http://ned.usgs.gov>). The same grid size was used for model simulation. The map of watercourses was generated from the flowline data of the USGS National Hydrography Dataset (NHD) (<http://nhd.usgs.gov>), which was “burned” into the DEM to force water flow into these watercourses. The sinks of the burned DEM were then filled and the flow direction was generated using the ArcGIS Hydrology tool. The flow direction data were the basis of calculating soil topographic index, tracking flow paths of surface runoff, and calculating travel time. The land cover map was the 2011 Cropland Data Layer (CDL) from the USDA

National Agricultural Statistics Service (<http://nassgeo.data.gmu.edu/CropScape>). The soil map was from the USDA SSURGO database (<http://www.nrcs.usda.gov>). The soil hydraulic conductivity was extracted from the soil input file of SWAT, and the effective porosity of soils was calculated with the SSURGO soil texture data (<http://web.ead.anl.gov/resrad/datacoll/porosity.htm>).

The weather data used in the DHM-WM were daily precipitation, and daily minimum and maximum temperatures. All of these data were from the USDA-ARS National Soil Erosion Research Laboratory in West Lafayette, Indiana. The precipitation gauges used included those located at sites AME, AS1, AXL and BME (Fig. 2). Thiessen-weighted precipitation of these gauges was used as model input to save calculation time. For temperature data, only gauge AS1 was used because it is located in the centre of the watershed and small differences were observed among the gauges. To make the moisture content ( $M$  and  $S$  variables) representative of the actual condition, a warm-up period covering several storms is needed. The period from December 2005 to March 2006 was used to warm up the model. The period from April 2006 to December 2009 was used for calibration, and the validation period was from January 2010 to December 2012.

### 3.3 Model calibration and evaluation

The DHM-WM was calibrated manually through a trial-and-error process by checking water balance and flow components as well as comparing the simulated time series of the streamflow at the watershed outlet with the observed streamflow monitored by ARS at the AXL gauge. Only the parameters in Table 2 were calibrated.

**Table 2.** DHM-WM calibrated parameters with and without tile flow module.

Parameters	Brief description	Value without tile flow module	Value with tile flow module
$S_{\text{abs}}$	Absolute retention parameter	$1.3 \cdot S_{\text{abs,default}}$	$1.3 \cdot S_{\text{abs,default}}$
$\lambda_{\text{MAX}}$	Maximum initial abstraction coefficient	0.13 for growing seasons 0.0055 for fallow seasons	0.12 for growing seasons 0.0050 for fallow seasons
$f_{c0}$	Coefficient of gravitational infiltration rate	0.035	0.016
$K_b$	Storage coefficient of baseflow	8 d	15 d
$m$	Watershed parameter for soil moisture distribution	0.3	—
$CN_{\text{adj}}$	CN adjustment factor for tile-drained areas	—	0.95
$t_{\text{drain}}$	Time to drain to field capacity	—	40 h for HSG D 10 h for HSG C
$g_{\text{drain}}$	Lag coefficient of tile flow	—	48 h

The distributed parameters  $S_{\text{abs}}$  and  $t_{\text{drain}}$  were calibrated as mentioned before. Other distributed parameters were not calibrated: those related to soil properties were extracted from the soil data, and Manning's roughness coefficient was determined from the literature. The observed streamflow covered the period from April 2006 to December 2012. Statistics ( $R^2$  and  $E_{\text{NS}}$ ) were calculated on both daily and monthly scales. The  $P_{\text{bias}}$  and the water budget were estimated on an annual basis.

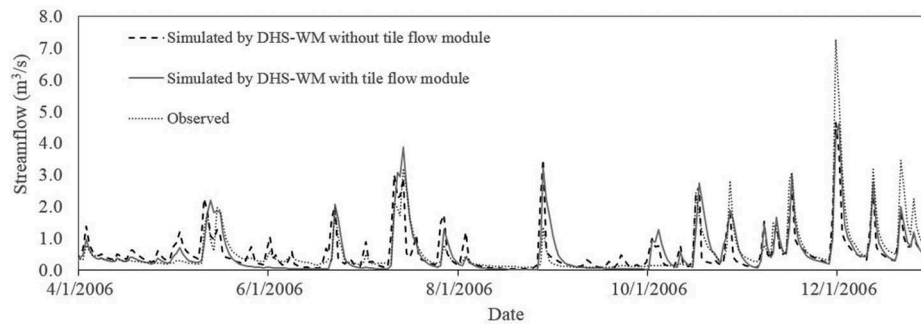
Model performance was evaluated by first checking whether the water budget was balanced on an annual basis and whether flow components were well represented. The patterns and closeness of match between simulated and observed data were then visually assessed with graphs. In addition, the statistics of the streamflow time series were analysed to substantiate model performance. If  $P_{\text{bias}}$  was less than  $\pm 15.0\%$ , and  $R^2$  and  $E_{\text{NS}}$  were greater than 0.70 and 0.65, respectively, on a monthly scale, the model performance was regarded as good; and  $R^2$  larger than 0.60 and  $E_{\text{NS}}$  larger than 0.50 were satisfactory (Moriassi *et al.* 2007b, 2015).

## 4 Results and discussion

### 4.1 Model calibration with and without tile flow module

For the calibration period, the streamflow simulated by the DHM-WM without using the tile flow module was composed of 49% baseflow and 51% surface flow. For the DHM-WM with the tile flow module, the simulated flow components were 29% baseflow, 45% tile flow and 26% surface flow. With or without the tile flow module, the general error and monthly variation of the streamflow simulated by the DHM-WM were both good after calibration. The  $R^2$  and  $E_{\text{NS}}$  of streamflow simulated without the tile flow module were 0.75 and 0.68, respectively, with a  $P_{\text{bias}}$  of 15.4% during the calibration period. When the tile flow module was included,  $R^2$  and  $E_{\text{NS}}$  of the calibration period were 0.73 and 0.70,

respectively, with a  $P_{\text{bias}}$  of 11.9%. However, the daily time series indicated some apparent differences. As shown in Figure 3, the streamflow simulated without the tile flow module tended to peak and drop more abruptly. This was as expected, since tile flow travels slower than surface flow and faster than baseflow. The tile flow simulated in the tile flow module became surface runoff and baseflow in the model without the tile flow module, which sharpened the peaking of streamflow and made the recession decline to a level value abruptly. The simulated results with the tile flow module better matched the peaking and recession trends of the observed data. The other differences shown in the time series were some small peaks simulated by the DHM-WM without the tile flow module that were not present in the observed streamflow. This was probably caused by the redistribution of subsurface moisture represented in the DHM-WM, which increased the moisture content of the lowland areas and made them more likely to generate runoff. However, the subsurface tiles actually drain the soil before water moves to the downslope areas. These two differences were also reflected by the statistics on a daily scale. The  $R^2$  and  $E_{\text{NS}}$  of the DHM-WM with the tile flow module were 0.58 and 0.56, respectively, better than the  $R^2$  of 0.53 and  $E_{\text{NS}}$  of 0.51 without the tile flow module. According to the Indiana Drainage Guide ID 160 ([http://irrigationtoolbox.com/ReferenceDocuments/BasicWaterManagement/f57\\_indiana\\_drainage\\_guide.pdf](http://irrigationtoolbox.com/ReferenceDocuments/BasicWaterManagement/f57_indiana_drainage_guide.pdf)), tile drainage systems are recommended for poorly and moderately drained fields. Previous research usually assumed tile drainage for agricultural fields with poorly drained or somewhat poorly drained soils in Indiana (Jiang *et al.* 2014, Boles *et al.* 2015). Based on this knowledge combined with the difference of simulated streamflow in this study, the tile flow was determined to be important in the Matson Ditch watershed. Thus, further statistical analysis in this study was based on the DHM-WM with the tile flow module. Prior to that, the calibrated parameters



**Figure 3.** Time series of observed streamflow and DHM-WM simulated streamflow with and without the tile flow module.

for the model with and without the tile flow module were compared (Table 2), to illustrate parameter meanings in the DHM-WM.

The absolute retention parameter ( $S_{abs}$ ) or its related CN values and the maximum initial abstraction coefficient  $\lambda_{MAX}$  directly determine the surface runoff. When the tile flow module was included, the CN values for the land areas that do not have tiles remained the same as the CN values without the tile flow module. For fields with tiles, CN values were reduced by multiplying the  $CN_{adj}$  of 0.95. The maximum initial abstraction coefficient  $\lambda_{MAX}$  was calibrated with respect to growing seasons (May–October) and fallow seasons (November–April), because much more water is intercepted by vegetation during the growing season. The initial infiltration before runoff occurred was also expected to be lower in the fallow season, since infiltration can be slowed or even prohibited by low temperature and frozen soils. Thus, the calibrated  $\lambda_{MAX}$  was much higher during the growing season. The model without the tile flow module had a slightly larger  $\lambda_{MAX}$ . This was reasonable since the average CN value for the model without tile flow was also larger. The positive relationship between CN values and  $\lambda_{MAX}$  here corresponded with the underlying logic of the NRCS CN method: the recommended CN values when  $\lambda$  equals 0.05 are lower than the CN values when  $\lambda$  is set to 0.2 (Hawkins *et al.* 2009).

The two parameters for baseflow,  $f_{c0}$  and  $K_b$ , were quite different for the models with and without the tile flow module. The  $f_{c0}$ , which defines how much water can infiltrate to the aquifer, was larger for the model without the tile flow module. In this case, part of the actual tile flow became baseflow. The  $K_b$ , which determines the lag time of baseflow, was shorter for the model without the tile flow module. Thus, the baseflow responded more quickly to the variation of precipitation.

The  $m$  value, which determines the subsurface moisture redistribution velocity, was set to be 0.3

in this study. That means it will take about 3.3 days (1 day divided by 0.3) to complete subsurface moisture redistribution. This parameter is not sensitive to the simulated streamflow at the watershed outlet, but significantly impacts the source areas of surface runoff. It should be noted that, without multiple gauges measuring streamflow and even soil moisture, the  $m$  parameter was not precisely calibrated in this case.

The depth to tile drains ( $D_{tile}$ ) was set to 1000 mm and the drainage coefficient (DC) was set to 10 mm  $d^{-1}$  according to Indiana Drainage Guide ID160 without calibration. For the other two tile flow parameters,  $t_{drain}$  was calibrated to be 40 h for HSG D and 10 h for HSG C, and  $g_{drain}$  was calibrated to be 48 h. These values were similar to the calibrated TDRAIN and GDRAIN parameters used in SWAT in a study in the same area (Boles *et al.* 2015).

## 4.2 Water budget and flow components

To evaluate the water balance simulated by the DHM-WM with tile flow, Table 3 lists precipitation, simulated flow output and evapotranspiration (ET) for each year. The watershed storage was calculated by subtracting flow output and ET from the precipitation. As can be seen, the watershed storage tended to be positive in wet years and negative for dry years, which was reasonable since extra water may be stored in the soils and the aquifers during wet years, and come out to satisfy ET or fill subsurface flow during dry years. For the period from 2006 to 2012, the watershed storage was  $-2.5$  mm, demonstrating simulated water balance.

Since an important goal of the DHM-WM is to represent the actual pathways of the flows, the simulated flow components are also listed in Table 3. In this watershed, tile flow simulated by DHM-WM was significant, contributing more than 40% of the

**Table 3.** Water budget and flow components simulated by the DHM-WM with tile flow module.

	Precipitation (mm)	Flow output (mm)	Water yield percentage (%)	Flow component (%)			ET (mm)	Watershed storage (mm)
				Baseflow	Tile flow	Surface flow		
Calibration period								
2006	987.4	479.5	48.6	28.0	48.3	23.8	490.6	17.3
2007	828.4	428.7	51.8	31.0	43.3	25.7	416.5	−16.8
2008	863.0	417.1	48.3	29.9	40.7	29.4	427.1	18.7
2009	1041.8	570.9	54.8	26.6	45.6	27.8	467.6	3.3
Validation period								
2010	800.5	340.2	42.5	29.2	44.1	26.7	490.2	−29.9
2011	1073.3	541.9	50.5	26.2	48.9	24.9	486.9	44.5
2012	604.5	218.4	36.1	43.2	37.9	18.9	428.9	−42.7

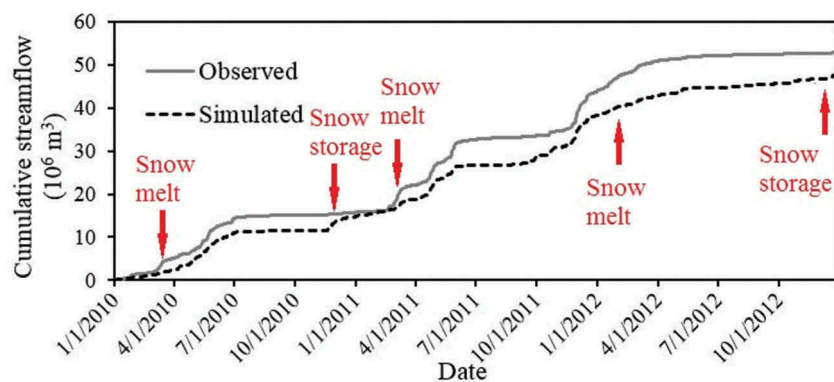
streamflow. The tile flow simulated in different years accounted for 13.7% to 25.0% of the annual precipitation, which was reasonable since the average percentage in studies with tile drains was reported to be 23.2% with a standard deviation of 13.9% (Boles *et al.* 2015). By contrast, the surface flow in the watershed was much smaller than the tile flow and varied greatly among years. The surface runoff component had a larger variation among years since it was directly impacted by the precipitation pattern, which was more variable than the subsurface water table that controlled tile flow. The baseflow simulated by DHM-WM was around 30% for most simulation years. When compared to literature values, the hydrograph also indicated the flow components were represented properly. The time to peak and recession were shortest for surface flow, medium for tile flow and longest for baseflow. The peaking and decreasing trends of simulated streamflow with the tile flow module in Figure 3 matched the observed trends in most cases, which indicated the flow components were reasonably simulated.

#### 4.3 Streamflow

Looking back to Figure 3, the peak of the simulated streamflow with the tile flow module matched observed

data in most storm events, but lagged by 1 day for several events. This is probably due to the summation scheme to approximate the surface flow hydrograph used in the model. The lagged events may have an early start time for precipitation and/or runoff generation, so that more runoff with travel time longer than 6 h reached the target outlet on the day of a precipitation event. The maximum number of days considered (7 d) in this study was long enough because most runoff reaches the outlet in 3 days. The travel time of runoff increases with watershed size, so that the maximum number of days considered needs to be larger for larger watersheds. Generally, 7 days is long enough for a watershed smaller than 200 km<sup>2</sup>.

The cumulative values of observed and simulated streamflow were compared in Figure 4 to evaluate the model capability of simulating the temporal variation of flow. Generally, the trend of the simulated line matched with the observed line. However, there were also some apparent discrepancies of the trends of the two lines. By checking the temperature and precipitation data, some of these discrepancies could be explained by snow storage and snowmelt that are not considered in the DHM-WM, indicated with red arrows in Figure 4. Other discrepancies, such as the overestimation in October 2011, may be caused by the model's simplification of other complex natural



**Figure 4.** Comparison of cumulative streamflow between observed values and simulated values by the DHM-WM with the tile flow module during the validation period.



processes. A general underestimation of streamflow for the winter periods was observed, causing the gap between the simulated and observed lines to become larger year after year. This might be attributable to processes that were significant only in winter and which are not accounted for in the model, or could also be due to some unrepresented change of parameters during winter. In addition, the uncertainty of the observed data might be another reason. During some of the winter months, one or two of the precipitation gauges were shut down. If there were no reliable gauges nearby to fill in the missing data, these gauges would have been omitted from the Thiessen weighting. The observed streamflow data also have a higher uncertainty during the winter months due to the cold temperature and freezing water which may damage the measurement instruments. In fact, when the winter periods (November–March) were excluded from statistics, the  $P_{\text{bias}}$  values for the calibration and validation periods were 0.2% and –2.0%, respectively.

Besides visual analysis of graphs, statistics were calculated for each year during both calibration and validation periods (Table 4). Generally, the statistics for both the calibration period and the validation period were good. The validation period had a lower  $P_{\text{bias}}$  and higher  $R^2$  and  $E_{\text{NS}}$ . However, looking into the statistics for each year, the model performance during the validation period showed more variation: the model predicted the flow variations well on both the daily scale and monthly scale in 2011; but in 2010 and 2012, the  $R^2$  and  $E_{\text{NS}}$  were just satisfactory on the monthly scale and not good on the daily scale. In this case, the good performance in the high flow year (2011) may outweigh the performance in the low flow years (2010 and 2012), making the overall performance during the validation period seem better. By contrast, the model performed more consistently during the calibration period. Still, model performance in 2006 and 2007 was better than in 2008 and 2009. The variation in

model performance among years was probably caused by the model's simplification of some natural processes, such as plant uptake of water, snow storage, snowmelt and soil freezing. For instance, the model performance in 2010 was mainly impacted by ignoring snow storage and snowmelt (Fig. 4). In addition, the uncertainty of the observed data discussed above may have contributed to the variations in performance. Considering the uncertainty, the model performed well generally, and the performance variations among years were considered acceptable.

#### 4.4 Spatial representation of the DHM-WM

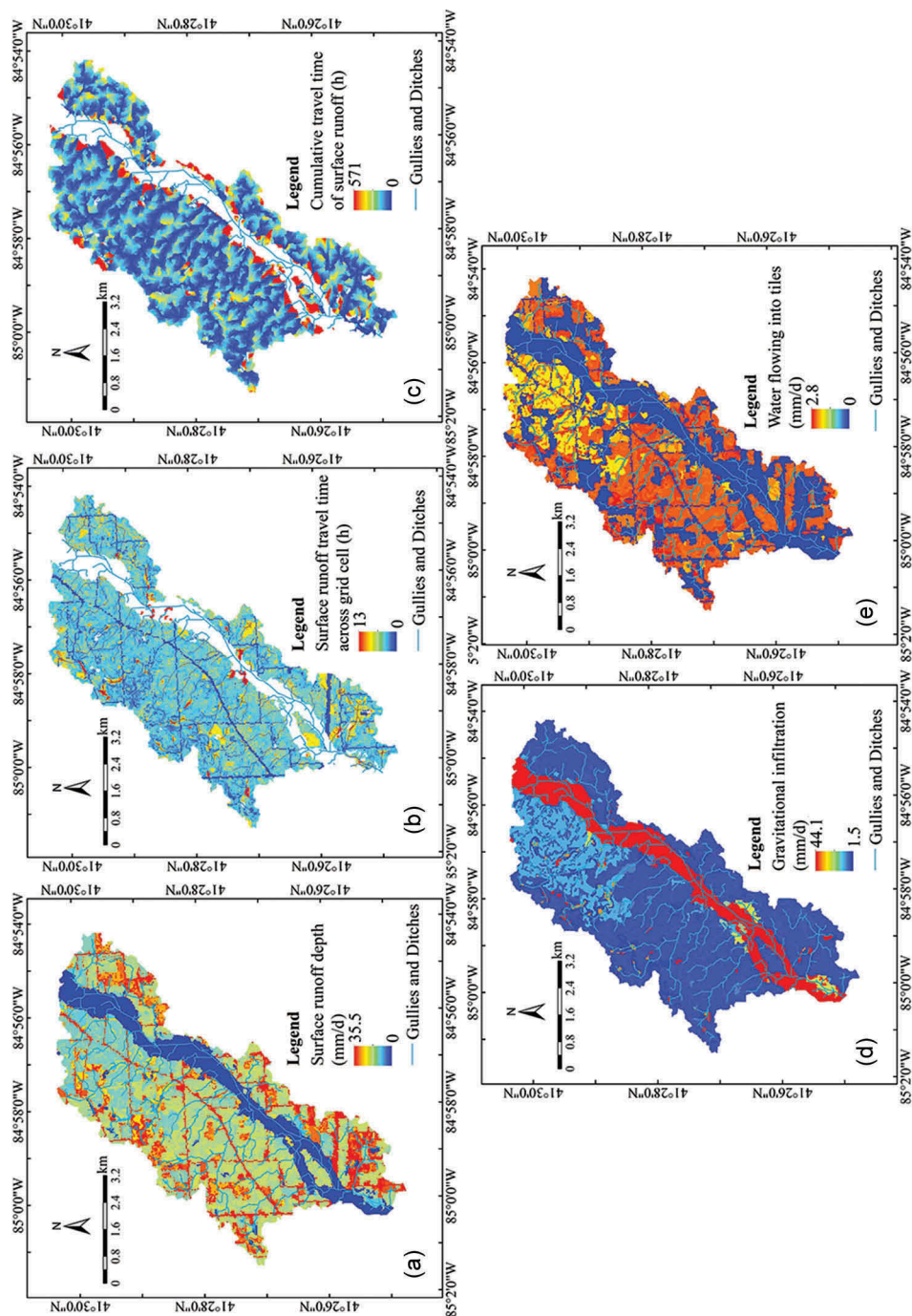
Two objectives of the DHM-WM are to identify the source areas of different flow components and to track the flow paths of surface runoff. These objectives were obtained by distributed representation of surface runoff and its travel time, infiltration to aquifer and tile flow, as illustrated in Figure 5. The surface runoff depth (Fig. 5(a)) was high from the roads, forests and grasslands with HSG C or HSG D, low from most agricultural fields, and nearly zero along Matson Ditch where soils were classified as HSG A or B. The lower surface runoff from most agricultural fields was contrary to normal conditions, because the tile drainage system underneath most of the fields in the watershed drained the soil and lowered the antecedent moisture condition, thus increasing initial abstraction and infiltration while decreasing surface runoff. The surface runoff travel time across a grid cell and the cumulative travel time to watershed outlet are shown in Figure 5(b) and (c), respectively. The travel time across a 10 m grid cell on the example day, 5 February 2008, had a mean value of 0.7 h and a standard deviation of 0.8 h (Fig. 5(b)). It varies with surface flow rate, slope and surface roughness, as indicated by Equation (11). The extremely long travel time occurred on lands that were very flat or generated very low surface runoff, and areas with high surface runoff depth generally had shorter travel time, such as the developed lands (roads). The cumulative travel time (Fig. 5(c)) was generally controlled by the topographic position, and was shorter along gullies and ditches and longer far away from these features. The extremely long travel time areas (in red) were flat lands with low surface runoff. Although these areas appeared to be close to gullies or ditches, it took a long time for the slow-moving surface runoff to flow into the watercourses. On the white areas, travel time was not applied since there was no surface runoff generated.

The gravitational infiltration (Fig. 5(d)), the source water for baseflow, was distributed mainly according to the soil hydraulic conductivity, and was extremely high

**Table 4.** DHM-WM performance in simulating streamflow with tile flow module.

	Daily		Monthly		$P_{\text{bias}}$ (%)
	$R^2$	$E_{\text{NS}}$	$R^2$	$E_{\text{NS}}$	
<i>Calibration period</i>					
2006–2009	0.58	0.56	0.73	0.70	11.9
2006	0.73	0.72	0.88	0.80	7.1
2007	0.62	0.61	0.69	0.65	15.0
2008	0.57	0.55	0.64	0.60	12.2
2009	0.54	0.48	0.77	0.75	11.8
<i>Validation period</i>					
2010–2012	0.63	0.62	0.77	0.74	9.6
2010	0.48	0.47	0.63	0.64	5.8
2011	0.71	0.68	0.81	0.74	16.8
2012	0.37	0.34	0.66	0.61	–6.6





**Figure 5.** Distributed maps of (a) surface runoff depth, (b) surface runoff travel time across a grid cell, (c) cumulative travel time of surface runoff, (d) gravitational infiltration and (e) water flowing into tiles in Matson Ditch watershed on 5 February 2008.

in the well-drained areas along Matson Ditch and low in the rest of the watershed. The northern part of the watershed tended to have higher gravitational infiltration than the southern part, because there is a large area of moderately well-drained soil, Rawson, in the north. In terms of the tile flow (Fig. 5(e)), this only occurred in the areas of agricultural fields with HSG C or D. The Rawson soil areas also had less water flowing into tiles. One reason was that more water infiltrated to the aquifer so that less was left on the surface and in the soil. The other reason may be that the larger effective porosity of Rawson soil allowed more water to be stored below the depth of tiles so that less water accumulated above the depth of tiles. It should be noted that the water flowing into the tiles on the example day, 5 February 2008, was generally low because the rainfall on that day (44.3 mm) needed some time to infiltrate and raise the soil water table. Actually, 10 mm of water flowed into the tiles the next day in almost all tile-drained areas in the watershed, reaching the maximum drainage coefficient set for the model.

#### 4.5 General discussion

The initial application of the DHM-WM in the Matson Ditch watershed showed satisfactory  $P_{bias}$ , monthly  $R^2$  and  $E_{NS}$  of streamflow simulation both when the tile flow module was included and when it was not. However, the simulated streamflow with the tile flow module better matched the observed data on a daily scale. These results indicated that it might not be sufficient to demonstrate the applicability of a model with only satisfactory monthly statistical indices of streamflow results, especially when water quality is the ultimate concern. By adjusting model parameters, streamflow may be well simulated even when water partitioning among flow components is being misrepresented. In fact, the calibrated parameters of this study showed that part of the actual tile flow was represented as baseflow when the tile flow module was not included. Although having limited impact on the monthly statistics of streamflow results, the misrepresented flow components could significantly affect simulation of agricultural chemical losses. Thus, when only monthly data are available for comparison, checking flow components becomes more important.

To test the DHM-WM performance in the study area, statistical indices of streamflow results, water budget, flow partitioning and parameters were all explored. The water budget was satisfactory and balanced for a 7-year period; water partitioning among flow components was representative; and the parameters were reasonable and within expected

ranges. Hence, the DHM-WM was demonstrated to perform well in the Matson Ditch watershed. Besides simulating reasonable water components at the watershed outlet, the DHM-WM can also provide spatial details of where surface flow, groundwater flow and tile flow are generated, and the travel time of surface runoff on a grid basis. This spatial information will be helpful for water resource management and nutrient control. The areas that contribute higher groundwater flow should be the priority areas if nitrate is the main concern, while areas with higher surface flow are potential critical sources of sediments and phosphorus. The tile flow source areas should be treated cautiously, since tile drains can be an important pathway not only for nitrate but also for phosphorus (Smith *et al.* 2015). The grid-cell-based travel routes and travel times of surface runoff recorded in the model are ready to be used for further estimation of sediment and chemical losses, allowing consideration of their impacts on the delivery processes. With this further analysis, this information can aid in the placement and design of field-scale management practices, such as vegetated filter strips, detention ponds, or grassed waterways, to reduce cost and improve effectiveness.

The distribution of soil moisture by incorporating the concepts of TOPMODEL was proposed for Modification 1 of the DHM-WM. This modification was, however, bypassed in the case study since tile drainage in the watershed artificially changed the moisture distribution. Hence, this modification was further developed and tested in a humid coastal watershed where soil moisture variation is largely controlled by natural subsurface water (Li *et al.* 2017). Since TOPMODEL is suitable for watersheds where subsurface flow is dominant to runoff generation, the applicability and necessity of the modification of using the TOPMODEL concept to distribute soil moisture may need further investigation in different watersheds. Furthermore, Modifications 1 and 3 are either/or options to deal with soil moisture update now. For future development, it would be better to make it possible to turn on or off the two modifications separately, so that the model would be suitable for mixed watersheds in which portions of them are tile drained and other parts are controlled by natural subsurface water movement.

The spatially distributed processes were only qualitatively analysed in this case study because there were not sufficient gauges of observed data to do a full quantitative justification. This is a common challenge for distributed modelling research (Beven 1999, Cao *et al.* 2006). The appropriate density of evaluation sites also depends on watershed characteristics. For a

watershed like this case study, where tile drainage prohibited natural subsurface water movement, the spatial distribution is mostly due to local characteristics; that is, the distribution of precipitation, land cover, soils and management practices. The ideal evaluation is to test the simulated results in each combination of the above characteristics. This situation is not easy, if possible, to reach. So, in practice, multi-site testing would be the option to improve the model calibration and validation as opposed to using only one site. Thus, it is recommended that the DHM-WM be calibrated and validated on multiple sites in watersheds where data are available.

For watersheds with limited data, another advantage of the DHM-WM is its simplicity and limited parameters that require calibration. If the tile flow module is not used, there are only five parameters to calibrate; and seven parameters if the tile flow module is included. All parameters have clear and well-understood meanings; therefore, given the simple model structure, users can expect these parameters to have impacts on specific flow components. This makes manual calibration easier and reduces the uncertainty from equifinality that is present with large parameter models, especially applied in areas where measured data of multiple hydrological processes are unavailable (Seibert and McDonnell 2003). Admittedly, the simplicity of the DHM-WM ignored some complex natural processes, such as soil freezing, snowmelt and plant uptake. This may impair model performance in areas or during time periods where these processes are critical. Though these processes probably caused some variation of the model performance in this case study, the overall model performance was good.

The simulation results of distributed hydrological models are considerably dependent on their spatial resolution (Soulis and Dercas 2010). The grid size of the DEM especially impacts the portrayal of the land surface and hydrological simulations; finer DEM resolutions generally lead to better representation of land surface and hydrological processes, but a grid size smaller than the DEM resolution may have negative effects (Zhang and Montgomery 1994). Hence, we recommended that the simulation grid size of the DHM-WM be based on the resolution of the DEM data. In addition, the grid size may have an impact on the calibrated values of the absolute retention parameters or the related curve numbers, since coarser grid sizes would lump several heterogeneous grid cells of a finer size into one homogeneous grid cell, resulting in composite CN values. Composite CNs lead to smaller runoff depths than distributed CNs because of the curvilinear relationship between CN and runoff depth

(Grove *et al.* 1998). Considering the reasonable range of CN values, there may also be a proper range for the grid cell sizes. This may need further investigation.

Another concern with grid-cell-based models is the computational efficiency. The computational time of the DHM-WM depends on the watershed size, the grid cell size, the number of target points selected to report hydrographs, and the computer properties. For a watershed of 43 km<sup>2</sup> with a grid cell resolution of 10 m and one target point to report in this case study, DHM-WM simulation runs take about 1.5 hours per simulation year on a laptop with an Intel® Core™ i5 2.30 GHz CPU and 8.00 GB RAM. The computational efficiency of the model is restricted by the use of ArcGIS functions in Python scripts and the sequential algorithm used in the programming, so it can be improved by modifying the algorithms.

## 5 Conclusion

A simple spatially and temporally distributed hydrological model on a grid scale (DHM-WM) was developed to provide detailed information about source areas of different flow components and flow pathways of surface runoff as well as simulating stream-flow time series. Two options of the DHM-WM, with or without a tile flow module, were tested in a small agriculture dominated watershed in northeastern Indiana. The results showed that both options of the DHM-WM performed well after calibration based on  $P_{bias}$  and monthly  $R^2$  and  $E_{NS}$ . Given the presence of significant amounts of tile drains, it was expected that the DHM-WM with the tile flow module would be more applicable in the watershed. This was seen in its performance based on the daily-scale hydrographs. Besides simulating reasonable flow components at the watershed outlet, the DHM-WM with the tile flow module can also provide spatial details of the sources of different flow components and the travel time of surface runoff on a grid basis, thus facilitating further water quality modelling and watershed management. As a simple distributed hydrological model with limited parameters needing calibration, the DHM-WM is a promising supplement to more complicated or semi-distributed models, providing more details to facilitate accurate and cost-effective watershed management.

## Disclosure statement

No potential conflict of interest was reported by the authors.



## Funding

This work was supported by the National Key Research and Development Program of China [2016YFD0800500]; National Natural Science Foundation of China [41471433]; the China Scholarship Council [201504910561].

## ORCID

Sisi Li  <http://orcid.org/0000-0002-9699-9028>

## References

- Arnold, J.G., *et al.*, 1998. Large area hydrologic modeling and assessment part I: model development. *JAWRA Journal of the American Water Resources Association*, 34 (1), 73–89. doi:10.1111/j.1752-1688.1998.tb05961.x
- Arnold, J.G., *et al.*, 1999. Validation of the subsurface tile flow component in the SWAT model. In: *ASAE Annual International Meeting*. St. Joseph, MI: American Society of Agricultural Engineers.
- Beven, K., 1999. How far can we go in distributed hydrological modelling? *Hydrology & Earth System Sciences*, 5 (1), 1–12. doi:10.5194/hess-5-1-2001
- Beven, K., 1995. TOPMODEL. In: V.P. Singh, ed. *Computer models of Watershed hydrology*. Colorado: Water Resources Publications, 627–668.
- Beven, K.J. and Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology /Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrological Sciences Bulletin*, 24 (1), 43–69. doi:10.1080/02626667909491834
- Bingner, R. and Theurer, F., 2001. *AnnAGNPS technical processes: documentation version 2*. Oxford, MS: Agricultural Research Service, US Department of Agriculture.
- Boles, C.M., Frankenberger, J.R., and Moriasi, D.N., 2015. Tile drainage simulation in SWAT2012: parameterization and evaluation in an Indiana Watershed. *Transactions of the ASABE*, 58 (5), 1201–1213.
- Bosznay, M., 1989. Generalization of SCS Curve Number method. *Journal of Irrigation and Drainage Engineering-Asce*, 115 (1), 139–144. doi:10.1061/(ASCE)0733-9437(1989)115:1(139)
- Buchanan, B.P., *et al.*, 2013. A phosphorus index that combines critical source areas and transport pathways using a travel time approach. *Journal of Hydrology*, 486, 123–135. doi:10.1016/j.jhydrol.2013.01.018
- Cao, W., *et al.*, 2006. Multi-variable and multi-site calibration and validation of SWAT in a large mountainous catchment with high spatial variability. *Hydrological Processes*, 20 (5), 1057–1073. doi:10.1002/hyp.5933
- Chescheir, G.M., *et al.*, 1991. Nutrient and sediment removal in forested wetlands receiving pumped agricultural drainage water. *Wetlands*, 11 (1), 87–103. doi:10.1007/BF03160842
- Chow, V.T., 1964. *Handbook of applied hydrology*. New York: McGraw Hill.
- Chow, V.T., Maidment, D.R., and Mays, L.W., 1988. *Applied hydrology*. New York: McGraw Hill.
- Davis, J.R. and Koop, K., 2006. Eutrophication in Australian rivers, reservoirs and estuaries—a southern hemisphere perspective on the science and its implications. *Hydrobiologia*, 559 (1), 23–76. doi:10.1007/s10750-005-4429-2
- Du, B., *et al.*, 2005. Development and application of SWAT to landscapes with tiles and potholes. *Transactions of the ASAE*, 48 (3), 1121–1133. doi:10.13031/2013.18522
- Dunne, T. and Black, R.D., 1970. Partial area contributions to storm runoff in a small New England watershed. *Water Resources Research*, 6 (5), 1296–1311. doi:10.1029/WR006i005p01296
- Gburek, W.J. and Sharpley, A.N., 1998. Hydrologic controls on phosphorus loss from upland agricultural watersheds. *Journal of Environmental Quality*, 27 (2), 267–277. doi:10.2134/jeq1998.00472425002700020005x
- Grove, M., Harbor, J., and Engel, B., 1998. Composite VS. Distributed curve numbers: effects on estimates of storm runoff depths. *Journal of the American Water Resources Association*, 34 (5), 1015–1023. doi:10.1111/jawr.1998.34.issue-5
- Hawkins, R.H., *et al.*, eds., 2009. *Curve number hydrology: state of the practice*. Reston, VA: American Society of Civil Engineers (ASCE).
- Horton, R.E., 1935. *Surface runoff phenomena, part I. Analysis of the hydrograph*. Horton hydrological laboratory publication 101. Ann Arbor, MI: Edward Bros.
- Jiang, R., 2001. *Investigation of Runoff Curve Number Initial Abstraction Ratio*. Thesis (M.S.). University of Arizona.
- Jiang, Y., *et al.*, 2014. Estimation of nonpoint source nitrate concentrations in Indiana rivers based on agricultural drainage in the watershed. *JAWRA Journal of the American Water Resources Association*, 50 (6), 1501–1514. doi:10.1111/jawr.12216
- Kling, C.L., *et al.*, 2014. LUMINATE: linking agricultural land use, local water quality and Gulf of Mexico hypoxia. *European Review of Agricultural Economics*, 41 (3), 431–459. doi:10.1093/erae/jbu009
- Li, S., *et al.*, 2014. Worldwide performance and trends in nonpoint source pollution modeling research from 1994 to 2013: a review based on bibliometrics. *Journal of Soil and Water Conservation*, 69 (4), 121A–126A. doi:10.2489/jswc.69.4.121A
- Li, S., *et al.*, 2017. Development of a soil moisture-based distributed hydrologic model for determining hydrologically-based critical source areas. *Hydrological Processes*. doi:10.1002/hyp.11276
- Lyon, S.W., *et al.*, 2004. Using a topographic index to distribute variable source area runoff predicted with the SCS curve-number equation. *Hydrological Processes*, 18 (15), 2757–2771. doi:10.1002/hyp.1494
- Melesse, A.M. and Graham, W.D., 2004. Storm runoff prediction based on a spatially distributed travel time method utilizing remote sensing and GIS. *Journal of the American Water Resources Association*, 40 (4), 863–879. doi:10.1111/jawr.2004.40.issue-4
- Mishra, S.K. and Singh, V.P., eds., 2003. *Soil Conservation Service Curve Number (SCS-CN) Methodology*. Netherlands: Springer, 84–146.
- Mishra, S.K. and Singh, V.P., 2004. Long-term hydrological simulation based on the soil conservation service curve number. *Hydrological Processes*, 18 (7), 1291–1313. doi:10.1002/(ISSN)1099-1085
- Moriasi, D.N., *et al.*, 2007b. Model evaluation guidelines for systematic quantification of accuracy in watershed

- simulations. *Transactions of the ASABE*, 50 (3), 885–900. doi:[10.13031/2013.23153](https://doi.org/10.13031/2013.23153)
- Moriasi, D.N., *et al.*, 2015. Hydrologic and water quality models: performance measures and evaluation criteria. *Transactions of the ASABE*, 58 (6), 1763–1785. doi:[10.13031/trans.58.10715](https://doi.org/10.13031/trans.58.10715)
- Moriasi, D.N., Arnold, J.G., and Green, C., 2007a. Incorporation of Hooghoudt and Kirkham tile drain equations into SWAT2005. In: *Proc. 4th Intl. SWAT Conf*, 139–147.
- Muzik, I., 1996. Flood modelling with GIS-derived distributed unit hydrographs. *Hydrological Processes*, 10 (10), 1401–1409. doi:[10.1002/\(ISSN\)1099-1085](https://doi.org/10.1002/(ISSN)1099-1085)
- Nachabe, M.H., 2005. Equivalence between TOPMODEL and the NRCS Curve Number method in predicting variable runoff source areas. *Journal of the American Water Resources Association (JAWRA)*, 42 (1), 225–235. doi:[10.1111/j.1752-1688.2006.tb03836.x](https://doi.org/10.1111/j.1752-1688.2006.tb03836.x)
- Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10 (3), 282–290. doi:[10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- NRCS, 2009. Hydrologic soil groups. In: *National engineering handbook: part 630—hydrology*. Washington, DC: USDA NRCS, 7.1–7.5.
- Scavia, D., *et al.*, 2014. Assessing and addressing the re-eutrophication of lake Erie: central basin hypoxia. *Journal of Great Lakes Research*, 40 (2), 226–246. doi:[10.1016/j.jglr.2014.02.004](https://doi.org/10.1016/j.jglr.2014.02.004)
- Seibert, J. and McDonnell, J.J., 2003. The quest for an improved dialog between modeler and experimentalist. In: Q. Duan, *et al.*, eds. *Calibration of watershed models*. Washington, DC: American Geophysical Union, 301–315.
- Smith, D.R., *et al.*, 2015. Surface runoff and tile drainage transport of phosphorus in the midwestern united states. *Journal of Environmental Quality*, 44 (2), 495–502. doi:[10.2134/jeq2014.04.0176](https://doi.org/10.2134/jeq2014.04.0176)
- Smith, V.H., 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research*, 10 (2), 126–139. doi:[10.1065/espr2002.12.142](https://doi.org/10.1065/espr2002.12.142)
- Soulis, K. and Dercas, N., 2007. Development of a GIS-based spatially distributed continuous hydrological model and its first application. *Water International*, 32 (1), 177–192. doi:[10.1080/02508060708691974](https://doi.org/10.1080/02508060708691974)
- Soulis, K. and Dercas, N., 2010. AgroHydroLogos: development and testing of a spatially distributed agro-hydrological model on the basis of ArcGIS. In: *International Congress on Environmental Modelling and Software, Modelling for Environment's Sake, Fifth Biennial Meeting (iEMSs)*, 5–8 July, Ottawa.
- Tian, F., Li, H., and Sivapalan, M., 2012. Model diagnostic analysis of seasonal switching of runoff generation mechanisms in the blue river basin, oklahoma. *Journal of Hydrology*, 418–419 (4), 136–149. doi:[10.1016/j.jhydrol.2010.03.011](https://doi.org/10.1016/j.jhydrol.2010.03.011)
- USDA-SCS, 1972. Hydrology. Section 4. In: *Soil Conservation Service National Engineering Handbook*. Washington, DC: U.S. Department of Agricultural-Soil Conservation Service.
- Uusi-Kamppa, J., *et al.*, 2000. Buffer zones and constructed wetlands as filters for agricultural phosphorus. *Journal of Environmental Quality*, 29 (1), 151–158. doi:[10.2134/jeq2000.00472425002900010019x](https://doi.org/10.2134/jeq2000.00472425002900010019x)
- Zhang, W. and Montgomery, D.R., 1994. Digital elevation model grid size, landscape representation, and hydrologic simulations. *Differentiation; Research in Biological Diversity*, 30 (4), 1019–1028.