

# A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling

Chaopeng Shen, Mantha S. Phanikumar<sup>\*</sup>

Department of Civil & Environmental Engineering, Michigan State University, East Lansing, MI 48824, USA

## ARTICLE INFO

### Article history:

Received 2 March 2010

Received in revised form 26 August 2010

Accepted 1 September 2010

Available online 15 September 2010

### Keywords:

Catchment hydrology

Distributed hydrological model

Multiscale models

Runoff generation

Process-based models

Mechanistic models

## ABSTRACT

Process-based watershed models are useful tools for understanding the impacts of natural and anthropogenic influences on water resources and for predicting water and solute fluxes exported from watersheds to receiving water bodies. The applicability of process-based hydrologic models has been previously limited to small catchments and short time frames. Computational demands, especially the solution to the three-dimensional subsurface flow domain, continue to pose significant constraints. This paper documents the mathematical development, numerical testing and the initial application of a new distributed hydrologic model PAWS (Process-based Adaptive Watershed Simulator). The model solves the governing equations for the major hydrologic processes efficiently so that large scale applications become relevant. PAWS evaluates the integrated hydrologic response of the surface–subsurface system using a novel non-iterative method that couples runoff and groundwater flow to vadose zone processes approximating the 3D Richards equation. The method is computationally efficient and produces physically consistent solutions. All flow components have been independently verified using analytical solutions and experimental data where applicable. The model is applied to a medium-sized watershed in Michigan (1169 km<sup>2</sup>) achieving high performance metrics in terms of streamflow prediction at two gages during the calibration and verification periods. PAWS uses public databases as input and possesses full capability to interact with GIS datasets. Future papers will describe applications to other watersheds and the development and application of fate and transport modules.

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

In recent years, the focus of hydrologic modeling is shifting toward using models to predict future scenarios including the impacts of climate change and human interventions on water resources. The effects of these forcings are modulated by complex and interrelated responses of hydrologic processes which are often characterized by feedbacks and thresholds. Process-based models, derived deductively from established physical laws [1], can capture the underlying dynamics of watersheds and may produce better predictions across a range of scales. Understanding water and solute fluxes is also important from a human health perspective. A variety of pollutants including chemical and biological agents pose threats to human and ecosystem health [2]. Pollutants can reach streams via point source discharge, non-point source (overland flow), or subsurface seepage. Process-based hydrologic models explicitly detail the various flow paths and thus provide the necessary information to predict contaminant loads at downstream receiving bodies including lakes [3] and oceans [4].

<sup>\*</sup> Corresponding author. Tel.: +1 517 432 0851; fax: +1 517 355 0250.

E-mail addresses: [shenchao@msu.edu](mailto:shenchao@msu.edu) (C. Shen), [phani@msu.edu](mailto:phani@msu.edu) (M.S. Phanikumar).

Previous hydrologic models, especially those intended for large scale simulations, tend to use conceptual representations of the groundwater compartment, ignoring the complexity of the groundwater and vadose zone flow problems in space, e.g. SWAT [5]. Some large-scale land surface models (LSMs), e.g. the VIC model [6] and Noah LSM [7], assume a leaky bottom for the land surface domain. These methods over-simplify the groundwater flow dynamics and lead to prediction errors when the water table is shallow or saturation excess is an important mechanism. On the continental scale, [8,9] added simple groundwater dynamics into LSMs and ran simulations for entire North America.

The subsurface flow is an integral component of the hydrologic cycle which should be studied using a holistic approach [10–12]. In shallow water table conditions, groundwater controls soil moisture and provides sources of water for ET. The soil moisture in turn exhibits heavy influences on surface energy fluxes and meteorological phenomena with feedback loops which may amplify the anomalies [13]. Coupled climate-hydrologic modeling indicated that shallow groundwater conditions in the humid regions and intermountain valleys in arid regions enhanced ET and precipitation [9,14]. Therefore, an accurate description of groundwater dynamics is crucial for understanding the nonlinear responses of

the hydrologic system and to describe impacts on regional climate.

There has been a growing interest in integrated surface–subsurface modeling. However, such research is still limited by computational constraints. The governing equation for three-dimensional subsurface flow is the Richards equation (RE) [15]. Rainfall, surface ponding and groundwater interact via the variably saturated soil zone (the vadose zone) where crucial processes such as infiltration, soil evaporation, root extraction and groundwater recharge/discharge take place. The RE automatically handles all relevant processes in a seamless fashion [16]. A number of watershed/regional-scale process-based models that solve the three-dimensional Richards equation have been recently developed to examine the interactions between surface and subsurface flow, e.g., InHM [17], Hydrogeosphere [18], CATHY [19], ParFlow [20,21], WASH123D [22], HYDRUS3D [23], MODHMS [24]. However, models that employ such a full 3D approach are faced with the problem of excessive computational demand, especially in large domains. Due to the strong non-linearity, the solution accuracy of the RE depends heavily on the spatial step size. In particular, the combination of heavy rain and a relatively dry soil surface can lead to the development of a steep wetting front along which soil moisture can change dramatically, making it necessary to use very fine grid cells. Downer and Ogden [25] studied the effect of vertical discretization on RE solution and concluded that to simulate infiltration accurately, the vertical cell size needs to be on the centimeter level near soil surfaces, but not throughout the soil column. The large matrix resulting from the 3-D discretization must be solved iteratively which is prohibitively expensive.

As a result, applications of the aforementioned process-based models were often restricted to small catchments (from plot scale to less than  $100 \text{ km}^2$ ) and short time frames. If the watershed size increased, the simulations took exceedingly long times. For example, the INHM model [17] or CAT3M [26] were only applied on a plot scale. The GSSHA model [27] was applied to small catchments ( $20 \text{ km}^2$  and  $3.64 \text{ km}^2$  in [25] and  $3 \text{ km}^2$  in [27]) over short time frames ( $<200$  days). The comparisons are also limited to stream-flow measurements. The applications of the MIKE-SHE model also ranged from a few  $\text{km}^2$  [28,29] to medium sized-watersheds of several hundred  $\text{km}^2$  [30]. Being a private-domain model, MIKE-SHE's source code is not accessible, which may have limited its use for research that links with other sciences. For medium/

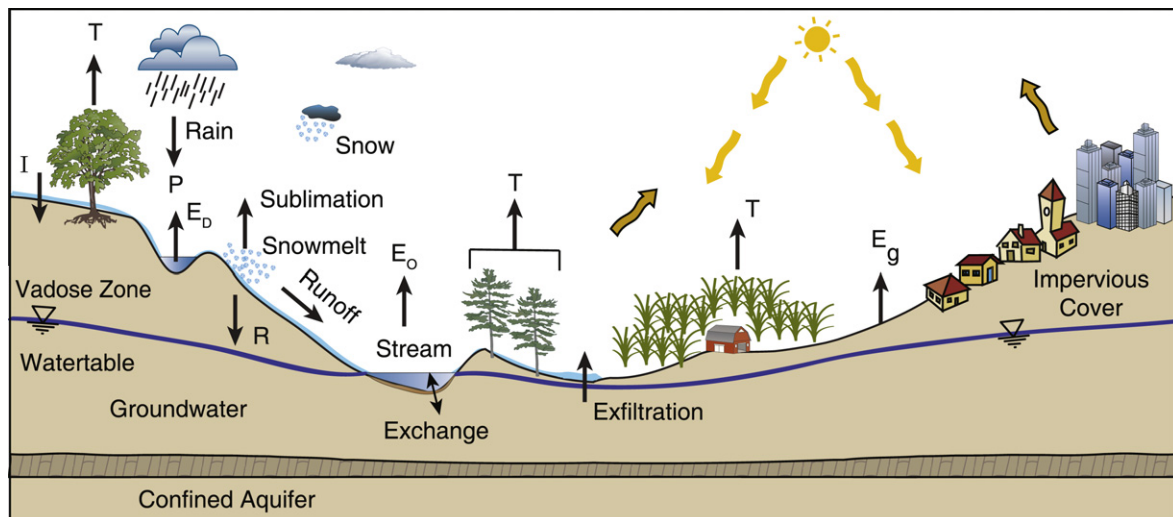
small-sized watersheds, models such as HydroGeoSphere [18] and CATHY [19] required many days to run when 3D RE is enabled. Some recent research makes use of parallel computing systems, e.g., ParFlow [20,21], WASH123D [22]. While such research offers useful insights, this means that the newly-available computing power will be consumed to solve the PDEs, making it difficult or impossible to carry out other equally important tasks such as model auto-calibration and uncertainty analysis. It is clear from the above review that widespread use of process-based models requires further advances in describing the subsurface physics in a computationally efficient manner. The aim of this paper is to describe a new process-based model for large watersheds that is based on a novel method for surface–subsurface coupling.

This paper documents the mathematical development, numerical testing and the initial application of a new process-based distributed hydrologic model, PAWS (Process-based Adaptive Watershed Simulator). The model solves governing equations for the major hydrologic processes efficiently so that large scale applications become relevant. PAWS is developed with the aims of long term simulations on medium ( $\sim 1000 \text{ km}^2$ ) to large ( $>5000 \text{ km}^2$ ) basins. With upscaling and parallelization, we expect to apply the model to larger (e.g., continental) scales in future. This paper serves as an introduction to the model and focuses mainly on the mathematical aspects of the model. Future papers will describe applications to other watersheds, data processing and the development and application of fate and transport modules.

## 2. Model development

### 2.1. General overview

PAWS solves physically-based conservation laws for major processes of the hydrologic cycle, which are depicted in Fig. 1 and summarized in Table 1. The eight compartments where most calculations take place are, respectively, surface ponding layer, canopy storage layer, impervious cover storage layer, overland flow layer, snowpack, soil moisture, groundwater aquifers and stream channels. The major state variables are summarized in Table 2. Clearly the vadose zone plays a central role in the model as it links surface water with groundwater and controls evapotranspiration. Having an efficient, flexible, and robust vadose zone module is thus critical to the success of the model.



**Fig. 1.** Definition sketch illustrating some of the key processes modeled in PAWS. T: transpiration, P: precipitation,  $E_o$ : evaporation from overland flow/stream,  $E_g$ : evaporation from bare soil;  $E_d$ : evaporation from depression storage; I: infiltration; R: Recharge.

**Table 1**  
Modeled processes.

Processes	Governing equations
Snowfall accumulation and melting	Mass and energy balance (UEB) [47]
Canopy interception	Bucket model, storage capacity related to Leaf Area Index (LAI)
Depression storage	Specific depth
Runoff	Manning's formula + Kinematic wave formulation + Coupled to Richards equation
Infiltration/exfiltration	Coupled to Richards equation. For heavy rain, Generalized Green and Ampt equation [45] is invoked
Overland flow	2D Diffusive wave equation
Overland/channel exchange	Weir formulation
Channel network	Dynamic wave or diffusive wave
Evapotranspiration	Penman Monteith + Root extraction
Soil water flow	Richards equation
Lateral groundwater flow	quasi-3D
Recharge/discharge (Vadose zone/Groundwater interaction)	Non-iterative coupling between Richards equation and groundwater flow equation
Stream/groundwater interaction	Conductance/leakance
Vegetation growth	Simplified growth cycle

**Table 2**  
Major state variables and their symbols.

State variable	Symbol	Unit
Pressure head	$h$	m
Overland flow depth <sup>a</sup>	$h$	m
Canopy storage	$CS$	m
Snow water equivalent,	$SWE$	m
Snow energy content	$U$	kJ/m <sup>2</sup>
Groundwater head	$H$	m
River cross sectional area	$A_c$	m <sup>2</sup>
Ponding storage on impervious cover	$h_i$	m

<sup>a</sup> There is no confusion between soil pressure head and overland flow depth since they are described in different sections, thus they share the same symbol  $h$ .

## 2.2. Spatial discretization and model structure

The spatial domain of the watershed is discretized into a structured grid. Rivers are modeled as separate objects in a network. Each river may be discretized with variable spatial step sizes. Exchange fluxes from land grid to river cell grid are computed on the river cells and transferred to intersecting land cells. This flexible strategy allows river and land domains to be discretized independently. Within each land cell, sub-grid heterogeneity is modeled using a method similar to the mosaic approach proposed in [31]. Depending on the grid size, a cell may possess a mixture of land use/land cover types. Landuse data is re-classified into model classes which are represented by several generic plant types (called the representative plant types or RPT). Currently, supported land use types and their RPTs include (RPTs are in the parentheses): Water (water), Urban (Urban), Evergreen Forest (Pine), Deciduous Forest (Oak), Bush and Shrub (Shrub), Herbaceous (Grass), Row Crops (Corn), Bare Soil (Bare Soil) and Forage Crops (Alfalfa). Five parameters completely characterize each RPT: crop coefficient for evapotranspiration ( $K_c$ ), canopy height ( $h_c$ ), rooting depth (RMX), Leaf Area Index (LAI) and growth periods ( $T_g$ ). For forest land uses,  $K_c$  is computed by applying the Penman–Monteith equation to standard tree types and normalizing the results by alfalfa ET. The values of these parameters for different RPTs are taken from literature values [5,32,33].

Water on the ground is separated into two sub-domains – the ponding water layer and the flow domain. The ponding layer will be subject to infiltration and evaporation whereas only water in the flow domain can move from one cell to another (including river cells). Runoff from ponding to flow domain can be calculated using several different methods described in Section 2.4.2. Depending on the climatic conditions, a layer of snow, quantified by the snow

water equivalent (SWE) and snow cover fraction, may also exist on the ground. After taking out the canopy interceptions, rainfall and snowmelt contribute to the ponding layer. One soil column is modeled for each cell. The soil column exchanges mass with the unconfined aquifer. Water can percolate further down to deeper aquifers and contribute to or gain from river water through river bed materials. Pumping activities can also directly extract water from the aquifer layers.

## 2.3. Hydrologic components

The four flow domains that the model considers are overland flow, channel flow, soil water flow and saturated groundwater flow. They are described below in detail. The other important component that will be discussed is evapotranspiration.

### 2.3.1. Surface ponding layer

The ponding layer on the ground surface serves as the entry point to look at these processes in detail. The mass balance equation for the ponding layer on the ground is written as:

$$\frac{\partial h_1}{\partial t} = P - \frac{dCS}{dt} + SNOM - E_g - I - F_g \quad (1)$$

where  $h_1$  is the water depth in the surface ponding layer (m) (also the first layer of the soil column),  $P$  is the rainfall rate (m d<sup>-1</sup>),  $CS$  is the canopy interception (m),  $SNOM$  is the rate of snowmelt (m d<sup>-1</sup>),  $E_g$  is the rate of evaporation on the ground (m d<sup>-1</sup>),  $I$  denotes infiltration (m d<sup>-1</sup>), and  $F_g$  is the runoff to the overland flow domain.

Runoff from ponding to flow domain is calculated together with infiltration inside the vadose zone model. The soil column module is responsible for computing infiltration from ground surface (the ponding domain), soil evaporation, root extraction and percolation into the unconfined aquifer. Water is assumed to move only vertically in the soil column but once in the aquifers it can move laterally. Eq. (1) for the surface ponding layer is solved simultaneously with the soil water fluxes in the vadose zone.

### 2.3.2. Evapotranspiration

The total ET demand is calculated as  $rET$ , the alfalfa (mature, 40 cm canopy height)-based reference ET, using the Penman–Monteith equation. The total ET demand is first used to evaporate any water in the canopy storage and on the ground. The ET demand on the impervious fraction of the cell is used to evaporate depression storage on the impervious fraction. This ET demand, if not depleted, is not used in any other way. The ET demand on the snow-covered fraction is also discarded as the energy is assumed to be used for melting the snow as described later. The maximum

Canopy storage of a cell is updated each day using its  $LAI$  using the Noilhan and Planton formulation [34]:

$$Wr_{mx} = 2 \times 10^{-4} \sum_{i=1}^{nRPT} \delta_i LAI_i \quad (2)$$

where  $Wr_{mx}$  is the maximum canopy storage (m) and  $LAI_i$  is the Leaf Area Index. The canopy interception is then calculated using a bucket model approach:

$$\frac{d(CS)}{dt} = \min \left( \frac{Wr_{mx} - CS}{\Delta t}, P \right) \quad (3)$$

in which,  $CS$  is the current canopy storage (m).

The transpiration for the  $j$ th soil layer is then computed as:

$$TP_j = \sum_{i=1}^{nRPT} Rt_{ij} \phi_j \delta_i \cdot K_c dET \quad (4)$$

where  $TP$  is the transpiration ( $m\ d^{-1}$ ),  $nRPT$  is the number of RPT in the current land cell,  $Rt$  is the root distribution function,  $\phi$  is the root efficiency function,  $\delta$  is the leaf-cover fraction of the  $i$ th RPT,  $K_c$  is the alfalfa-based crop coefficient whose value can be found from literature [33] and  $dET$  is the remaining ET demand after considering canopy storage and ground surface evaporation and snow cover. The leaf-cover fraction of the RPT is calculated using the Beer Lambert law:

$$\delta_i = f_i(1 - \exp(-0.5LAI_i)) \quad (5)$$

Plant roots have difficulty extracting water when soil moisture is low. The amount of water that roots can extract depends on the soil moisture and this relationship is called the root efficiency. The following root efficiency function is used, Lai and Katul [35]:

$$\phi(\theta) = \left( \frac{\theta - \theta_w}{\theta_s} \right)^{\Upsilon / (\theta - \theta_w)} \quad (6)$$

in which  $\theta$  is the soil moisture content,  $\theta_s$  is the saturated water content,  $\theta_w$  is the wilting point and  $\Upsilon$  is an empirical parameter. Braud et al. [36] examined this root efficiency function and found it to perform well compared to a field-measured soybean dataset.

The evaporation on the bare soil, on the other hand, is applied to the area that is not covered by leaves, for  $j$ th layer:

$$E_j = \phi_j^e [EDF(zd_j) - EDF(zd_{j-1})] \left( 1 - \sum_{i=1}^{nRPT} \delta_i \right) rET \quad (7)$$

The soil evaporation distribution function and water constraint function are taken from [5]:

$$EDF(d_z) = \frac{d_z}{d_z + \frac{\exp(2.374 - 7.13d_z)}{1000}} \quad (8)$$

$$\phi_j^e = \begin{cases} 1, & \theta \geq \theta_f \\ \exp \left( 2.5 \frac{\theta - \theta_f}{\theta_f - \theta_w} \right), & \theta < \theta_f \end{cases} \quad (9)$$

Here  $d_z$  is the depth of the soil layer from the ground surface and  $\theta_f$  is the field capacity. In PAWS, the transpiration and soil evaporation are applied in the vadose zone model as a source term in the Richards equation.

### 2.3.3. Vegetation

A simple vegetation growth module is included in the current version of PAWS but more complex functionality can be easily added. The  $LAI$ , rooting depth ( $RMX$ ), crop ET coefficient ( $K_c$ ) and canopy height ( $h_c$ ) are updated daily according to a piecewise linear function described in [37]:

$$V = \begin{cases} V_{\min} & JD < JP(1) \\ V_{\min} + \frac{JD - JP(1)}{JP(2) - JP(1)} (V_{\max} - V_{\min}) & JP(1) \leq JD < JP(2) \\ V_{\max} & JP(2) \leq JD < JP(3) \\ V_{\max} - \frac{JD - JP(3)}{JP(4) - JP(3)} (V_{\max} - V_{\min}) & JP(3) \leq JD < JP(4) \\ V_{\min} & JP(4) \leq JD \end{cases} \quad (10)$$

where  $V$  is the variable of either  $LAI$ ,  $RMX$ ,  $K_c$  or  $h_c$ ,  $V_{\max}$  is the maximum value of this variable,  $V_{\min}$  is the minimum value of the variable.  $JD$  is the Julian day, and  $JP$  are the Julian days of the control points,  $JP(1)$ ,  $JP(2)$ ,  $JP(3)$  and  $JP(4)$  which correspond, respectively, to the starting day of growth, the day on which the plant reaches its maximum canopy, the day on which the leaves begin to wilt, and the day on which canopy comes back to its minimum. Maximum and minimum  $LAI$ ,  $RMX$  and  $h_c$  have been obtained from literature values [5,32].

### 2.3.4. Overland flow

Overland flow is modeled using the two-dimensional Diffusive Wave equation (DWE) [38]:

$$\begin{aligned} \frac{\partial h}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} &= s \\ 0 &= -g \frac{\partial h}{\partial x} + g(S_{0x} - S_f) \\ 0 &= -g \frac{\partial h}{\partial y} + g(S_{0y} - S_f) \end{aligned} \quad (11)$$

where,  $h$  is the overland flow water depth (m),  $u$  and  $v$  are the  $x$ - and  $y$ -direction water velocities ( $m\ s^{-1}$ ),  $g$  is the gravitational acceleration ( $m\ s^{-2}$ ),  $s$  is the source term ( $m\ s^{-1}$ ) including runoff contribution from the ponding layer and evaporation,  $S_0$  is the slope ( $-$ ),  $S_f$  is the frictional slope ( $-$ ). The use of Eq. (11) implies sheet flow conceptualization of overflow. However, if we take  $h$  as an effective depth and slope  $S_0$  as the effective slope of the cell, Eq. (11) may also be interpreted as rill flow conceptualization, governed by the same equation on a macro-scale.

Surface flow components need to be run using small time steps in order to satisfy the Courant condition. A time step of 10 min is commonly employed (e.g. [31]). For long-term simulations, the computational efficiency of the surface flow modules is of great importance. An accurate, efficient and robust Runge–Kutta Finite Volume (RKfV) scheme is used to solve the Diffusive Wave equation for long term surface flow routing. Eq. (11) can be discretized in space as:

$$\frac{\partial h}{\partial t} = - \left[ \frac{h_{i-1/2,j} u_{i-1/2,j} - h_{i+1/2,j} u_{i+1/2,j}}{\Delta x} + \frac{h_{i,j-1/2} v_{i,j-1/2} - h_{i,j+1/2} v_{i,j+1/2}}{\Delta y} \right] + S_{ij} \quad (12)$$

We can compute  $u_{i+1/2,j}$  and  $v_{i,j+1/2}$ , the velocities at the cell interfaces, from  $h_{i+1/2,j}$  and  $h_{i,j+1/2}$  using the Manning's equation. Given the water depth on two sides of the boundary as  $h_L$  and  $h_R$ , and the free surface elevation as  $\eta_L$  and  $\eta_R$  ( $\eta = z + h$  where  $z$  is the elevation of the cell), we use the following logic to determine the flow depth at the boundary,  $h_{1/2}$ :

$$\eta_L > \eta_R? \begin{cases} \text{Yes}(Y), h_L > 0? \begin{cases} Y, C_d h_L > h_R? \begin{cases} Y, h_{1/2} = 0.5 * (h_L + h_R) \\ N, h_{1/2} = h_L \end{cases} \\ N, h_{1/2} = 0 \end{cases} \\ \text{No}(N), h_R > 0? \begin{cases} Y, C_d h_R > h_L? \begin{cases} Y, h_{1/2} = 0.5 * (h_L + h_R) \\ N, h_{1/2} = h_R \end{cases} \\ N, h_{1/2} = 0 \end{cases} \end{cases} \quad (13)$$



The interface flow depth in the  $y$  direction can be calculated using the same approach. The first condition,  $\eta_L > \eta_R$ ?, is to decide the upwind direction. For the DWE, the free surface elevation always determines the direction of flow. The second conditional statement is to decide if the upstream cell is dry. The third selection,  $C_d h_L > h_R$ ?, where  $C_d$  is a stability coefficient, is to ensure the interface flux does not exceed what is available for outflow during the explicit updating. Larger  $C_d$  values allow more frequent choice of the more accurate average depth, instead of the upwinding water depth which is diffusive but stable. From extensive numerical tests we found a value of  $C_d = 4$  works well for almost all scenario.

Once  $h_{1/2}$  is obtained, we use it in Eq. (14) to calculate interface velocities  $u$  and  $v$  using Manning's formula [38,39]:

$$u_{i+1/2,j} = -\text{sgn}(\eta_{i+1,j} - \eta_{i,j}) \frac{1}{n_m} h_{i+1/2,j}^{2/3} \left| \frac{\eta_{i+1,j} - \eta_{i,j}}{\Delta x_{i,j+1/2}} \right|^{1/2} \quad (14)$$

The  $y$ -direction velocity is computed similarly. After computing the velocities, Eq. (12) is used to march in time using the Runge–Kutta method. Such a design is chosen mainly for its efficiency, simplicity and stability which make it practical for large scale and long term simulations. A semi-implicit semi-Lagrangian (SISL) scheme [40] is also implemented to route water through surface water features such as shallow lakes and wetlands.

### 2.3.5. Channel flow and the river network

The channel flow model is based on the one dimensional Diffusive Wave equations:

$$\begin{aligned} \frac{\partial A_c}{\partial t} + \frac{\partial(uA_c)}{\partial x} &= \frac{P}{86400} w + q_{oc} + q_{gc} + q_t \\ 0 &= -g \frac{\partial h_r}{\partial x} + g(S_{0x} - S_f) \end{aligned} \quad (15)$$

where  $A_c$  is the cross-sectional area of the channel ( $\text{m}^2$ ),  $u$  is the flow velocity ( $\text{m s}^{-1}$ ),  $q_t$  is lateral inflow from overland flow ( $\text{m}^3 \text{m}^{-1} \text{s}^{-1}$ ),  $q_{gc}$  is the groundwater contribution ( $\text{m}^3 \text{m}^{-1} \text{s}^{-1}$ ),  $q_t$  is the contribution from tributaries and  $w$  is the width of the river. The equations are solved using the RKFDV scheme similar to the one described above for overland flow. We employ an operator splitting technique for the calculation of  $q_{gc}$ , which is computed after the rest of the terms are updated, as described later in Section 2.4. The SISL scheme is also available for the solution of full nonlinear Dynamic Wave equation, and it is more affordable in river routing because the problem is 1D in nature. However, for the results reported in this paper, the simple RKFDV is used in order to demonstrate its effectiveness.

The river models are run in an upstream–downstream cascade sequence, so the contributions from tributaries are always computed before the model is run for the main river. The tributary inflows are converted into a source term to the main river:

$$q_{t,i}^{n+1} \rightarrow q_{t,i}^{n+1} + \frac{1}{\Delta x_i \Delta t} \sum_{j=1}^{nt} \int_{t^n}^{t^{n+1}} q_{t,j} dt \quad (16)$$

In which,  $q_{t,i}$  is the source term throughout the new time step of the confluence cell on the main river ( $\text{m}^2 \text{s}^{-1}$ ),  $\Delta t$  and  $\Delta x$  are the temporal and spatial steps of the main river and  $\int_{t^n}^{t^{n+1}} q_{t,j} dt$  is the accumulative inflow of its  $j$ th tributary during this time step and  $nt$  denotes the number of tributaries. The above boundary condition implicitly assumes that the downstream river stage does not change much during the time step, which is generally valid since the main river is a larger river. During high flow periods the time step of the river network may need to be reduced to maintain stability.

### 2.3.6. Unsaturated vadose zone model

Vertical water moisture movement in the soil compartment is described by the Richards equation [41,42]:

$$C(h) \frac{\partial h}{\partial t} = \frac{\partial}{\partial z} \left[ K(h) \left( \frac{\partial h}{\partial z} + 1 \right) \right] + W(h) \quad (17)$$

in which,  $K(h)$  is the unsaturated hydraulic conductivity ( $\text{m d}^{-1}$ ),  $h$  is the soil water pressure head ( $\text{m}$ ),  $W(h)$  is the volumetric source or sink term, which includes contributions from evaporation and plant root extraction,  $z$  is the vertical coordinate (positive upward) ( $\text{m}$ ) and  $C(h) = \partial \theta / \partial h$ , the differential water capacity ( $\text{m}^{-1}$ ).

The unsaturated hydraulic conductivity and pressure head are both functions of the soil moisture content. We choose the Mualem–van Genuchten formulation:

$$\begin{aligned} \Theta &= \frac{\theta(h) - \theta_r}{\theta_s - \theta_r} = \left( 1 + |\alpha h|^N \right)^{-(N-1)/N} \\ K(h) &= K_s \Theta^\lambda \left[ 1 - (1 - \Theta^{N/(N-1)})^{(N-1)/N} \right]^2 \end{aligned} \quad (18)$$

where  $\Theta$  is the relative saturation,  $\theta$  is the soil moisture content,  $\theta_s$  is the saturated water content,  $\theta_r$  is the residual water content,  $N$  is a measure of the pore-size distribution,  $\alpha$  is a parameter related to the inverse of the air entry suction and  $\lambda$  is a pore tortuosity/connectivity parameter [43]. Here we employ the widely used implicit iteration scheme of Celia et al. [42].

$$\theta_i^{n+1} - \theta_i^n = C_i^{n+1,p-1} (h_i^{n+1,p} - h_i^{n+1,p-1}) + \theta_i^{n+1,p-1} - \theta_i^n \quad (19)$$

in which  $n$  denotes the time level,  $i$  denotes the soil layer number,  $p$  is the iteration number and  $C$  is evaluated as:

$$C^{n+1,p-1} = \frac{\partial \theta}{\partial h} = - \frac{(\theta_s - \theta_r)}{(1 + |\alpha h^{n+1,p-1}|^N)^{2N-1}} \frac{|\alpha h^{n+1,p-1}|^N (N-1)}{h^{n+1,p-1}} \quad (20)$$

where the soil layer number  $i$  is omitted. In other words, the differential water capacity  $C(\theta)$  is updated at every iteration using the newly obtained  $h^{n+1,p-1}$ . The iteration repeats until both  $h$  and  $\theta$  converge. Note that  $\theta^{n+1,0} = \theta^n$ . This stable implicit scheme can guarantee mass balance errors to be close to machine round off errors [41,42]. If we apply the procedure in Eq. (19) to (17) we obtain the following:

$$\begin{aligned} &\frac{1}{\Delta t} \left[ C_i^{n+1,p-1} (h_i^{n+1,p} - h_i^{n+1,p-1}) + \theta_i^{n+1,p-1} - \theta_i^n \right] \\ &= \left[ \frac{K_{i-1/2}^n}{\Delta z_i} \left( \frac{h_{i-1}^{n+1,p} - h_i^{n+1,p}}{\Delta z_{i-1/2}} + 1 \right) - \frac{K_{i+1/2}^n}{\Delta z_i} \left( \frac{h_i^{n+1,p} - h_{i+1}^{n+1,p}}{\Delta z_{i+1/2}} + 1 \right) \right] + W_i^n \end{aligned} \quad (21)$$

The above system can be solved efficiently using the Thomas algorithm [44].

### 2.3.7. Infiltration under heavy rain

Under normal conditions, infiltration is simulated simultaneously with the soil column by the Richards equation (RE), as described later in Section 2.4.2. Under heavy rainfall at dry surfaces, however, RE is expected to produce significant numerical error with the vertical discretization we can generally afford. This scenario is thus handled by the Green and Ampt equation (GA). The basic assumption of GA is a piston-type saturated wetting front. This means that (a) after water enters the soil column it fills any pore spaces between ground surface and wetting front, and (b) the infiltration rate is determined by Darcy's law using the effective saturated conductivity of the soil as well as suction head at the wetting front. Jia [45] developed a more generalized version of GA that is applicable for multiple layers to address the vertical heterogeneity of soils, and named the method Generalized Green and Ampt method (GGA).

In PAWS we invoke a modified version of Jia's GGA method when rainfall is heavy and the soil column is relatively dry. The modifications allow GGA and RE to be applied during the same

time step, each solving a portion of the soil column. Source term is also taken into account. The modified GGA proceeds as follows.

Let the wetting front position (*WFP*) be the location of the piston wetting front (m, measured from soil surface), in the *m*th soil layer. For a given soil column, the infiltrating capacity (*IFC*,  $\text{m d}^{-1}$ ), is only a function of *WFP* and can be calculated as [31]:

$$IFC = \frac{dF}{dt} = K_{S,m} \left( 1 + \frac{A_{m-1}}{B_{m-1} + F} \right) \quad (22)$$

in which,  $K_{S,m}$  is the saturated conductivity of the *m*th layer ( $\text{m d}^{-1}$ ),  $F(m)$  is the accumulated infiltration (or total moisture volume from surface to wetting front),  $\theta_{0,m}$  is the water content in layer *m* prior to the current time step and  $A_{m-1}$ ,  $B_{m-1}$  are coefficients:

$$\begin{aligned} A_{m-1} &= \left( L_{m-1} - \frac{K_{S,m} L_{m-1}}{\bar{k}_{m-1}} + SW_m \right) (\theta_{s,m} - \theta_{0,m}) \\ B_{m-1} &= \frac{K_{S,m} L_{m-1}}{\bar{k}_{m-1}} (\theta_{s,m} - \theta_{0,m}) - F_{m-1} \\ L_{m-1} &= \sum_{i=1}^{m-1} \Delta z_i, \quad \bar{k}_{m-1} = L_{m-1} / \left( \sum_{i=1}^{m-1} \frac{\Delta z_i}{K_{S,i}} \right), \\ F_{m-1} &= \sum_{i=1}^{m-1} \Delta z_i (\theta_{s,i} - \theta_{r,i}) \end{aligned} \quad (23)$$

$L_{m-1}$ ,  $\bar{k}_{m-1}$  and  $F_{m-1}$  are, respectively, depth to the bottom of the soil layer *m* – 1, harmonic mean saturated conductivities and accumulated infiltration of layers 1 through *m* – 1. To save computational time, they can be pre-computed and stored. The detailed derivations of the above terms are available in [45].  $SW_m$  is the average suction of the wetting front. As theoretically justified in [46],  $SW$  can be calculated as

$$SW(h) = \int_h^0 \frac{K(\xi)}{K_s} d\xi \quad (24)$$

For the van Genuchten hydraulic conductivity formulation in Eq. (18), the values of  $SW$  are numerically tabulated for different suction heads and soil parameters, and  $SW(h)$  is then queried from the table at the time of GGA execution.

Two flow conditions are distinguished: (a) infiltration under maximum capacity (*IFC*) (b) infiltration rate equals the rainfall rate. When there is ponding or rainfall exceeds *IFC*, infiltration proceeds under maximum capacity according to Eq. (22). Integrating Eq. (22), we obtain a nonlinear equation:

$$(F - F^{p-1}) = k_m \Delta t + A_{m-1} \ln \left( \frac{A_{m-1} + B_{m-1} + F}{A_{m-1} + B_{m-1} + F^{p-1}} \right) \quad (25)$$

in which  $F^{p-1}$  is the cumulative infiltration from the previous time step.  $F^{p-1}$  may be computed from a fractional time step that is smaller than the macro model time step, as the flow conditions may change a few times in a macro time step.  $F$  can be solved for using Newton iteration. Eq. (25) is valid only when *WFP* is within the *m*th layer and water is infiltrating under maximum capacity. It is possible that present ponding is exhausted during the time step. We may predict the time at which this happens by solving the following nonlinear equation:

$$\begin{aligned} h_1 + P \Delta t_p &= k_m \Delta t_p + A_{m-1} \\ &\times \ln \left( \frac{A_{m-1} + B_{m-1} + F^{p-1} + h_1 + P \Delta t_p}{A_{m-1} + B_{m-1} + F^{p-1}} \right) \end{aligned} \quad (26)$$

In this case,  $\Delta t_p$  is the time after which the ponding layer becomes dry. It is also possible that the current soil layer is filled up within the time step and the wetting front moves to the next unsaturated layer. Since soil properties may change when wetting front moves into the next layer, this time,  $\Delta t_f$ , may be solved from Eq. (25), substituting  $F$  with  $F_m$ .

On the other hand, if there is no ponding (case b) and rainfall is less than *IFC*, infiltration rate is simply the rain intensity:  $F = F^{p-1} + \Delta t P$ . As *IFC* decreases with infiltration, we need to predict the time to ponding and time to fill the *m*th cell:

$$\Delta t_p = \frac{1}{P} (F_p - F^{p-1}), \quad F_p = \frac{A_{m-1}}{P/k_m - 1} - B_{m-1} \quad (27)$$

In this case,  $\Delta t_p$  is the time to ponding. For time coordination with other components of the model, we first calculate  $\Delta t = \min(\Delta t_f, \Delta t_p, t_{end} - t)$ , ( $t_{end}$  is the prescribed ending time of the time step) and use the corresponding equation to calculate  $F$ . The algorithm may need to go through several fractional steps in order to switch between different infiltration rates or soil layers to arrive at the prescribed ending time. Source/sink terms, which are generally small during the heavy rain period, are added to (or subtracted from) the rainfall rate. This completes the description for the modified GGA method.

After GGA is solved, the average soil moisture of the *WFP* cell is calculated and provided to the Richards equation as initial condition. The RE updates soil moisture from the *m*th layer to the bottom layer, using a zero-flux top boundary condition. This is similar to using an operator-splitting method. The resulting soil moisture will be converted back to *WFP* and  $\theta_m$  in a mass-conservative fashion in the next time step if GGA is again invoked.

### 2.3.8. Saturated groundwater flow model

The saturated aquifers are conceptualized as a series of vertical layers. In each vertical layer, we solve the 2-dimensional groundwater equation:

$$S \frac{\partial H}{\partial t} = \frac{\partial}{\partial x} \left[ T \left( \frac{\partial H}{\partial x} \right) \right] + \frac{\partial}{\partial y} \left[ T \left( \frac{\partial H}{\partial y} \right) \right] + R + W - Dp \quad (28)$$

where  $S$  is the storativity (dimensionless),  $T = Kb$  is the transmissivity of the aquifer ( $\text{m}^2 \text{d}^{-1}$ ), where  $K$  is the horizontal saturated hydraulic conductivity ( $\text{m d}^{-1}$ ) and  $b$  is the saturated thickness of the aquifer (m),  $H$  is hydraulic head (m),  $R$  is recharge or discharge ( $\text{m d}^{-1}$ ),  $W$  is the source and sink term due to pumping, root extraction and exchange with channel flow ( $\text{m d}^{-1}$ ) (inflow as positive) and  $Dp$  is percolation into deeper aquifers ( $\text{m d}^{-1}$ ). The above equation is discretized using standard backward-in-time, center-in-space finite difference scheme:

$$\begin{aligned} \frac{S_{ij}}{\Delta t} (H_{ij}^{n+1} - H_{ij}^n) &= \left[ \frac{T_{i-1/2,j}}{\Delta x_{ij}} \left( \frac{H_{i-1,j}^{n+1} - H_{i,j}^{n+1}}{\Delta x_{i-1/2,j}} \right) - \frac{T_{i+1/2,j}}{\Delta x_{ij}} \left( \frac{H_{i,j}^{n+1} - H_{i+1,j}^{n+1}}{\Delta x_{i+1/2,j}} \right) \right] \\ &+ \left[ \frac{T_{i,j-1/2}}{\Delta y_{ij}} \left( \frac{H_{i,j-1}^{n+1} - H_{i,j}^{n+1}}{\Delta y_{i,j-1/2}} \right) - \frac{T_{i,j+1/2}}{\Delta y_{ij}} \left( \frac{H_{i,j}^{n+1} - H_{i,j+1}^{n+1}}{\Delta y_{i,j+1/2}} \right) \right] \\ &+ R_{ij}^n + W_{ij}^n - Dp_{ij}^n \end{aligned} \quad (29)$$

The thickness of the unconfined aquifer is taken as  $b = H^{n-1} - z_0$  where  $H^{n-1}$  is the water table elevation solved from the last time step and  $z_0$  is the bottom elevation of the unconfined aquifer. Here  $i$  and  $j$  are  $x$  and  $y$  coordinate indices, respectively and  $n$  is the time level.  $W_{ij}^n = \int_{t_n}^{t_{n+1}} W_{ij} dt$  is the integral source term throughout the time step.

We solve for the percolation term implicitly:

$$Dp_{ij}^n = -K_l \frac{(H_{ij}^{n+1} - H_{ij}^n)}{\Delta z_l} \quad (30)$$

where  $K_l$  is the vertical hydraulic conductivity ( $\text{m d}^{-1}$ ) of the aquitard beneath the current aquifer,  $\Delta z_l$  is the thickness of the aquitard,

$h_i^n$  is the hydraulic head in the aquifer below. The  $x$ -direction interface transmissivity is taken as:

$$\frac{T_{i-1/2,j}}{\Delta x_{ij}} = \frac{2T_{ij}T_{i-1,j}}{\Delta x_{ij}T_{ij} + \Delta x_{i-1,j}T_{i-1,j}} \quad (31)$$

The  $y$  direction interface transmissivity is defined in a similar fashion. The resulting matrix system is solved using the Conjugate Gradient algorithm [44].

### 2.3.9. Snowpack model

The UEB (Utah Energy Balance) snow model [47–49] is adapted as the snowpack module of the new model. This model keeps track of mass and energy balance of the snowpack as well as snow age (for the calculation of snow albedo). For snowmelt calculations, the model uses physically-based accounting of radiative, sensible, latent and advective heat exchanges. The version we adapted uses a lumped representation of sub-grid variability [47,49] and simulates the hysteresis effect caused by new snowfall and snowmelt. The governing equations of the UEB model are:

$$\begin{aligned} \frac{dU}{dt} &= Q_{sn} + Q_{li} + Q_p + Q_g - Q_{le} + Q_h + Q_e - Q_m \\ \frac{d(SWE)}{dt} &= P - SNOM - E_{sub} \end{aligned} \quad (32)$$

In the energy balance equation, the terms are:  $U$ , the snowpack energy content of the snow-covered area ( $\text{kJ m}^{-2}$ ), all other terms in this equation are in  $\text{kJ m}^{-2}\text{d}^{-1}$ ;  $Q_{sn}$ , net shortwave radiation;  $Q_{li}$ , incoming longwave radiation;  $Q_p$ , advected heat from precipitation;  $Q_g$ , ground heat flux;  $Q_{le}$ , outgoing longwave radiation;  $Q_h$ , sensible heat flux;  $Q_e$ , latent heat flux due to sublimation/condensation; and  $Q_m$ , advected heat removed by meltwater. In the mass balance equation, the terms are:  $SWE$ , snow water equivalent (m);  $SNOM$ , meltwater outflow from the snowpack ( $\text{m d}^{-1}$ ); and  $E_{sub}$ , sublimation from the snowpack ( $\text{m d}^{-1}$ ). The system of two equations is solved using a predictor–corrector scheme. More details of the snow model are available in [47,49].

### 2.4. Interactions among domains

The above sections complete the descriptions of solvers for individual domains. The hydrologic system is a fully coupled system such that fluxes in one domain are linked to those in other domains. In this section we discuss how these different domains are coupled in the model. In particular, we describe the theoretical development of the coupling between the vadose zone and the saturated groundwater flow, which serves as a major step to reduce computational demand by avoiding iteration at the equation level.

#### 2.4.1. Coupling of the unsaturated Richards equation and the groundwater flow equation

Consider the fully three-dimensional Richards equation:

$$\begin{aligned} (C(h) + S_s) \frac{\partial h}{\partial t} &= \left[ \frac{\partial}{\partial z} \left( K_z(h) \left( \frac{\partial h}{\partial z} + 1 \right) \right) + \frac{\partial}{\partial x} \left( K_x(h) \frac{\partial h}{\partial x} \right) \right. \\ &\quad \left. + \frac{\partial}{\partial y} \left( K_y(h) \frac{\partial h}{\partial y} \right) \right] + W(h) \end{aligned} \quad (33)$$

where  $h$  is the pressure head (m),  $K_x$ ,  $K_y$ ,  $K_z$  are hydraulic conductivities ( $\text{m d}^{-1}$ ),  $S_s$  is the specific storage ( $\text{m}^{-1}$ ) and  $W(h)$  is the volumetric source or sink term ( $\text{d}^{-1}$ ). We seek to lower the dimensionality of Eq. (33) by separating it into the one-dimensional Richards equation that governs the unsaturated soil column (17) and the two-dimensional groundwater flow equation that governs flow in the saturated unconfined aquifer (28). One approach is to solve the Richards Eq. (17), taking the solution to the groundwater flow Eq. (28) from last time step as a Dirichlet boundary condition.

Several examples of this approach are available [27,31,50]. This method implicitly assumes that the head in the unsaturated zone does not change throughout the time step. However, such an assumption does not hold when recharge from the vadose zone or lateral flow of groundwater cause large changes in the water table position and may induce numerical instabilities. In addition, some theoretical inconsistency occurs as a head difference is always required to maintain recharge to the aquifer. We will illustrate this in Section 3.1.4.

Another method is to iteratively solve the two equations until their solutions converge. Some researchers [51,52] coupled the two equations via the deep percolation flux at the bottom of the saturated domain and the vadose zone model. When iteration is involved on the equation level, the method can be much more expensive. Moreover, the coupling is ad hoc as apparently, if an impervious layer sits at the bottom and the deep percolation is zero, such a coupling is not possible. Here we seek a method that avoids equation-level iterations.

We conceptualize the three-dimensional unsaturated vadose zone as an array of one dimensional soil columns connected to the unconfined aquifer at the bottom. Two assumptions are invoked to make the problem more manageable as illustrated in Fig. 2. The first assumption is that water moves only vertically in the unsaturated zone. Denoting the lateral flow into a unit volume of the soil column as  $q$  ( $\text{T}^{-1}$ ), this can be written as Assumption 1:

$$q(x, y, z) = 0, \quad z > H(x, y) \quad (34)$$

where  $q = \frac{\partial}{\partial x} (K_x(h) \frac{\partial h}{\partial x}) + \frac{\partial}{\partial y} (K_y(h) \frac{\partial h}{\partial y})$  and  $H(x, y)$  is the elevation of the water table. For simplicity, we omit the coordinates  $(x, y)$ . The rationale behind this assumption is that in the unsaturated domain gravity dominates over lateral moisture diffusion.

The next assumption is that the lateral flow is uniformly distributed along the saturated thickness of the unconfined aquifer, Assumption 2:

$$\frac{\partial q}{\partial z} = 0 \quad \text{and} \quad \int_{z_0}^H q dz = DR, \quad z_0 \leq z \leq H \quad (35)$$

where  $z_0$  is the bottom elevation of the unconfined aquifer,  $DR$  is the integrated lateral drainage term ( $\text{m d}^{-1}$ ), positive for inflow. Integrating Eq. (33) from  $z_0$  to  $H$  and using (35) and (28), we get:

$$DR \approx \frac{\partial}{\partial x} \left[ T \left( \frac{\partial H}{\partial x} \right) \right] + \frac{\partial}{\partial y} \left[ T \left( \frac{\partial H}{\partial y} \right) \right] \quad (36)$$

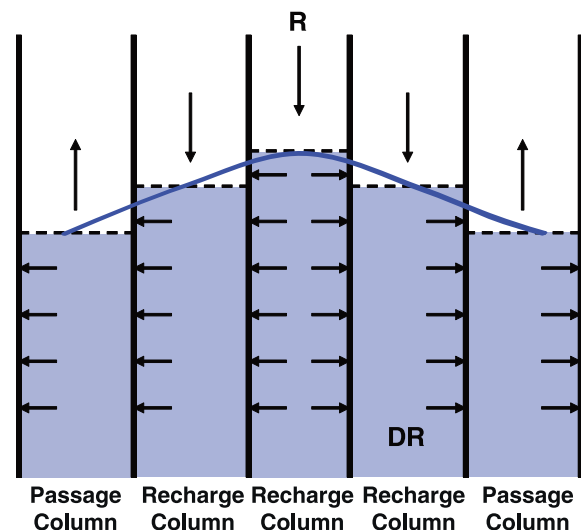


Fig. 2. Illustration of Assumptions 1 and 2 for the unsaturated/saturated coupling scheme.

Solving for  $q$  from (35), we get:

$$q = \frac{DR}{b}, \quad z_0 \leq z \leq H \quad (37)$$

where again  $b = (H - z_0)$  is the thickness of the unconfined aquifer (m). This condition was used by several researchers earlier to bring the effects of regional groundwater flow into the Richards equation, e.g. [52]. Assumption 2 is similar to, but not the same as the Dupuit-Forchheimer (DF) assumption. The original DF assumption states that the vertical hydraulic gradient is zero in the saturated zone, i.e.  $\partial(h+z)/\partial z = 0$  when  $z_0 \leq z \leq H$ . This assumption is not valid when lateral flow is present. In order to provide a physically consistent lower boundary condition, we consider the mass balance of a unit volume in the saturated portion at steady state, using Eqs. (33), (36) and assuming  $W' = 0$  and a constant  $K_z$  in the saturated zone:

$$(C(h) + S_s) \frac{\partial h}{\partial t} = 0 = \left[ \frac{\partial}{\partial z} \left( K_z \left( \frac{\partial h}{\partial z} + 1 \right) \right) + q \right] \quad (38)$$

Considering deep percolation (from the bottom of the unconfined aquifer), the above equation forms a second order boundary value problem (BVP) that can be easily solved. The boundary conditions for the BVP are:

$$K_z \left( \frac{\partial h}{\partial z} + 1 \right) \Big|_{z=0} = q_{bot} \quad (39)$$

$$h(x, y, H) = 0$$

The steady state solution to the above system is:

$$h'(x, y, z) = \left( 1 - \frac{q_{bot}}{K_z} \right) (H - z) + \frac{1}{2} \frac{q}{K_z} [(H - z_0)^2 - (z - z_0)^2] \quad (40)$$

where  $q_{bot}$  is the deep percolation ( $\text{m d}^{-1}$ , positive for downward percolation) and  $z$  is the vertical coordinate. As we can see,  $h'$  is a quadratic function of  $H$  when head loss due to drainage is considered instead of the linear relationship that results from the DF assumption. Fig. 3 shows an example of  $h'$  as a function of  $z$  for positive, zero and negative  $DR$  values. Although (40) is derived from the steady state case, we will use it in a quasi-steady manner using the latest values of  $q$  and  $H$ .

With the above two assumptions and Eq. (40) we are ready to lay out the steps of our coupling scheme. The 3D Richards Eq. (33) is re-written for the soil column as:

$$(C(h) + S_s) \frac{\partial h}{\partial t} = \frac{\partial}{\partial z} \left( K_z(h) \left( \frac{\partial h}{\partial z} + 1 \right) \right) + q + W'(h) \quad (41)$$

This equation is simply the 1D Richards Eq. (17) with the addition of a drainage flux term:

$$q(x, y, z) = \begin{cases} 0 & z > H \\ DR/b & z_0 \leq z \leq H \end{cases} \quad (42)$$

The last computational cell of the soil column serves as the linkage between the two flow domains. It extends from the bottom of the unconfined aquifer to  $z_u$ , a cell boundary immediately below the water table, and fluctuates with the water table. As the water table progresses or retreats from the soil column cells, the saturated zone must share at least one cell with the soil columns. We use this cell to simulate the unconfined aquifer and to bring the dynamics of the regional groundwater flow into the soil column. The center of this cell should always be below the water table. Integrating Eq. (41) from  $z_0$  to  $z_u$  leads to the following mass balance equation:

$$S_{nz} \frac{\partial}{\partial t} \int_{z_0}^{z_u} h \, dz \approx S_{nz} \frac{\partial h_{nz}}{\partial t} = K_z \left( \frac{\partial h}{\partial z} + 1 \right) \Big|_{z_0}^{z_u} - K_z \left( \frac{\partial h}{\partial z} + 1 \right) \Big|_{z_u} + \frac{z_u - z_0}{b} DR + \int_{z_0}^{z_u} W' \, dz \quad (43)$$

The parameter  $S_{nz}$  mimics the response of the unconfined aquifer, whose stage will rise after receiving recharge.

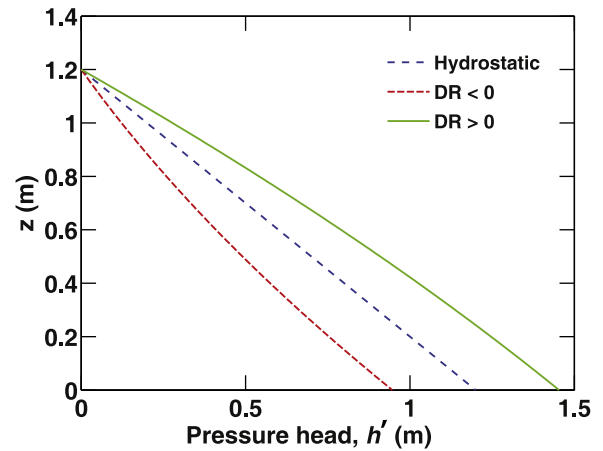


Fig. 3. Pressure head  $h$  as a function of  $z$  and  $DR$  values.

The next step is to solve Eq. (43) for the last cell together with the unsaturated zone Eq. (41). Applying a semi-implicit discretization to Eq. (43), we obtain:

$$\frac{S_{nz}}{\Delta t} (h_{nz}^{n+1} - h_{nz}^n) = K_{nz-1/2} \left( \frac{h_{nz-1}^{n+1} - h_{nz}^{n+1}}{\Delta z_{nz-1/2}} + 1 \right) - K_l \left( \frac{h_{nz}^{n+1} - h_l^n}{\Delta z_l} + 1 \right) + \frac{z_u - z_0}{b} DR + (z_u - z_0) W_{nz}^m \quad (44)$$

in which,  $nz$  is the index of the bottom cell,  $S_{nz}$  is the storativity (dimensionless),  $h_l^n$  is the pressure head of the confined aquifer beneath the unconfined aquifer, and  $K_l$  is the hydraulic conductivity of the aquitard between the unconfined and confined aquifers. We treat  $h_l^n$  and  $DR$  explicitly because their temporal changes tend to be much smaller than the rest of the terms. This equation is linked to the state of the groundwater by the inclusion of  $h_{nz}^n$ , which we compute from Eq. (40), substituting  $z$  with  $z_{nz}$ , the elevation of the center of the last cell.

Both  $H$  and the lateral flow term  $DR$  in (42) have already been computed in the groundwater flow model in (29), lagged by one time step. Thus Eq. (44) presents a lower boundary condition to Eq. (41).

After this equation is solved, we calculate the recharge as the water leaving the bottom of the soil column:

$$R = \int_{z_u}^H q \, dz + K_{nz-1/2} \left( \frac{h_{nz-1}^{n+1} - h_{nz}^{n+1}}{\Delta z_{nz-1/2}} + 1 \right) \quad (45)$$

The coupling procedure can be summarized as follows:

1. At the beginning of each time step, extract the water table location  $H$ , and the drainage flux  $DR$  from the groundwater flow solution from last time step
2. Use Eqs. (40) and (37) to calculate  $h_{nz}$  and  $q$
3. Solve the system resulting from (41) and (43) using the method described in Section 2.3.6.
4. Calculate  $R$  with (45), accumulate it over time, and divide the result by the time step of the groundwater solver
5. With the predicted  $R$ , solve the groundwater flow Eq. (28)
6. Calculate  $DR$
7. Go to step 1 and repeat

In essence, we use the last cell to simulate the behavior of the groundwater system and to quantify the flux between the vadose zone and the saturated zone. Thus iteration on the equation level is avoided. Although the head in this last cell is updated while solving the Richards equation in the vadose zone, its final state is not



important. Only the flux through the interface is used later to calculate recharge in the groundwater model (29). An advantage of the approach is that since changes in the groundwater system tend to be slow, Eq. (36) calculated from the last time step represents a good approximation to  $DR$ . Thus we have decomposed the full 3D system into simpler, low-dimensional subsystems to enable large scale simulation, while retaining the salient features of the 3D equation. The actual performance of this approach will be described in a later section.

#### 2.4.2. Interaction between the surface ponding layer and soil moisture

The surface ponding layer mass balance Eq. (1) is solved simultaneously with the soil column:

$$\frac{dh_1}{dt} = SS_1 - K_{1/2} \left( \frac{h_1 - h_2}{\Delta z_{1/2}} + 1 \right) - F_g \quad (46)$$

where,  $h_1$  is the surface ponding depth (the topmost cell of the soil column),  $h_2$  denotes the pressure head of the second cell of the soil column,  $SS_1 = P + SNOM - dCS/dt - E_g$  is the source term accounting for precipitation, snowmelt, canopy storage and ground surface evaporation,  $F_g$  is surface runoff ( $m d^{-1}$ ) and  $K_{1/2}$  is the surface hydraulic conductivity, which is calculated as the geometric mean of the saturated and the current-state unsaturated conductivity of the first soil layer (layer 2), considering the fraction of pervious area:

$$K_{1/2} = (1 - f_{imp}) \sqrt{K_1 K_{S,1}} \quad (47)$$

The surface runoff is the contribution from surface ponding to various flow paths. If we assume the total length of overland flow paths in the cell as  $l$ , as described in [24,53], the surface runoff contribution can be computed as:

$$F_g(h_1) = \frac{(h_1 - h_o) u_l l}{86400 A} \quad (48)$$

where  $u_l$  is the runoff flow velocity ( $m s^{-1}$ ),  $A$  is the area of the cell ( $m^2$ )  $A = \Delta x \Delta y$ ,  $h_o$  is a minimum depth of water for surface runoff to occur (understory and ground interception, m) and the constant 86400 is for unit conversion from  $m s^{-1}$  to  $m d^{-1}$ .  $l$  is the average length of surface flow paths which can be parameterized based on land surface characteristics and adjusted during calibration.

We use the Manning's formula to compute flow velocity and apply the kinematic wave concept ( $S_0 = S_f$ ):

$$u_l = \frac{1}{n_m} (h_1 - h_o)^{2/3} S_0^{1/2} \quad (49)$$

where  $S_0$  is the average slope of the cell and  $n_m$  is the Manning's roughness coefficient. Eq. (46) is a nonlinear equation that can be solved iteratively together with the rest of the soil profile. However, if the surface is dry or the ponding water is exhausted in the time step, a flux boundary condition is applied instead.

#### 2.4.3. Interaction between overland flow and channel flow

A close-up view helps illustrate the exchange between the river and the land cells (Fig. 4). Interactions between overland flow and the channel can be modeled using the framework for flow over a wide rectangular weir [24]. We have developed an efficient and stable procedure to compute river/land exchanges on a physical basis using this approach. The cross-sectional and plane sketches of a river cell are given in Fig. 4.  $z_{Bank}$  is the elevation of the bank (m),  $h$  is the depth of the overland flow (m) as defined in Eq. (11),  $z$  is the average elevation of the cell (m),  $\eta$  is the average free surface elevation of the cell (m),  $h_r$  is the channel flow depth and  $\eta_{ch}$  is the stage of the channel (m). The exchange volume  $M$  ( $m^3$ ) between land and channel is computed with the procedure outlined in Fig. 5.

In the diagram, the first condition (1) is to determine the direction of the flow. If it is from the land to the river ( $\eta > \eta_{ch}$ ) and there

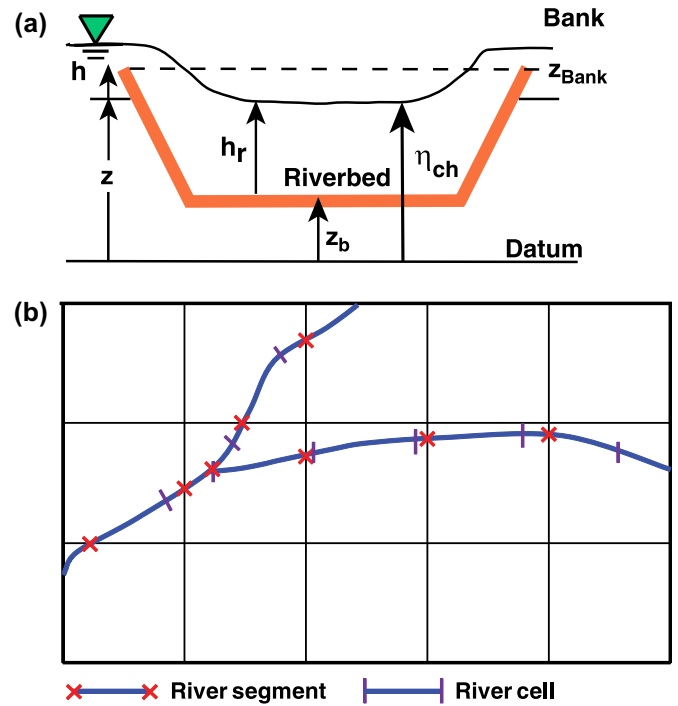


Fig. 4. Sketch of a river cell in the model: (a) cross-sectional view (b) plane view.

is water on the land (condition 2), we first attempt to compute the contribution explicitly using the diffusive wave equation:

$$M_{ex} = \frac{2L_c \Delta t}{n_m} h^{5/3} \left| \frac{\partial \eta}{\partial x} \right|^{1/2} = \frac{2L_c \Delta t}{n_m} h^{5/3} \left| \frac{\eta - \max(\eta_{ch}, z_{Bank})}{(\Delta x/2)} \right|^{1/2} \quad (50)$$

in which,  $L_c$  is the length of the channel segment that overlaps with the land cell (m), and other terms are defined as in Fig. 4. However, this flux cannot exceed the amount of currently available water on the land cell,  $M_a = A \cdot h$ .

Also, there is an equilibrium state, at which the river stage will be the same as the land free surface elevation. Denoting this stage as  $\eta^*$ , the mass transfer equation is written as:

$$(\eta^* - \eta_{ch}) A_b = (\eta - \eta^*) A \quad (51)$$

where  $A$  is again the area of the land cell and  $A_b$  is the area of the river segment that spans this land cell. Then we can find the mass transfer that leads to  $\eta^*$ :

$$M_E = (\eta^* - \eta_{ch}) A_b = \left( \frac{\eta A + \eta_{ch} A_b}{(A + A_b)} - \eta_{ch} \right) A_b = \frac{(\eta - \eta_{ch}) A_b}{(1 + A_b/A)} \quad (52)$$

Thus the exchange mass will be the minimum of  $M_{ex}$ ,  $M_E$ ,  $M_a$ .

On the other hand, if the river stage rises higher than the land free surface elevation (and also the bank elevation), flooding would occur, which is solved using a backward Euler implicit approach to enhance stability. The exchange mass solved by the implicit method is denoted as  $M_{im}$ . Again using the diffusive wave formulation, the mass exchange function can be written in the form of two ordinary differential equations:

$$\begin{aligned} A \frac{d\eta}{dt} &= -\frac{M_{im}}{\Delta t} = -\frac{2L_c}{n_m} \frac{\eta_{ch} - \eta}{\Delta x/2} (\eta_{ch} - z_{Bank})^{5/3} \\ A_b \frac{d\eta_{ch}}{dt} &= -A \frac{d\eta}{dt} \end{aligned} \quad (53)$$

These two equations can be solved by either Picard or Newton iteration. The resulting scheme is very stable.

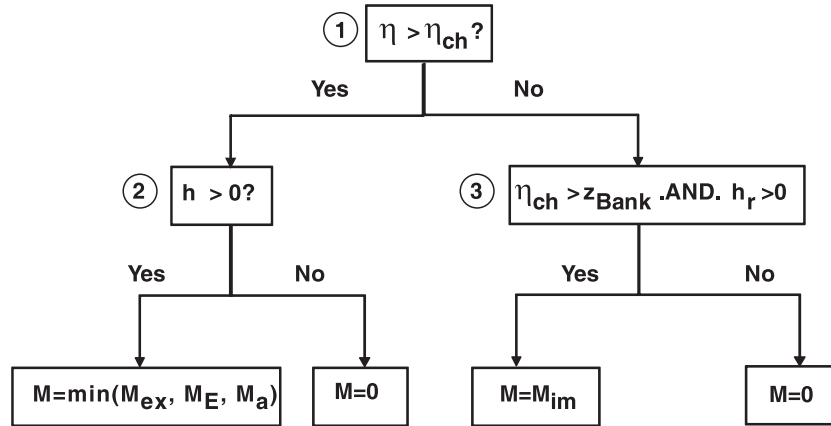


Fig. 5. Calculation of river/land exchange.

#### 2.4.4. Interaction between groundwater and channel flow

Equations describing interactions between groundwater flow and channel flow are solved immediately after the channel flow step. We use operator splitting to couple the river flow and exchange with groundwater. After the river flow model is solved explicitly with a Runge–Kutta approach, an implicit step is used to calculate the exchange between streams and groundwater using the conductance concept [53]. The governing equation for the fractional step is written as:

$$\frac{dh_r}{dt} = \begin{cases} K_r \frac{H - (z_b + h_r)}{\Delta z_b} & H > (z_b - \Delta z_b) \\ -K_r \frac{(h_r + \Delta z_b)}{\Delta z_b} & H \leq (z_b - \Delta z_b) \end{cases} \quad (54)$$

in which,  $z_b$  is the river bed elevation (m),  $\Delta z_b$  is the thickness of the river bed material (m),  $K_r$  is the river bed conductivity ( $\text{m s}^{-1}$ ) and  $H$  is again the groundwater table elevation (m). The above equation can be discretized implicitly as:

$$\frac{h_r^{n+1} - h_r^*}{\Delta t} = K_r \frac{H^* - (z_b + h_r^{n+1})}{\Delta z_b} \quad (55)$$

$h_r^*$  is computed from the solution to Eq. (15) (without the  $q_{gc}$  term). A similar equation can be written for the groundwater table  $H^*$  which can be solved simultaneously with (55). Then the lateral flow from groundwater is computed as:

$$q_{gc} = w_R (h_r^{n+1} - h_r^*) \quad (56)$$

where  $w_R$  is the wetted perimeter (m), approximately the river width for rivers wider than 10 m.  $q_{gc}$  is accumulated and subtracted from the source term in the groundwater flow equation.

The model marches in time using calendar days. The hydrologic processes, including evaporation, transpiration, snowpack and unconfined aquifer are updated on an hourly basis. The vadose zone model solves infiltration, runoff and depression storage together with soil moisture. It uses an hour as a base time step but adjusts it depending on flow conditions and convergence rate, as in [41]. The overland flow and river network also have the ability to adaptively select time steps to ensure stability and computational efficiency.

The computational subroutines of PAWS are written in FORTRAN while data preparation, certain control functions and display utilities are written in MATLAB®. PAWS allows model variables, including spatial fields of fluxes and state variables, river stages and time history to be displayed real-time. The goal is to keep the model as an open-ended, dynamically-evolving and interactive environment in which the modeler can be engaged with the modeling process on a real-time basis. A software package with

Graphical User Interface (GUI) has been developed for the present model to help interface with data. PAWS was developed using an objected-based programming approach, although it does not explicitly enforce concepts such as encapsulation and inheritance. The aim was to easily allow the addition of new functionality while keeping the code modular.

### 3. Results and discussion

#### 3.1. Test cases

Numerical codes must be carefully tested and verified before they can be put to use. Following the Freeze and Harlan blueprint [54], we require that each component be independently verified. There are two levels of code testing. In the first level, numerical code is compared to available analytical solutions to ensure that there are no bugs in the code and that the numerical schemes solve the PDEs with acceptable accuracy (within the range of parameters covered by the test problem). At this level a good agreement between the numerical and analytical solutions is expected. In the second level, the numerical code is compared against experimental data to examine the extent to which the idealized governing equations approximate the real-world. Due to experimental error and simplifications of the processes, some deviations are commonly observed.

In this section, we present comparisons with analytical solutions or experimental data for all of the flow domains. The coupling between overland flow and channel flow is tested together with overland and channel flow codes in the V-catchment test problem. Since analytical solutions for 3D saturated/unsaturated flow are not available, the coupling between the vadose zone and groundwater flow is tested by comparing model results with experimental data reported in Vauclin et al. [55] as well as numerical solutions based on the fully 3D equations for the subsurface. More details of the test cases are available in Shen [56].

##### 3.1.1. V-catchment

The V-catchment problem is a standard test case for overland flow models [57,58]. The domain consists of two inclined planes draining into a sloping channel. Both of the planes are 800 m in the lateral direction and 1000 m in the longitudinal direction and the slope is 0.05. The channel has a slope of 0.02. The Manning's coefficient  $n_m$  used was 0.015 for the planes and 0.15 for the channel. Due to symmetry, only half of the domain needs to be modeled. Flow over the inclined plane was solved using the 2D RKFFV scheme and the channel flow was computed on a separate channel grid, using the channel version of the RKFFV scheme. The solutions

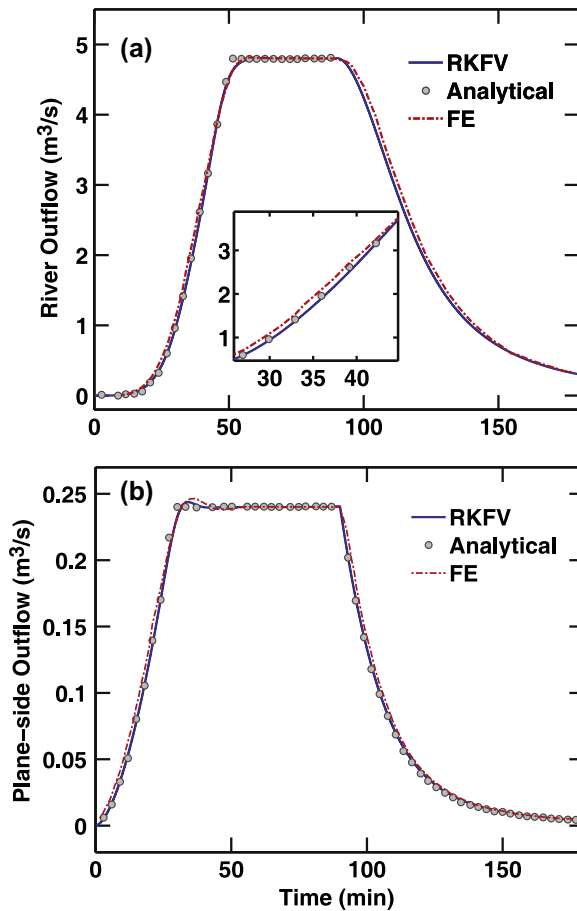


Fig. 6. Solution to the V-catchment test problem: comparison of river outflow hydrograph with analytical solution and the finite element (FE) solution reported in [57]. (a) River outflow (b) Plane side outflow. A definition sketch showing the geometry of the test problem is included in the Supplemental Information.

obtained from using the SISL scheme are almost identical and thus will not be reported here. In order to test not only the overland and river flow models but also the land/river exchange scheme of the proposed scheme, we used the same procedure described in Section 2.4.3. The land grid uses a  $\Delta x$  of 50 m and a  $\Delta t$  of 5 s. Because the channel has a milder slope and a large  $n_m$ , a  $\Delta t = 10$  s was used. The channel hydrograph is shown in Fig. 6a, and compared to the analytical solution and Finite Element (FE) solution reported in [57]. The analytical solution is not available for the declining limb of the hydrograph. We observe that the RKFV scheme matches the analytical solution very well during the initial rising period, slightly better than the FE solution. The plane-side outflow rate, the rate at which land contributes to the river, is shown in Fig. 6b. There is a small oscillation around  $t = 30$  min due to the mass exchange scheme. This is more pronounced in the FE solution. In watershed-scale modeling, such a small feature is expected to have little or no impact [56]. We also note that the declining limb is captured noticeably better by the RKFV scheme. The accuracy, simplicity and computational efficiency of the RKFV scheme make it attractive for watershed modeling.

### 3.1.2. Infiltration test problem

This widely-used test problem demonstrates the ability of the 1D Richards equation to predict infiltration into very dry soils [42,58]. It is one of the few analytical solutions available to the Richards equation. The soil column in the test problem has a total length of  $L = 0.3$  m, with the initial condition:

Table 3

van Genuchten soil parameters for the test cases.

Test case (Section)	$\theta_r$ (–)	$\theta_s$ (–)	$\alpha$ (m <sup>–1</sup> )	$N$ (–)	$K_s$ (m d <sup>–1</sup> )	$\lambda$ (–)
Infiltration (3.1.2)	0.102	0.368	3.35	2.0	7.97	0.5
Vauclin (3.1.4)	0.01	0.30	3.30	4.1	8.40	0.5
Sulis (3.1.5)	0.08	0.40	1.00	2.0	1.00 <sup>a</sup>	0.5

<sup>a</sup> Varied in cases (c) and (d) in Section 3.1.5.

$$h(z, 0) = -10 \text{ m} \quad (57)$$

and Dirichlet boundary conditions were used on both ends:

$$\begin{aligned} h(0, t) &= -0.75 \text{ m} \\ h(-L, t) &= -10 \text{ m} \end{aligned} \quad (58)$$

The soil parameters for this test case are given in Table 3. The solutions (pressure head) obtained with three different spatial step sizes ( $\Delta x = 0.075, 0.25$  and  $0.6$  cm) and the analytical solution at the final simulation time  $t = 6$  h are presented in Fig. 7. As noted by many other studies, the vertical resolution does have an impact on the solution quality [25]. The solver performs very well at  $\Delta x = 0.075$  cm and  $0.25$  spatial resolution. The quality of the solution for  $\Delta x = 0.25$  is very similar to that reported in [59]. A coarse spatial resolution will not be able to resolve the sharp wetting front predicted by the homogeneous Richards equation. One of the reasons is that the solver, used by many other researchers earlier, is first order accurate in space and time. More important, at the wetting front the soil hydraulic conductivity changes dramatically due to the strong non-linearity of the soil water retention and hydraulic models. Solving the Richards equation accurately using large spatial size remains a major challenge. However, there are much larger uncertainties, e.g., the existence of macropores and the parameterization of soil water properties that dominate numerical errors. The Richards equation represents the major dynamics of the soils. Some parameter adjustment may ultimately compensate for the numerical error.

### 3.1.3. Pumping near an impervious wall

The classic Theis solution [60] was used to check the accuracy of the groundwater solver. This test problem shows that when an impervious wall is present, the drawdown in the confined aquifer will be greater because water supply is limited. The analytical solution to this problem can be obtained by modifying the classical Theis solution, considering the symmetry of the problem and setting a 'shadow' well in the mirror position of the pumping well. The time-series comparison with analytical solution at the observation well location is presented in Fig. 8. As we can see, the two solutions are in good agreement.

### 3.1.4. Coupled saturated/unsaturated model

We described a novel approach to couple the unsaturated/saturated models in an earlier section but its effectiveness in replacing the 3D equation must be tested. As analytical solutions to such a coupled system are not available, we can only compare our results to those obtained from an experiment, or solutions from a fully three-dimensional Richards model. Here we first present a comparison with the experimental data from Vauclin et al. [55]. This dataset was employed by many researchers to verify their 3D Richards equation codes, e.g. [61,62]. The experiment consists of a soil slab  $6.0$  by  $2.0$  meters. Initially the soil slab has established hydrostatic equilibrium with the water table at  $0.65$  meters from the bottom. At the center  $1$  m on the surface of the soil slab, a constant flux of  $3.5 \text{ m d}^{-1}$  is applied. On the left and right boundaries, water level is maintained at  $0.65$  m. As in other studies, we model only the right half of the domain due to the symmetrical nature of the

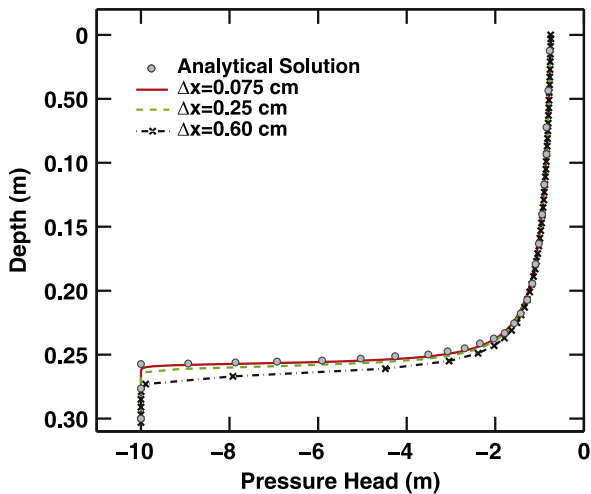


Fig. 7. Infiltration into very dry soil test problem at  $t = 6$  h. Circle: Analytical solution; solid line:  $\Delta x = 0.075$  cm; dashed line:  $\Delta x = 0.25$  cm; dashed line with cross:  $\Delta x = 0.6$  cm.

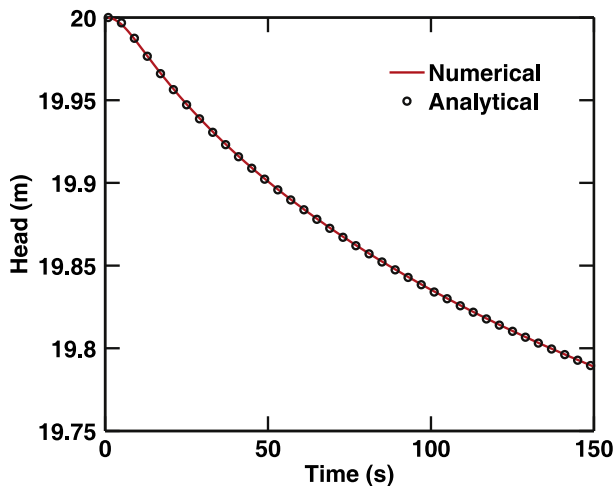


Fig. 8. Comparisons for the groundwater test problem for the case of pumping near an impervious wall. A definition sketch showing the geometry of the test problem is available in the Supplemental Information.

experiment. Thus, the no-flow boundary condition was applied to the left boundary. The soil retention and conductivity data published in the original paper in [55] was fitted to the Van Genuchten formulation by Dogan and Motz [61], Clement et al. [62] and provided here in Table 3. For easier discussion we name the soil columns that are directly receiving inflow as recharge columns and the other columns as passage columns.

The water table simulated using the proposed coupling method using a 121 ( $x$ ) by 21 ( $z$ ) grid is shown in Fig. 9. It is observed that the solutions at hours 3, 4 and 8 match well with the experimental data, whereas at hour 2 the model over-estimated the head at the left boundary. This over-estimation can be explained by the deviation of the coupled system from the actual 3D system. At the beginning stages of recharge, the passage columns are dry as compared to the recharge columns. Lateral moisture diffusion causes the inflow to be re-distributed to the passage columns before it reaches the unconfined aquifer and the unsaturated region stores a portion of the inflow. However, in the coupled system, water can move laterally only *after* it enters the unconfined aquifer. Thus at early stages the “recharge” is over-estimated. However as the water

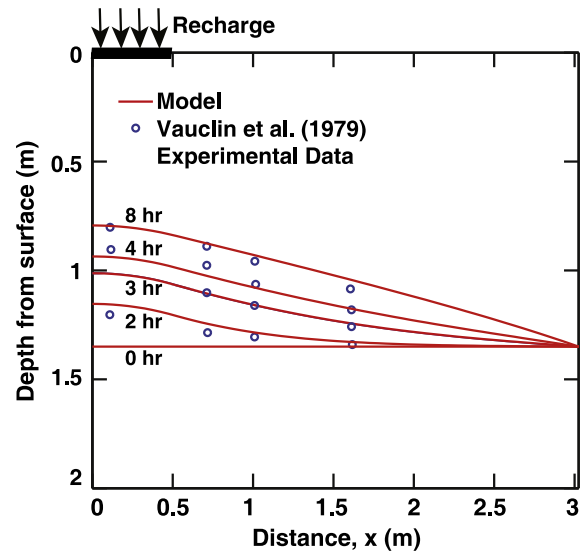


Fig. 9. Comparison of PAWS results with experimental results from [55].

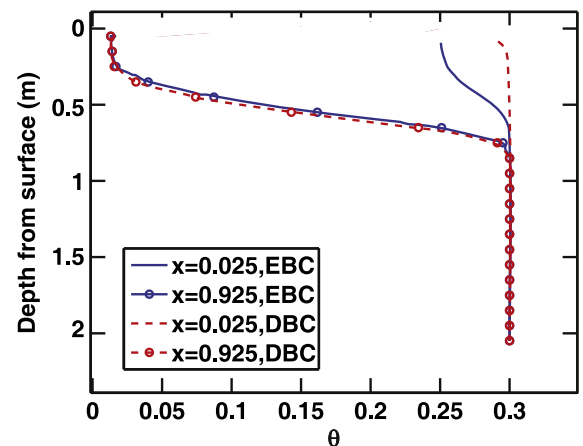


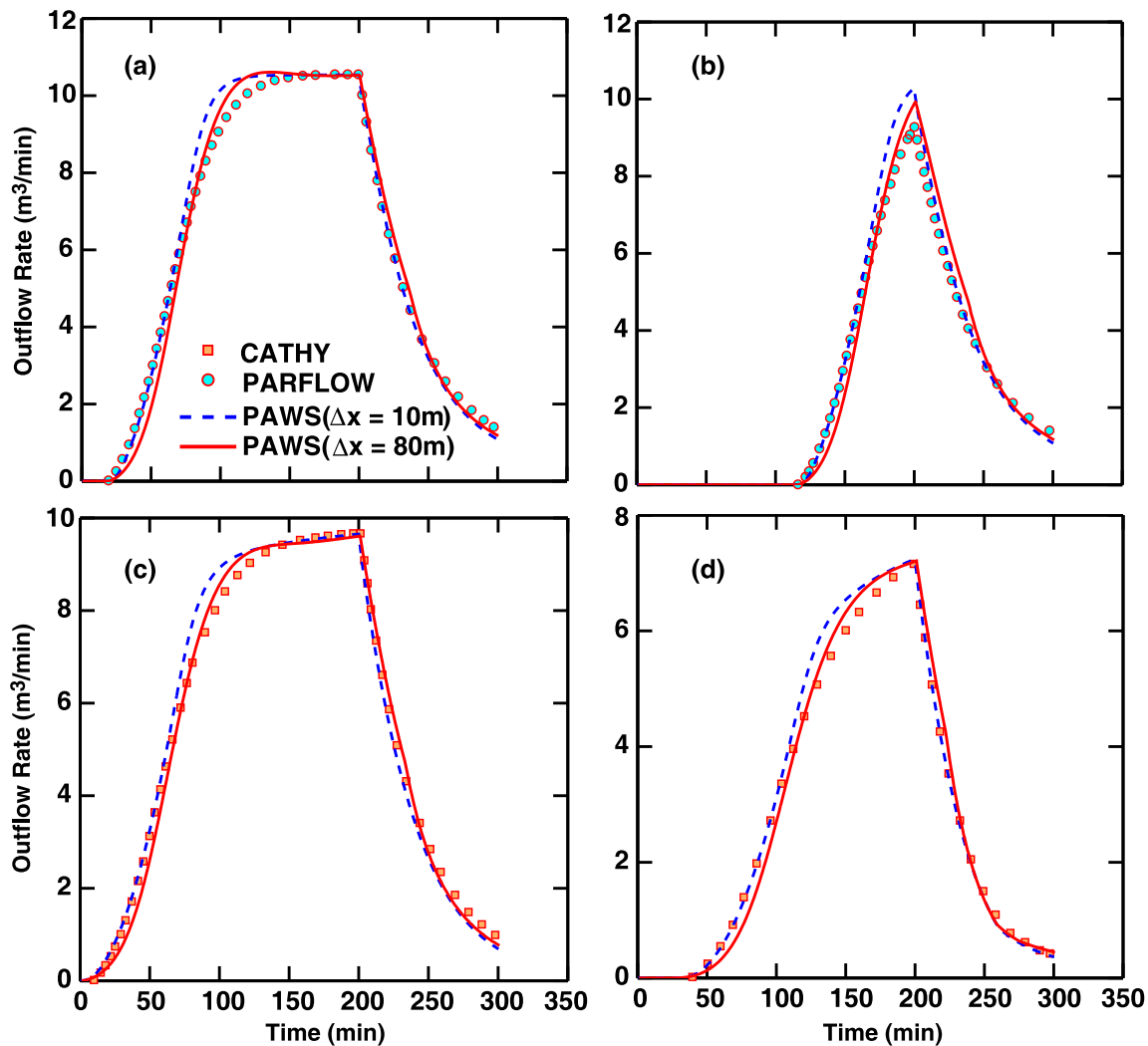
Fig. 10. Comparison of soil moisture profiles obtained using the approaches based on the proposed and traditional methods,  $\theta_s = 0.3$ , EBC = equation boundary condition (Present approach); DBC = Dirichlet boundary condition (Conventional approach).

table rises, the drainage term  $DR$  increases, the unconfined aquifer is able to transport more and more water, and the system gets closer and closer to steady state. The solution thus matches the experimental data better at later times. This diffusion effect is expected to be less noticeable when the method is applied to the watershed-scale, where the ratio of lateral flow flux to recharge may be much smaller compared to this experimental scenario.

Previous studies have directly used the location of the water table as a Dirichlet boundary condition to the Richards equation. An important implication of this approach is that, even at steady state, the saturated/unsaturated interface found in the soil profile must be higher than  $H$  because a head difference must exist in the soil profile to supply the required recharge to the saturated zone. This can lead to an erroneous soil moisture profile because the saturated/unsaturated interface should be placed right at the groundwater table.

Finally, Fig. 10 shows a comparison of the soil moisture profile at hour 8 between the proposed coupling method and the traditional method, in which the hydraulic head of the unconfined aquifer is provided as a Dirichlet boundary condition to the soil column. We notice that the soil moisture values in the recharge





**Fig. 11.** Comparison of the saturated/unsaturated flow coupling method in PAWS with solutions produced by two full 3D models, CATHY [19] and ParFlow [20,21] as reported in [63]. In cases (a) and (b), saturation excess was produced, slope =  $5 \times 10^{-4}$  and  $K_s = 1$  ( $\text{m d}^{-1}$ ). IWT was 0.5 m for case (a) and 1 m for case (b). Cases (c) and (d) correspond to infiltration excess with slope =  $5 \times 10^{-4}$  and IWT = 1 m.  $K_s$  is 0.01 ( $\text{m d}^{-1}$ ) for case (c) and 0.1 ( $\text{m d}^{-1}$ ) for case (d).

columns ( $x = 0.025$ ) based on the conventional method are much higher. Whereas in one of the passage columns ( $x = 0.925$ ) the soil moisture is lower than the steady state value. Both deviations are explained by the formulation of Eq. (40). With the present approach, the water table location (point at which soil moisture becomes saturated) agrees well with the head in the unconfined aquifer (in Fig. 7). Compared to the approach based on the Dirichlet boundary condition, the new coupling method is found to be very stable.

### 3.1.5. Coupled surface/subsurface flow

To further verify that the surface/subsurface flow coupling scheme gives reasonable solutions, our numerical results are compared to those reported in [63] using two models, CATHY [19] and ParFlow [20]. Both models solved the full 3D Richards equation. A hypothetical inclined plane with a fixed vertical soil thickness along the  $x$ -direction slope was used in the tests. There is no  $y$ -direction slope. Uniform precipitation falls on the plane and free outflow was allowed at the downstream edge. The plane is permeable so that water may infiltrate or exfiltrate. The sides and bottom of the soil column are no-flow boundaries. The initial condition consists of a phreatic groundwater table at a certain depth underneath the surface, and the unsaturated portion of the soil column

was assumed to be at hydrostatic state. The length of the slope is 400 m, the width ( $y$ -direction) is 320 m and the thickness of the soil column is 5 m. A vertical discretization of 0.05 m was used in our study, with two  $\Delta x$  values tested: 10 m and 80 m. The slope of the plane, the saturated hydraulic conductivity ( $K_s$ ) and the initial depth of the water table (IWT) were varied to allow different flow conditions. Complete descriptions of the numerical experiments are available in [63]. In PAWS, surface flow is simulated by the 2D diffusive wave equation; subsurface is modeled by 1D Richards equation, 2D groundwater flow equation and the coupling method described in Section 2.4.1; excess water at the end of each subsurface step becomes available for overland flow.

We have simulated four sets of test cases reported in [63], namely, two saturation excess cases with different IWT and two infiltration excess cases with different conductivities. For all these cases, a  $S_{nz}$  value of  $5 \times 10^{-4}$  was used. The outflow rates simulated by PAWS in comparison with results from CATHY and ParFlow models using identical spatial and temporal resolutions are shown in Fig. 11 for these four cases. PAWS predicts slightly higher flow rates and sharper rising limbs than the 3D models. Two reasons could have accounted for this difference: (a) the 2D overland solver of PAWS is a second accurate scheme whereas the two 3D models both employ first order accurate schemes (b) PAWS does not

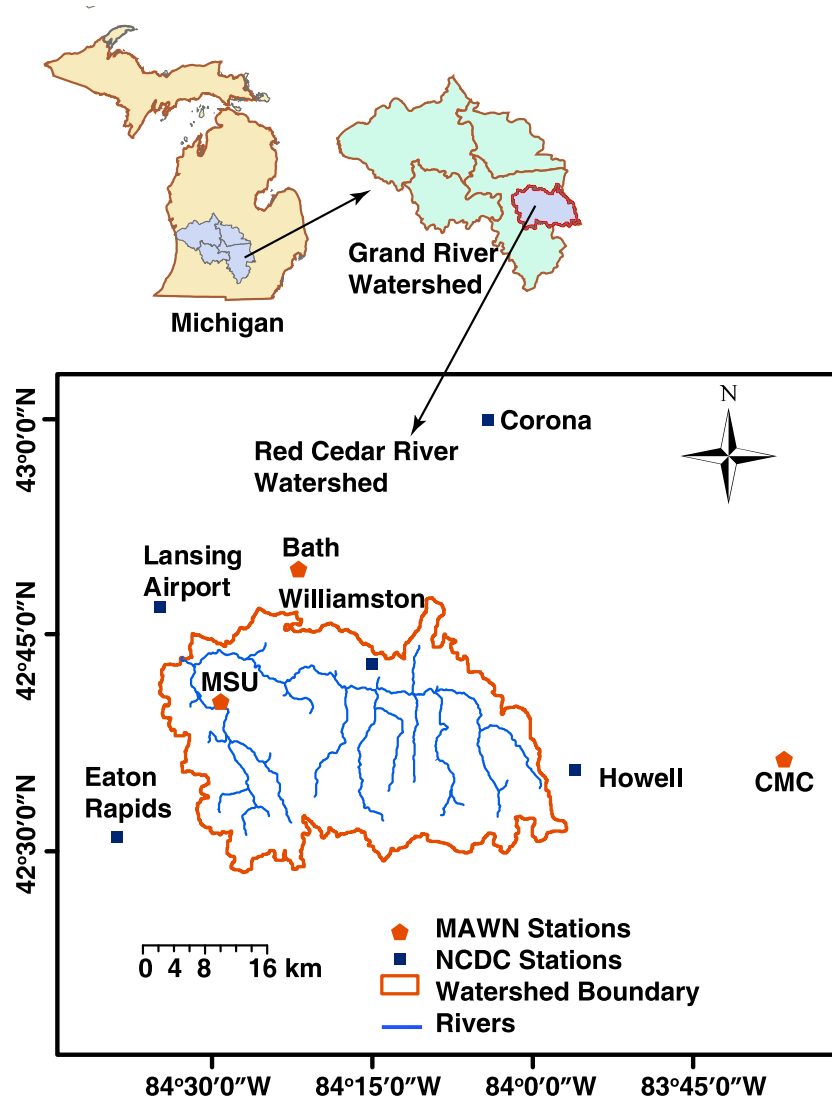


Fig. 12. Map showing the location of the Red Cedar River Watershed and the weather stations.

simulate lateral soil water movement above the water table, therefore more water exited as surface flow. In spite of these differences, we observe that PAWS produces results that are comparable with the two full 3D models in all four cases. This indicates that the surface–subsurface interaction and the simplified subsurface flow formulation of PAWS indeed approximate the more complex 3D equation reasonably well.

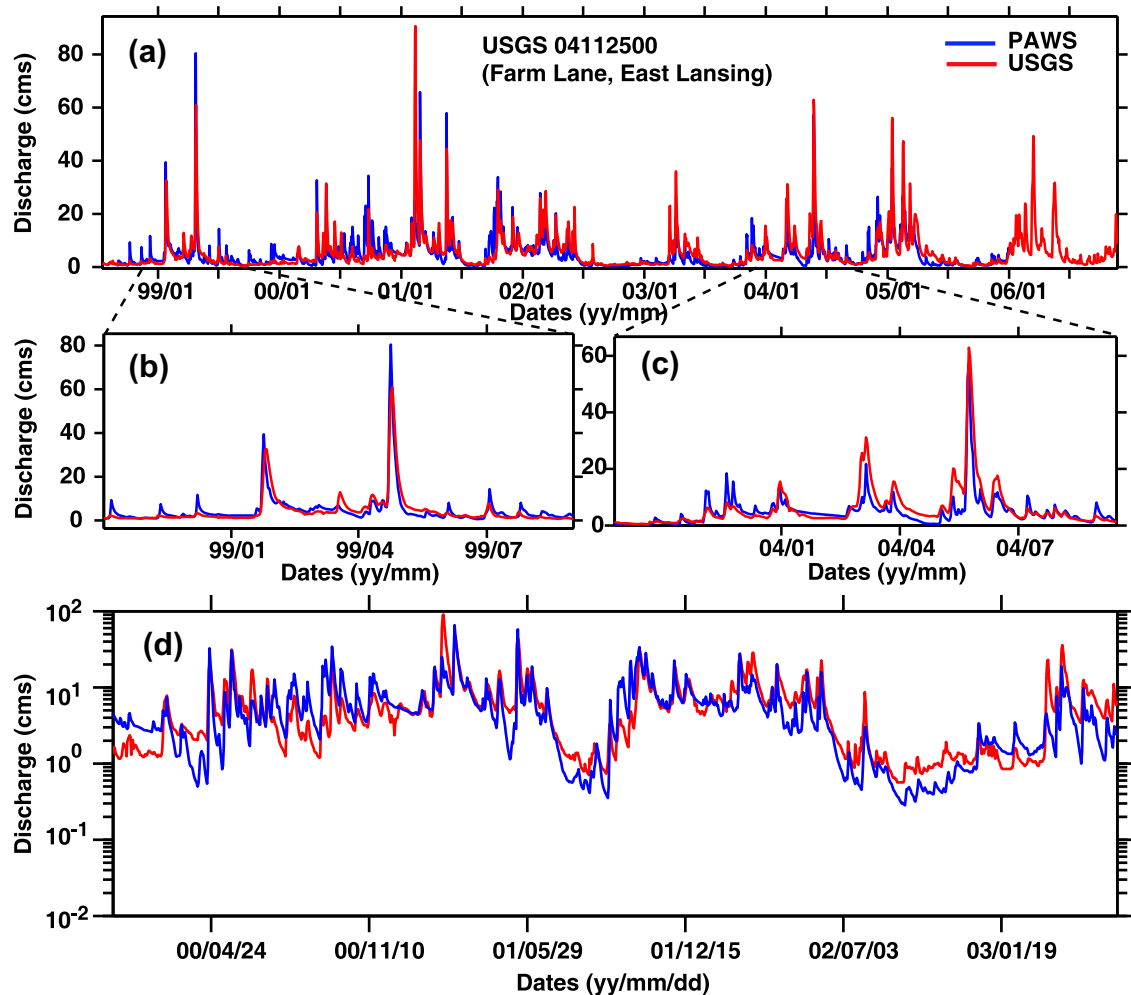
### 3.2. Application of the model to a Michigan watershed

#### 3.2.1. Study site and input data

In this section, we apply the PAWS model to a medium-sized watershed in Michigan and evaluate its performance as an integrated watershed model. The Red Cedar River watershed is located in the Grand River watershed in Michigan (Fig. 12). The total area of the watershed is 1169 km<sup>2</sup>. The watershed has a relatively low relief with the maximum elevation recorded as 324 m and a minimum of 249 m. The watershed is characterized by the humid continental climate with ample precipitation and distinct temperatures in different seasons. The land use of the Red Cedar River watershed is predominantly agricultural. Row crops plus forage crops add up to 56% of the watershed. Urban areas are in the

northwest while forested areas are scattered throughout the watershed.

A 900 m × 900 m grid was used for the computations. Daily or sub-daily weather data were obtained from various sources including the National Climatic Data Center [64] and the Michigan Automated Weather Network (MAWN) [65]. The locations of the stations are shown in Fig. 12. A suite of subroutines were developed to process data from readily available national databases. The National Elevation Dataset (30 m resolution NED) was used to establish elevation of the land grid. The 30 m resolution IFMAP land use land cover data [66] which was derived from classification of Landsat Thematic Mapper™ imagery was used to provide land use information. SSURGO data were used to provide initial values for the soil properties. Eight major rivers from the National Hydrography Dataset (NHD) were manually selected as the modeled river network. River bed elevation was obtained by subtracting the bankfull river depth from the local elevation values (taken as the bank elevation). It was found that a 90 m resolution DEM often had difficulty capturing thin features such as streams and creeks. Bankfull river depths were estimated using an empirical relation from [67] that relates depth to river width and slope. Field measured data were used to fit the equation. Examination of the river bed elevations estimated based on the DEM and NED



**Fig. 13.** Comparison of the observed and simulated daily flow at USGS gage 04112500 at East Lansing. X-axis label is dates in YY/MM/DD format (a) Hydrograph comparison in linear scale for the entire simulation period (b and c) Close-up views for selected simulation periods from (a) and (d) Hydrograph comparison in log-scale.

**Table 4**

Performance metrics evaluating the model applied to the RCR watershed.

Periods (YYYYMMDD)	NASH	RMSE	RNASH	ME	$R^2$
Calibration (20030101 ~ 20051231)	0.694	3.855	0.693	−1.122	0.742
Verification (19990101 ~ 20021231)	0.585	4.816	0.661	−0.232	0.591
Verification (excluding 20010208 to 20010222)	0.699	3.415	0.685	−0.084	0.726
Williamston (20030101 ~ 20051231)	0.665	1.936	0.677	−0.545	0.714

using the above approach for the Red Cedar River, as well as the groundwater table provided by well head observations from the Michigan Wellogic database [68,69] indicated that the NED-based estimates closely correspond to the groundwater heads while the DEM based values generally tend to be much higher, especially in reaches with a large drop in elevation. Close examination of the discrepancies revealed that the places where the two estimates differed the most are often places where the river feature is missing from the DEM map due to its resolution. Therefore, we conclude that NED-based values are much more accurate.

Generally, some information about the aquifer configuration is necessary for constructing the conceptual groundwater model. In the case of the RCR watershed, a layer of relatively permeable glacial deposit with a thickness ranging approximately from 5 meters to 30 meters overlies the bedrock. This information allows us to

divide the groundwater model into two layers. The upper layer is the glacial deposit, taken as the unconfined aquifer; the lower layer is the bedrock, taken as the confined aquifer. For the RCR watershed and Michigan in general, valuable groundwater-related data have been collected, logged and filed with the Michigan Department of Environmental Quality (MDEQ) when water wells in the state of Michigan were drilled. These records were used to estimate local hydraulic conductivities, groundwater heads and thicknesses of the glacial drift layer [69]. Spatial fields were generated by interpolating these well records using kriging [70] after noises were removed.

### 3.2.2. Model calibration

Flow records at the USGS gage 04112500, East Lansing from 2002/09/01 to 2005/12/31 were used to calibrate the model while

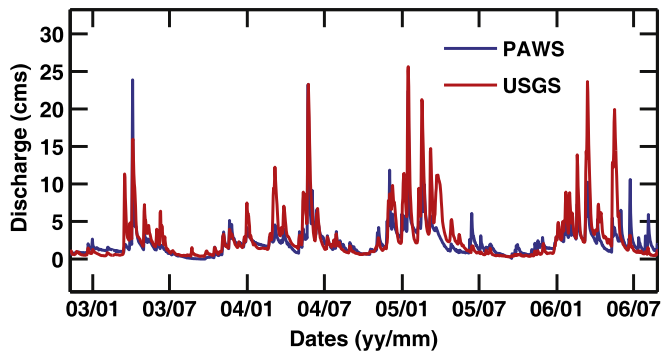


Fig. 14. Comparison of the observed and simulated daily flow at USGS gage 04111379 at Williamston.

the years prior to this period were used for model verification. The reason for using water years 2002 to 2005 was that in these three consecutive years the watershed experienced high, normal and unusually low precipitation periods. Parameters estimated during calibration are listed in Table S1 in the Supplementary Material. The parameter  $S_{nz}$  was set to  $5 \times 10^{-4}$ . The average computer time for 10 years of simulation for the RCR basin was 0.85 h on a PC with a 3.0 GHz Intel Xeon processor and 4 GB of RAM. Auto-calibration of the model was done using the Differential Evolution (DE) algorithm [71] with the computational resources at the High Performance Computing Center (HPCC) at Michigan State University. The original DE code was modified and implemented for the parallel computing environment.

### 3.2.3. Hydrograph comparison

Since one of the main objectives of this paper is to describe the mathematical basis of PAWS, here we report the initial evaluation of the model as it is applied to a real watershed and focus primarily on streamflow comparisons. Future publications will provide detailed comparisons with spatial and temporal groundwater data, as well as soil moisture and snow cover observations. Model performance is evaluated against USGS streamflow data. The metrics we use to measure the performance include the coefficient of determination ( $R^2$ ), root mean squared error (RMSE), the Nash–Sutcliffe model efficiency coefficient (NASH), the mean error (ME) and the square-root transformed NASH (RNASH). RNASH is defined as:

$$RNASH = 1 - \frac{\sum_{t=1}^T (\sqrt{Q_o^t} - \sqrt{Q_m^t})^2}{\sum_{t=1}^T (\sqrt{Q_o^t} - \sqrt{Q_o})^2} \quad (59)$$

The daily observed ( $Q_o$ ) and simulated flows ( $Q_m$ ) ( $m^3 d^{-1}$ ) at the USGS gage 04112500 (Farm Lane, East Lansing) for the 10 year period from 1998 to 2007 are compared in Fig. 13. The new model did a good job of predicting streamflows, especially during the baseflow/low flow periods (Fig. 13d shows a semi-log plot of the hydrograph comparison). The performance metrics of the calibration, verification and entire simulation periods are summarized in Table 4.

Close inspection of the hydrograph indicated that some mismatch is mostly due to the underestimation of streamflow peaks during snowmelt periods. Because the year 2003 was a dry year, the underestimation of spring snowmelt peak has been quite serious and is the main reason for NASH reduction during the calibration period. On the other hand, the model failed to describe a snowmelt event in spring 2001 which is almost solely responsible for the significant reduction of NASH in the verification period. Excluding this single event brings the NASH of the verification

period back to 0.699 (Table 4). Several reasons may have contributed to the mismatch. First, some uncertainty exists with the measurement methods of snowfall depths at MAWN stations which may have resulted in under-estimation of snowfall. This is most likely since the underestimation of snowmelt events extends into their recession periods, showing an underestimation of total mass. Thus it is less likely to be due to errors associated with the separation of runoff and infiltration. Secondly, some parameter adjustment within the UEB snowpack module may be needed. However, since the UEB model is intended as a model with no parameters for adjustment, adjusting this aspect of the simulation is difficult. Further in-depth research is needed to understand the exact reason for this under-estimation and to improve the results. Fig. 14 shows the hydrograph comparison at the inner gage Williamston. The model performs equally well at this gage. Additional comparisons and more detailed modeling will be reported in future papers.

## 4. Limitations/Future research

It is an assumption in our model that soil moisture does not flow laterally in the unsaturated part of the soil column. The model performance may deteriorate in catchments dominated by perched water table dynamics. Detailed analyses including comparisons with the SWAT model indicated that soil lateral flow is a minor component of the hydrologic cycle for the watershed described in this paper [56], however it may be important for other watersheds and the process will be added in a future version of the model. In addition, soil macropore flow will be considered. Macropore (or bypass flow) can be significant for some types of soils. A simple vegetation module is used in the current version of the model and a more comprehensive module will be developed in future versions to better model the interactions between vegetation and the hydrologic cycle. The model will also be applied to larger watersheds and watersheds with different land use/ geology/soil/climate. The effects of spatial resolution on the results will be examined in future work. Previous studies have shown that process-based models are marginally sensitive to grid sizes [30]. Finally, long term and short term transport modules will be added to the model.

## Acknowledgements

This research was funded by the NOAA Center of Excellence for Great Lakes and Human Health. We thank Jie Niu for assistance with data processing and GUI development. Computer time on the High Performance Computer Center at MSU and help from HPCC staff Dirk Colbry and Andrew Keen are gratefully acknowledged.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.advwatres.2010.09.002](https://doi.org/10.1016/j.advwatres.2010.09.002).

## References

- [1] Beven K. Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrol Process* 2002;16(2):189–206.
- [2] U.S.EPA. National Primary Drinking Water Regulations: Ground Water Rule; Proposed Rules. 40 CFR Parts 141 and 142. 2000. p. 30194–274.
- [3] Thupaki P, Phanikumar MS, Beletsky D, Schwab DJ, Nevers MB, Whitman RL. Budget analysis of *Escherichia coli* at a Southern Lake Michigan Beach. *Environ Sci Technol* 2010;44(3):1010–6. [doi:10.1021/es902232a](https://doi.org/10.1021/es902232a).
- [4] Grant SB, Kim JH, Jones BH, Jenkins SA, Wasyl J, Cudaback C. Surf zone entrainment, along-shore transport, and human health implications of pollution from tidal outlets. *J Geophys Res-Oceans* 2005;110(C10):C10025. [doi:10.1029/2004ic002401](https://doi.org/10.1029/2004ic002401).



- [5] Neitsch SL, Arnold JG, Kiniry JR, Williams JR. Soil and water assessment tool theoretical documentation version 2005. Temple, Texas: US Department of Agriculture – Agricultural Research Service; 2005.
- [6] Liang X, Xie ZH. A new surface runoff parameterization with subgrid-scale soil heterogeneity for land surface models. *Adv Water Resour* 2001;24(9–10):1173–93.
- [7] Chen F, Dudhia J. Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon Weather Rev* 2001;129(4):569–85.
- [8] Miguez-Macho G, Fan Y, Weaver CP, Walko R, Robock A. Incorporating water table dynamics in climate modeling: 2. Formulation, validation, and soil moisture simulation. *J Geophys Res-Atmos* 2007;112(D13):D13108. doi:10.1029/2006jd008112.
- [9] Jiang XY, Niu GY, Yang ZL. Impacts of vegetation and groundwater dynamics on warm season precipitation over the Central United States. *J Geophys Res-Atmos* 2009;114:D06109. doi:10.1029/2008jd010756.
- [10] Maxwell RM, Miller NL. Development of a coupled land surface and groundwater model. *J Hydrometeorol* 2005;6(3):233–47.
- [11] Gulden LE, Rosero E, Yang ZL, Rodell M, Jackson CS, Niu GY, et al. Improving land-surface model hydrology: is an explicit aquifer model better than a deeper soil profile? *Geophys Res Lett* 2007;34(9):L09402. doi:10.1029/2007gl029804.
- [12] Maxwell RM, Kollet SJ. Interdependence of groundwater dynamics and land-energy feedbacks under climate change. *Nature Geosci* 2008;1(10):665–9. doi:10.1038/Ngeo315.
- [13] Entekhabi D, Rodriguez-Iturbe I, Castelli F. Mutual interaction of soil moisture state and atmospheric processes. *J Hydrol* 1996;184(1–2):3–17.
- [14] Anyah RO, Weaver CP, Miguez-Macho G, Fan Y, Robock A. Incorporating water table dynamics in climate modeling: 3. Simulated groundwater influence on coupled land-atmosphere variability. *J Geophys Res-Atmos* 2008;113(D7):D07103. doi:10.1029/2007jd009087.
- [15] Richards LA. Capillary conduction of liquids in porous mediums. *Physics* 1931;1:318–33.
- [16] Downer CW, Ogden FL, Martin WD, Harmon RS. Theory, development, and applicability of the surface water hydrologic model CASC2D. *Hydrol Process* 2002;16(2):255–75.
- [17] Vanderkwaak JE, Loague K. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resour Res* 2001;37(4):999–1013.
- [18] Goderniaux P, Brouyere S, Fowler HJ, Blenkinsop S, Therrien R, Orban P, et al. Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves. *J Hydrol* 2009;373(1–2):122–38. doi:10.1016/j.jhydrol.2009.04.017.
- [19] Camporese M, Paniconi C, Putti M, Orlandini S. Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data. *Water Resour Res* 2010;46:W02512. doi:10.1029/2008wr007536.
- [20] Kollet SJ, Maxwell RM. Integrated surface-groundwater flow modeling: a free-surface overland flow boundary condition in a parallel groundwater flow model. *Adv Water Resour* 2006;29(7):945–58. doi:10.1016/j.advwatres.2005.08.006.
- [21] Maxwell RM, Kollet SJ. Quantifying the effects of three-dimensional subsurface heterogeneity on Hortonian runoff processes using a coupled numerical, stochastic approach. *Adv Water Resour* 2008;31(5):807–17. doi:10.1016/j.advwatres.2008.01.020.
- [22] Yeh GT, Huang GB, Cheng HP, Zhang F, Lin HC, Edris E, et al. A first-principle, physics-based watershed model: WASH123D. In: Singh VP, Frevert DK, editors. *Watershed models*. Boca Raton, FL: CRC Press; 2006. p. 211–44.
- [23] Simunek J, Van Genuchten MT, Sejna M. The HYDRUS software package for simulating two- and three-dimensional movement of water, heat, and multiple solutes in variably-saturated media, version 1.0, technical manual, PC Progress. Prague, Czech Republic, 2006.
- [24] Panday S, Huyakorn PS. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Adv Water Resour* 2004;27(4):361–82.
- [25] Downer CW, Ogden FL. Appropriate vertical discretization of Richards' equation for two-dimensional watershed-scale modelling. *Hydrol Process* 2004;18(1):1–22.
- [26] Weill S, Mouche E, Patin J. A generalized Richards equation for surface/subsurface flow modelling. *J Hydrol* 2009;366(1–4):9–20. doi:10.1016/j.jhydrol.2008.12.007.
- [27] Downer CW, Ogden FL. GSSHA: model to simulate diverse stream flow producing processes. *J Hydrol Eng* 2004;9(3):161–74.
- [28] McMichael CE, Hope AS, Loaiciga HA. Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. *J Hydrol* 2006;317(3–4):307–24. doi:10.1016/j.jhydrol.2005.05.023.
- [29] Thompson JR, Sorenson HR, Gavin H, Refsgaard A. Application of the coupled MIKE SHE/MIKE 11 modelling system to a lowland wet grassland in southeast England. *J Hydrol* 2004;293(1–4):151–79. doi:10.1016/j.jhydrol.2004.01.017.
- [30] Vazquez RF, Feyen L, Feyen J, Refsgaard JC. Effect of grid size on effective parameters and model performance of the MIKE-SHE code. *Hydrol Process* 2002;16(2):355–72.
- [31] Jia YW, Ni GH, Kawahara Y, Suetsugi T. Development of WEP model and its application to an urban watershed. *Hydrol Process* 2001;15(11):2175–94.
- [32] Breuer L, Eckhardt K, Frede HG. Plant parameter values for models in temperate climates. *Ecol Model* 2003;169(2–3):237–93. doi:10.1016/S0304-3800(03)00274-6.
- [33] Allen RG, Pereira LS, Raes D, Smith M. Crop evaporation. Guidelines for computing crop water requirement–FAO Irrigation and Drainage paper 56. Rome: Food and Agriculture Organization of the United Nations; 1998.
- [34] Noilhan J, Planton S. A simple parameterization of land surface processes for meteorological models. *Mon Weather Rev* 1989;117(3):536–49.
- [35] Lai CT, Katul G. The dynamic role of root-water uptake in coupling potential to actual transpiration. *Adv Water Resour* 2000;23(4):427–39.
- [36] Braud I, Varado N, Olliso A. Comparison of root water uptake modules using either the surface energy balance or potential transpiration. *J Hydrol* 2005;301(1–4):267–86.
- [37] Dingman SL. Physical hydrology. Prentice Hall; 2002.
- [38] Singh VP. Kinematic wave modeling in water resources: surface water hydrology. Wiley-Interscience; 1996.
- [39] Gottardi G, Venutelli M. An accurate time integration method for simplified overland flow models. *Adv Water Resour* 2008;31(1):173–80. doi:10.1016/j.advwatres.2007.08.004.
- [40] Casulli V. Semi-implicit finite-difference methods for the 2-dimensional shallow-water equations. *J Comput Phys* 1990;86(1):56–74.
- [41] van Dam JC, Feddes RA. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. *J Hydrol* 2000;233(1–4):72–85.
- [42] Celia MA, Bouloutas ET, Zarba RL. A general mass-conservative numerical-solution for the unsaturated flow equation. *Water Resour Res* 1990;26(7):1483–96.
- [43] van Genuchten MT. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci Soc Amer J* 1980;44(5):892–8.
- [44] Press WH, Flannery BP, Teukolsky SA, Vetterling WT. Numerical recipes in FORTRAN 77: the art of scientific computing. Cambridge University Press; 1997.
- [45] Jia YW. Modeling infiltration into a multi-layered soil during an unsteady rain. *J Hydrosol Hydraulic Eng, JSCE (Japan)* 1998;16(2).
- [46] Neuman SP. Wetting front pressure head in infiltration-model of Green and Ampt. *Water Resour Res* 1976;12(3):564–6.
- [47] Luce CH, Tarboton DG. The application of depletion curves for parameterization of subgrid variability of snow. *Hydrol Process* 2004;18(8):1409–22.
- [48] Luce CH, Tarboton DG, Cooley RR. The influence of the spatial distribution of snow on basin-averaged snowmelt. *Hydrol Process* 1998;12(10–11):1671–83.
- [49] Luce CH, Tarboton DG, Cooley KR. Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model. *Hydrol Process* 1999;13(12–13):1921–33.
- [50] Twarakavi NKC, Simunek J, Seo S. Evaluating interactions between groundwater and vadose zone using the HYDRUS-based flow package for MODFLOW. *Vadose Zone J* 2008;7(2):757–68.
- [51] Stoppelenburg FJ, Kavar K, Pastoors MJH, Tiktak A. Modeling the interactions between saturated and unsaturated groundwater. Off-line coupling of LGM and SWAP, RIVM Rep. 500026001/2005. RIVM. Bilthoven, The Netherlands; 2005.
- [52] van Dam JC, Groenendijk P, Hendriks RFA, Kroes JG. Advances of modeling water flow in variably saturated soils with SWAP. *Vadose Zone J* 2008;7(2):640–53.
- [53] Gunduz O, Aral MM. River networks and groundwater flow: a simultaneous solution of a coupled system. *J Hydrol* 2005;301(1–4):216–34.
- [54] Freeze RA, Harlan RL. Blueprint for a physically-based, digitally-simulated hydrologic response model. *J Hydrol* 1969;9:237–58.
- [55] Vauclin M, Khanji D, Vachaud G. Experimental and numerical study of a transient, 2-dimensional unsaturated-saturated water table recharge problem. *Water Resour Res* 1979;15(5):1089–101.
- [56] Shen C. A process-based distributed hydrologic model and its application to a Michigan watershed. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Michigan State University, East Lansing; 2009.
- [57] DiGiannmarco P, Todini E, Lamberti P. A conservative finite elements approach to overland flow: the control volume finite element formulation. *J Hydrol* 1996;175(1–4):267–91.
- [58] Haverkamp R, Vauclin M, Touma J, Wierenga PJ, Vachaud G. Comparison of numerical-simulation models for one-dimensional infiltration. *Soil Sci Soc Amer J* 1977;41(2):285–94.
- [59] Lehmann F, Ackerer PH. Comparison of iterative methods for improved solutions of the fluid flow equation in partially saturated porous media. *Transp Porous Media* 1998;31(3):275–92.
- [60] Freeze AR, Cherry JA. Groundwater. Prentice Hall; 1979.
- [61] Dogan A, Motz LH. Saturated-unsaturated 3D groundwater model. II: verification and application. *J Hydrol Eng* 2005;10(6):505–15. doi:10.1061/(ASCE)1084-0699(2005)10:6(505).
- [62] Clement TP, Wise WR, Molz FJ. A physically-based, 2-dimensional, finite-difference algorithm for modeling variably saturated flow. *J Hydrol* 1994;161(1–4):71–90.
- [63] Sulis M, Meyerhoff SB, Paniconi C, Maxwell RM, Putti M, Kollet SJ. A comparison of two physics-based numerical models for simulating surface water-groundwater interactions. *Adv Water Resour* 2010;33(4):456–67. doi:10.1016/j.advwatres.2010.01.010.

- [64] NCDC. National Climatic Data Center. Available at <<http://www.ncdc.noaa.gov/oa/climate/climatedata.html#daily>>; 2010 [accessed 10.08.2010].
- [65] MAWN. Michigan Automated Weather Network, East Lansing, Michigan, 48823. Available at <<http://www.agweather.geo.msu.edu/mawn/>>; 2010 [accessed 10.08.2010].
- [66] MDNR. 2001 IFMAP/GAP Lower Peninsula Land Cover. Available at <<http://www.mcgi.state.mi.us/mgdl/?rel=thext&action=thmname&cid=5&cat=Land+Cover+2001>>; 2010 [accessed 28.11.2009].
- [67] Bjerklie DM. Estimating the bankfull velocity and discharge for rivers using remotely sensed river morphology information. *J Hydrol* 2007;341(3–4): 144–55. doi:[10.1016/j.jhydrol.2007.04.011](https://doi.org/10.1016/j.jhydrol.2007.04.011).
- [68] GWIM. State of Michigan Public Act 148 Groundwater Inventory and Mapping Project (GWIM). Technical Report. 2006.
- [69] Simard A. Predicting groundwater flow and transport using Michigan's statewide wellogic database. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Michigan State University, East Lansing; 2007.
- [70] Li SG, Liu Q, Afshari S. An object-oriented hierarchical patch dynamics paradigm (HPDP) for modeling complex groundwater systems across multiple-scales. *Environ Model Software* 2006;21(5):744–9. doi:[10.1016/j.envsoft.2005.11.001](https://doi.org/10.1016/j.envsoft.2005.11.001).
- [71] Price K, Storn R, Lampinen J. Differential evolution: a practical approach to global optimization. Berlin: Springer-Verlag; 2005.