Water Resources Research

RESEARCH ARTICLE
10.1029/2022WR032198

A Generalized Multistep Dynamic (GMD) TOPMODEL

Salim Goudarzi[1], **David Milledge**[1], **and Joseph Holden**[2]

[1]Department of Civil Engineering, Newcastle University, Newcastle, UK, [2]School of Geography, University of Leeds, Leeds, UK

Special Section:
Advancing process representation in hydrologic models: Integrating new concepts, knowledge, and data

Key Points:

- Hydrological models should be formulated in a constraint-handling Ordinary Differential Equation form where possible
- Similarity-based models cannot constrain spatially-distributed information. "Iso-basin" discretization helps overcome this limitation
- Arbitrary Topographic Index (TI) discretization can significantly alter model parameters. More objective TI classifications are needed

Correspondence to:
S. Goudarzi,
salim_goodarzi@yahoo.com

Citation:
Goudarzi, S., Milledge, D., & Holden, J. (2023). A Generalized Multistep Dynamic (GMD) TOPMODEL. *Water Resources Research*, *59*, e2022WR032198. https://doi.org/10.1029/2022WR032198

Received 14 FEB 2022
Accepted 3 JAN 2023

Author Contributions:
Conceptualization: Salim Goudarzi, David Milledge
Formal analysis: Salim Goudarzi, Joseph Holden
Investigation: Salim Goudarzi, Joseph Holden
Methodology: Salim Goudarzi, David Milledge
Resources: Salim Goudarzi
Software: Salim Goudarzi
Supervision: Salim Goudarzi, David Milledge, Joseph Holden
Validation: Salim Goudarzi
Visualization: Salim Goudarzi

Abstract There is a lack of Ordinary Differential Equation (ODE) formulations in numerical hydrology, contributing to the lack of application of canned adaptive timestep solvers; hence the continued dominance of fixed (e.g., Euler) timestep techniques despite their fundamental problems. In this paper, we reformulate Dynamic-TOPMODEL into a constraint-handling ODE form and use MATLAB's advanced adaptive ODE-solvers to solve the resulting system of equations. For wider applicability, but based on existing research and/or first principles, we developed Generalized Multistep Dynamic TOPMODEL which includes: iso-basin spatial discretization, diffusion wave routing, depth-dependent overland flow velocity, relaxing the assumption of water-table parallelism to the ground surface, a power-law hydraulic conductivity profile, new unsaturated zone flux, and a reference frame adjustment. To demonstrate the model we calibrate it to a peat catchment case study, for which we also test sensitivity to spatial discretization. Our results suggest that (a) a five-fold improvement in model runtime can result from adaptive timestepping; (b) the additional iso-basin discretization layer, as a way to further constrain spatial information where needed, also improves performance; and (c) the common-practice arbitrary Topographic Index (TI) discretization substantially alters calibrated parameters. More objective and physically constrained (e.g., top-down) approaches to TI classification may be needed.

1. Introduction

TOPMODEL (Beven & Kirkby, 1979) continues to be one of the most widely applied hydrological models in both research and practice (Beven et al., 2021). It assumes that, independent of their specific spatial locations, some parts of the catchment will manifest similar tendencies to become, and remain, saturated during a storm. Based on a catchment-specific Topographic Index (TI) of similarity, these areas are grouped together to form Hydrologically Similar Units (HSUs), for which the numerical computations are performed when estimating a catchment's runoff response to rainfall. This grouping has enabled the dramatic reduction in runtimes associated with TOPMODEL, compared to fully spatially distributed models such as those of Freeze and Harlan (1969), Loague (2010), and Gao et al. (2015).

Though fast, TOPMODEL is a lumped model in that only a single value of catchment average subsurface storage is updated in each timestep, from which individual HSU subsurface storage values are back-calculated according to the deviation of their TI value from the catchment average TI value (Beven, 2011, p. 211); hence, it can be construed as a quasi single-store (equally a pseudo multi-store) model. Implicit in such application of a time-invariant TI is the assumption (a) that the transients of water-table between HSUs across the catchment are fast enough, that in each timestep water table can be approximated by a steady-state configuration. At the same time the (quasi) single-store representation leads to assumption (b) that during rainfall there is always downslope flow at each and every point in the catchment, equal to the recharge rate from all of the upslope areas draining to that point (i.e., there is always downslope connectivity everywhere, all the way to the stream network) (Beven et al., 2021). Both assumptions (a) and (b) are clearly approximations.

Dynamic-TOPMODEL of Beven and Freer (2001) relaxed both of the above assumptions by allowing subsurface storage of individual HSUs to vary locally and independently of both the catchment average storage and TI, by incorporating a time-dependent kinematic wave solution to the subsurface flow. However, since its introduction 20 years ago, and despite significantly improving catchment representation, the original steady-state version has remained the preferred choice (albeit sometimes with modifications/improvements) (Arenas-Bautista et al., 2018; Fu et al., 2018; Gil & Tobón, 2016; Jeziorska & Niedzielski, 2018; J. Wang et al., 2020; Lane & Milledge, 2013; Li et al., 2019; Mukae et al., 2018; Park et al., 2019; Rogelis et al., 2016; Xue et al., 2018; Zhang et al., 2016). Aside from the considerable additional complexity in numerical implementation of the dynamic

Writing – original draft: Salim Goudarzi
Writing – review & editing: Salim Goudarzi, David Milledge, Joseph Holden

versus the steady-state version, the lack of momentum in transitioning is most likely due to the substantially slower runtimes, making the dynamic version much less attractive for calibration purposes.

Although, naturally, longer runtimes are to some extent unavoidable due to additional CPU cost of numerically solving a more complex system, in Dynamic-TOPMODEL's case they are also partly due to the model setup requiring (at least in principle) thousands of HSUs at catchment scale. More specifically, the kinematic wave solution adapted by Dynamic-TOPMODEL requires HSUs to be "strictly" ordered downslope, because kinematic waves are unidirectional (have one characteristic direction) downslope, as opposed to for example, dynamic or diffusion waves which can travel upslope and downslope, hence better mimicking natural waves (see Lighthill & Whitham, 1955 for more information). Consequently, all Digital Elevation Model (DEM) cells in HSU#1 must be upslope of cells in HSU#2, and #2 upslope of #3, and so on. This poses a severe limitation on the way DEM cells can be grouped into HSUs while maintaining their "strict" downslope ordering status; resulting in the need for many HSUs at catchment scale. Few authors report numbers of HSUs in Dynamic-TOPMODEL applications. Among those who reported their numbers of HSUs, the downslope requirement is clearly only "approximately" enforced (e.g., Page et al. (2007) and Freer et al. (2004), 27 and 41 HSUs, respectively). Possibly to reduce runtimes, but these studies do not report any test (such as sensitivity analysis) or discussion of the potential repercussions of such approximation.

In an attempt to improve runtimes while allowing a more flexible discretization, Metcalfe et al. (2015) proposed a flow distribution matrix (FDM) allowing for interactions between HSUs, where interactions are based on Quinn et al.'s (1991) multiple-direction downslope surface flow paths between DEM cells. The FDM essentially enforces a "strictly" downslope flow for any, and all HSU discretisations, such that fewer numbers of HSUs can be used without violating the subsurface kinematic wave's requirement. Surprisingly, however, while the FDM approach reduced the number of HSUs and subsequently lead to a significantly more spatially flexible discretization, it did not improve model runtimes, rendering their Dynamic-TOPMODEL "currently inadequate" (Metcalfe et al., 2015) for calibration purposes or evaluation over longer time scales. This suggests that the additional matrix operations required in each timestep to account for HSU interactions (i.e., using FDM) counters the computational savings due to reduced numbers of HSUs.

We argue that this, at least in part, is because HSU interactions are evaluated unnecessarily in many of the timesteps, due to inefficient use of adaptive multistep numerical schemes. In their setup, Metcalfe et al. (2015) use an outer loop with fixed timesteps equal to the steps in the input rainfall to advance the solution in time. LSODE (Livermore Solver for Ordinary Differential Equations [ODEs], Petzold, 1983) is then invoked in an inner loop to distribute subsurface flow between units (with an arbitrary maximum of four substeps, and thus no "numerical error control"). This leads to timestepping inefficiency because advanced algorithms such as LSODE are capable of maintaining a user-specified time integration error while maximizing the timestep size, for example, in the inter-storm period when the catchment is in, or close to, a steady-state and these solvers can take much larger steps without loss of accuracy.

In parallel to the inefficiency of fixed timestepping, which is widespread in numerical hydrology (La Follette et al., 2021), for nearly two decades there have been urgent calls to move away from these schemes (and toward adaptive multistep methods), due to their significant adverse impact on the numerical solution accuracy/stability (Clark & Kavetski, 2010), distorted parameter distributions obtained via calibration (Kavetski & Clark, 2010), and inconsistency of performance when applied across different calibration data resolutions (Kavetski et al., 2011; Schoups et al., 2010), in particular under extreme precipitation scenarios (La Follette et al., 2021). Despite having long been accepted as "essential" (Kahaner et al., 1989; Shampine, 1994) in other areas of science and engineering, adaptive schemes have rarely been applied in conceptual hydrological modeling; and even where they have, the applications have been limited to simpler models (e.g., simplified version of steady-state TOPMODEL; Clark & Kavetski, 2010; Kavetski et al., 2003) in conjunction with relatively simple adaptive schemes (e.g., Clark & Kavetski, 2010; Kavetski & Clark, 2010; Schoups et al., 2010).

To our knowledge, application of more complex models in general, and Dynamic-TOPMODEL in particular, together with more advanced libraries of canned solvers, such as MATLAB's ODE suite (Shampine & Reichelt, 1997) has not been attempted before. The lack of application is primarily due to code complexity and/or a lack of open source access (Clark & Kavetski, 2010), making it difficult to retrofit these sophisticated algorithms to handle problem-specific ODE physical solution constraints; not least in a theoretically based and efficient manner (Clark & Kavetski, 2010). Alternatively, the mathematical model can be reformulated into a

constraint-handling form to allow the use of such solvers without the need for retrofitting the solver code. This is the motivation for our paper. Here, we propose Generalized Multistep Dynamic (GMD) TOPMODEL, by modifying and reformulating Dynamic-TOPMODEL of Metcalfe et al. (2015), and solve the resulting system of equations using MATLAB's ODE suite.

Note that while there are core similarities between GMD-TOPMODEL proposed here and the Dynamic-TOPMODEL of Metcalfe et al. (2015), there are also key differences intended to: extend the model's applicability to a wider range of landscapes where previously Dynamic-TOPMODEL's assumptions would break down; provide flexibility in better representing spatial heterogeneity; and improve physical process representations. More specifically, these are: (a) an "iso-basin" hillslope discretization which allows flexibility to represent spatial heterogeneity where such information is available or is the study focus; (b) a flexible diffusion-wave hydraulic channel routing coupled to the hydrologic model within the ODE solution (requiring derivation of both upslope and downslope FDMs, as well as a diffusion matrix); (c) relaxation of the assumption of water-table parallelism to ground surface, and derivation of different FDMs for surface and subsurface flows (previously subsurface flows were distributed using surface FDM); (d) adoption of a general power-law vertical hydraulic conductivity profile as opposed to the more limiting exponential form traditionally used; (e) continuous representation of hydraulic conductivity across saturated and unsaturated zones (previously the two were unrelated); (f) depth dependent hillslope and channel velocities using Manning's equation (previously a constant); and (g) tilting the frame of reference to the more physically correct version, that is, along the slope, as opposed to a horizontal frame typically used, leading to extended applicability of the model to steeper catchments (previously only applicable to moderate slopes).

The justification for the above modifications can be derived either from first principles, or from existing studies, and are presented in the body of the article. Note that whether such modifications lead to improved model performance under different conditions/catchments is a separate question which requires elaborate multi-catchment multi-period hypothesis testing and parameter-switch experiments to answer and is beyond our scope. However, by making our model open-source, we hope that these new model features will be put to the test fairly quickly as the model is applied to different catchment types. Importantly, here we demonstrate the successful implementation of GMD-TOPMODEL for a peatland catchment test case. We assess the model's performance within a Generalized Likelihood Uncertainty Estimation (GLUE) framework with limits of acceptability (LOA). We report model runtimes, curve fitting capabilities and sensitivity to resolution of both iso-basins and TI.

2. Mathematical Model Underlying GMD-TOPMODEL

2.1. Underlying Assumptions

2.1.1. Dynamic-TOPMODEL

The core assumptions of Dynamic-TOPMODEL (Beven & Freer, 2001; Metcalfe et al., 2015) are: (a) phreatic surface is always parallel to the ground surface, thus the hydraulic gradient in the saturated zone is approximated using the mean local "surface" slope $\overline{\beta}$ (°); and (b) lateral transmissivity decays exponentially with depth. Under these assumptions, during a rainfall event local propensity to saturation at DEM cell k is estimated using the TI of M. Kirkby (1975):

$$\gamma_k^* = \ln\left(\frac{a_k}{\tan\left(\overline{\beta}_k\right)}\right) \tag{1}$$

where $\overline{\beta}_k$ is the local slope at cell k, and a_k (L) is the upslope contributing area per unit contour length (contour length is taken to be equal to DEM cell size c_s (L)) draining to cell k. We recall the prerequisites for a justified application of the above two assumptions, which are (Beven et al., 2021): (a) shallow systems with (b) moderate slopes, where (c) exponential hydraulic conductivity profile is likely to be valid.

Our interpretation of the shallowness requirement (a) is that in landscapes where the thickness of the permeable upper layer is small relative to hillslope length, there is very little piezometric head buildup possible before water reaches the surface. Meaning that before the phreatic surface gradient deviates significantly from that of the ground surface, water is out of the subsurface domain. But the same is not true for deeper systems where

approximation of phreatic surface slope using local surface gradient can lead to significant under/over-estimation (C. Harman & Sivapalan, 2009; C. J. Harman et al., 2010).

Regarding the "moderate" slope requirement (b), note that appearance of $\tan\left(\overline{\beta}_k\right)$, as opposed to $\sin\left(\overline{\beta}_k\right)$, in Equation 1 is the consequence of choosing a horizontal frame of reference independent of the hillslope orientation. Using $\tan\left(\overline{\beta}_k\right)$ is equivalent to imposing a lateral subsurface hydraulic gradient of $\tan\left(\overline{\beta}_k\right)$, that induces a base flow horizontally over the plan distance c_s. Whereas in the more physically correct frame of reference wherein subsurface water moves parallel to the phreatic line, the true hydraulic gradient is $\sin\left(\overline{\beta}_k\right)$ Montgomery and Dietrich (2002, 1994), Borga et al. (2002), and Chirico et al. (2003). This explains why Dynamic-TOPMODEL (Beven & Freer, 2001; Metcalfe et al., 2015) should not be applied to very steep catchments, because the error in approximation increases with slope. But Dynamic-TOPMODEL is also not applicable at very gentle (near flat) catchments, because in such terrain diffusion dominates the flow rather than topographic gradient as Dynamic-TOPMODEL assumes; hence the applicability of Dynamic-TOPMODEL only to "moderate" slopes (not very steep or flat).

Regarding requirement (c), note that the choice exponential profile is a legacy of the original TOPMODEL's (Beven & Kirkby, 1979) steady-state and spatial uniformity of runoff assumptions (assumptions (a) and (b) discussed in the introduction), in which case an exponential profile is mathematically and physically more consistent, because it leads to very rapid transients, quickly leading to a steady state, in turn leading to rapid dissipation of any initial spatial nonuniformity in runoff (M. J. Kirkby, 1997). However, due to relaxation of both of (a) and (b) assumptions, Dynamic-TOPMODEL no longer requires the exponential profile for its consistent mathematical derivation or physical interpretation. In fact, the recession limb of hydrographs is not always of the first-order hyperbolic function of time that an exponential profile implies (Beven et al., 2021). Thus, a more general profile (such as power-law (Duan & Miller, 1997; Iorgulescu & Musy, 1997)) is an obvious extension to fully utilize a "dynamic" model.

For a more general-purpose model, we seek to free GMD-TOPMODEL from the requirements (a), (b), and (c), as discussed in the following section.

2.1.2. GMD-TOPMODEL

To extend the applicability of GMD-TOPMODEL to systems with deeper subsurface layers, we relax assumption (a). We still assume that base flow runs parallel to the phreatic line of slope $\overline{\alpha}_k$, but the phreatic line is not necessarily parallel to the surface slope $\overline{\beta}_k$. To estimate $\overline{\alpha}_k$ from the available DEM, similar to Hjerdt et al. (2004), we also recognize that, in general, the phreatic surface tends to be smoother than the ground surface. The level of this smoothness can be controlled by a parameter which determines how much of the surface microtopography is to be filtered out in order to reconstruct an estimate of phreatic surface gradient from the available data, that is, surface topography.

To do this Hjerdt et al. (2004) defines a reference elevation drop ΔZ_{ref} (L) then finds the distance that a parcel of water has to travel along the single direction downslope surface flowpath, for its elevation to drop ΔZ_{ref} meters (see Hjerdt et al., 2004 for visual representation). They then define mean phreatic surface gradient $\overline{\alpha}_k^*$ as:

$$\overline{\alpha}_k^* = \tan^{-1}\left(\frac{\Delta Z_{\text{ref}}}{d_k}\right) \tag{2}$$

where d_k (L) is the downslope distance from cell k along the single direction surface flowpath over which hydraulic head drops by ΔZ_{ref}. While this method is an improvement in that it allows a distinction between surface and subsurface hydraulic gradients, it has two drawbacks. First, it assumes subsurface flow follows the downslope surface flow path, but there is no objective reason to suggest that this is always the case, especially in deeper systems where there tends to be less correlation between surface topography and water-table shape (Beven et al., 2021). Second, it ignores the gradient in other directions and assumes all subsurface water flows to the single steepest downslope cell, which is also generally not the case (Quinn et al., 1991).

Here, we do not assume that subsurface water necessarily follows surface flowpaths, nor that it flows only in the direction of steepest descent. Instead, for DEM cell k, we consider the nearest distance in eight directions in which the elevation drops by ΔZ_{ref} meters. Clockwise from target cell k these are: north (0° or 360°), north east (45°),

east (90°), south east (135°), south (180°), south west (225°), west (270°), and north west (315°). Subsurface hydraulic gradient in each direction is then given by ($r = 1, 2, \ldots, 8$ denotes the direction IDs):

$$\alpha_k^r = \tan^{-1}\left(\frac{\Delta Z_{\text{ref}}}{d_k^r}\right) \tag{3}$$

Consider an X-Y coordinate system wherein each DEM cell's location is defined by a unique (X, Y) combination. The positional angles of cell k with respect to all other DEM cells will not necessarily be equal to the eight directions defined above. For this reason, after calculating the positional angles between cell k and all other DEM cells, we group these angles into eight bins. When searching for the nearest distance in direction r, that is, d_k^r, we take the nearest distance experiencing the head drop ΔZ_{ref} among all cells that fall into bin $\#r$.

Note that depending on the ΔZ_{ref} value, some or most DEM cells may not experience a head drop ΔZ_{ref} in all eight directions, no matter how far in those directions the search is extended, within the bounds of the catchment (treated as no-flow boundaries), for example, on local ridges. In extreme cases where a DEM cell does not experience ΔZ_{ref} head drop in any direction (such as at or near catchment boundaries), to avoid having cells that eternally accumulate water, ΔZ_{ref} value is progressively halved until the cell experiences a head drop in at least one direction. This way each cell k will have at least one, and at most eight α values. We then take the mean hydraulic gradient for cell k, $\overline{\alpha}_k$, to be the linearly weighted mean of downslope gradients, where larger gradients get larger weights and vice versa.

This approach comes with the caveat that it relies on the subjective choice of ΔZ_{ref}. Hjerdt et al. (2004) suggest that the typical elevation difference between adjacent cells can give an indication of the proper value for ΔZ_{ref} in a catchment-specific way. We further propose that the water-table depth during inter-storm periods (i.e., indicative of maximum water-table depth), if known, would be another useful indicator of the proper ΔZ_{ref} value, because it provides an objective estimate of the range within which the water table is expected to be able to fluctuate. In other words, "reasonable" ΔZ_{ref}, and the resulting "reasonable" search radius in the eight directions, should be dependent on both the DEM relief and soil thickness (or depth to the impermeable bed boundary), such that if the thickness is less than typical relief, then it provides an upperbound on ΔZ_{ref}, otherwise the average elevation difference can be used as a guide value. For example, our peatland catchment case study (introduced in Section 4), has a shallow water table, that is, no more than 1m deep, whereas DEM relief can be much larger than 1 m in some parts. Therefore in this case we set $\Delta Z_{\text{ref}} = 1$.

Despite the above caveat, the eight-directional subsurface gradients obtained using our method have two important implications: (a) for the first time, they allow calculation of FDM (containing information regarding the fraction of flow going from each HSU, to other HSUs—will be discussed in Section 2.3.1) for subsurface flow, that differs from surface flow FDM. The level of deviation of the two is controlled by ΔZ_{ref} such that larger values will produce a smoother phreatic surface compared to the ground surface. (b) When calculating the TI using this new subsurface hydraulic gradient (see Equation 4), we are taking into account not only the upslope controls on local saturation (through upslope contributing area), but also downslope controls (reflected in $\overline{\alpha}_k$) known to be a significant control in some cases (Hjerdt et al., 2004; Lanni et al., 2011; Speight, 1980).

To extend applicability of GMD-TOPMODEL to steeper hillslopes, we use the correct frame of reference wherein water moves along the slope and not horizontally over a plan distance, which allows us to better account for slope when calculating surface, unsaturated zone and saturated zone fluxes. We discuss these in detail in the next section where we introduce the model equations.

Finally, we also extend GMD-TOPMODEL's applicability to a wider range of lateral transmissivity (depth integrated hydraulic conductivity) profiles by relaxing assumption (b) of Section 2.1.1. We do so by assuming that the transmissivity profile takes a more general power-law form, of which the default TOPMODEL's exponential profile is a special case, as shown by Duan and Miller (1997) and Iorgulescu and Musy (1997).

Taking into account these modifications, the new TI for cell k takes the more general form:

$$\gamma_k = \left(\frac{a_k}{\sin\left(\overline{\alpha}_k\right)}\right)^{1/d} \tag{4}$$

where d (-) is the power-law exponent. In the following section, we derive the governing equations based on the generalized assumptions discussed here.

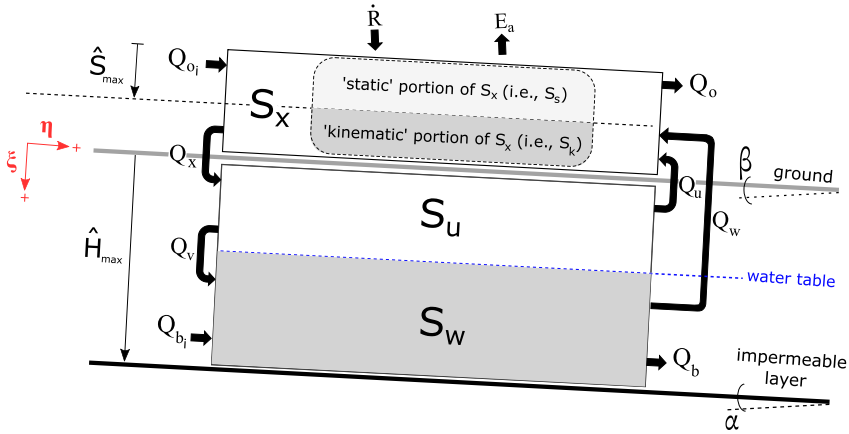

Figure 1. Schematic of the parameters and fluxes within each Hydrologically Similar Unit in Generalized Multistep Dynamic-TOPMODEL. S_x, S_u, and S_w are solution variables and therefore known in each timestep. Given the input parameters listed in Table 1, all other variables/fluxes can be calculated.

2.2. Inside Each HSU

In Figure 1, η (L) and ξ (L) define the reference frame for the model: η is the slope parallel, and ξ (L) is the slope perpendicular direction. Both are expressed as lengths but (in common with previous TOPMODEL formulations), η is a physical length while ξ is a "water column," or saturation thickness. To convert ξ to a physical thickness it must be divided by the porosity of the medium.

We represent each HSU with three main stores S_W (L), S_u (L), and S_x (L). Storage is always expressed as volume of water per unit area (thus with dimensions of length). S_w is the total storage in the saturated zone (i.e., below the water-table). S_u is the total storage in the unsaturated zone (i.e., between the ground surface and the water table). S_x is the total surface excess storage.

GMD-TOPMODEL has seven uncertain input parameters which are to be obtained via calibration. Note that these parameters are spatially averaged across the area of each HSU (i.e., lumped at the HSU level). These are listed in Table 1. For tractability, in the upcoming subsections we introduce parameters and fluxes shown in Figure 1 separately for each store. To better follow the derivations, keep in mind that S_x, S_u, and S_w are our primary solution variables (more details in Section 3), meaning that the values of these three variables are known in each timestep. Given the set of input parameters listed in Table 1, all other variables and fluxes can be found. Thus, what follows describes how model internal parameters and fluxes are obtained as a function of the three primary variables and input parameters. Finally, to highlight and distinguish model input parameters from other variables/parameters, we denote input parameters with (ˆ) throughout this paper (as presented in Table 1).

Table 1
Generalized Multistep Dynamic-TOPMODEL's Uncertain Model Parameters and Their Units

Parameter	Description	Unit
$\hat{d}$	Power-law exponent of transmissivity decay with depth	m
$\hat{T}_{max}$	Maximum transmissivity at saturation (at the surface)	$m\ s^{-1}$
$\hat{e}_p$	Annual average daily potential evapotranspiration rate	$m\ day^{-1}$
$\hat{S}_{max}$	Maximum static surface storage	m
$\hat{n}_{hs}$	Hillslope Manning's roughness coefficient	$s\ m^{-\frac{1}{3}}$
$\hat{n}_{ch}$	Channel Manning's roughness coefficient	$s\ m^{-\frac{1}{3}}$
$\hat{H}_{max}$	Catchment average maximum total subsurface storage	m

2.2.1. At the Surface

With reference to Figure 1, surface excess storage, S_x, is filled by rainfall rate $\dot{R}$ (LT^{-1}), overland inflow from other HSUs Q_{o_i} (LT^{-1}), and fluxes from the subsurface, Q_u (LT^{-1}) and Q_w (LT^{-1}). It is emptied by overland outflow Q_o (LT^{-1}), infiltration rate, Q_x (LT^{-1}), as well as actual evapotranspiration E_a (LT^{-1}).

Following Goudarzi et al. (2021), we conceptualize the total surface storage as the sum of a "static" storage, S_s (L), and a "kinematic" storage, S_k (L):

$$S_s = \begin{cases} S_x, & S_x \le \hat{S}_{\max} \\ \hat{S}_{\max}, & S_x > \hat{S}_{\max} \end{cases} \tag{5}$$

where $\hat{S}_{\max}$ (L) is the maximum possible static storage, and kinematic storage is given by:

$$S_k = S_x - S_s \tag{6}$$

Conceptually, "static" storage represents the portion of total surface water, S_x, that is unable to flow to other locations and is thus spatially fixed/immobile, that is, interception storage by vegetation canopy, storage in plant roots and pond storage. In contrast, kinematic storage, S_k, is the portion of S_x that is in motion at any given point in time; it is representative of the thickness of the sheet of water flowing over the surface. Note that S_s has to be filled before any water can flow either laterally (through Q_o) and/or vertically into the soil (through Q_x).

Given the above internal structure of our surface storage, we follow the previous versions of TOPMODEL (Beven & Kirkby, 1979; Metcalfe et al., 2015) in defining actual evapotranspiration e_a (LT^{-1}) as the potential (maximum) evapotranspiration e_p (LT^{-1}) scaled by water availability for evapotranspiration:

$$e_a = \left(\frac{S_s}{\hat{S}_{\max}}\right) e_p \times (24 * 60 * 60)^{-1} \tag{7}$$

e_p (LT^{-1}) is the potential "daily" evapotranspiration rate, hence the conversion from m/day to m/s above. A range of models of varying complexity exist to predict e_p. We choose to account for variation of e_p across a calendar year with the simple and commonly used sine curve approach (Ponce, 1989, pp. 222), and a similar sine wave for intra-daily variations, such that it is at its maximum at noon time and at its minimum at night time (respectively, the first and second square brackets below):

$$e_p = \frac{\hat{e}_p}{2}\left[2 + \sin\left(\frac{2\pi \times \text{day}}{365} - \frac{3\pi}{4}\right)\right] \times \left[2 + \sin\left(\frac{2\pi \times \text{hour}}{24} - \frac{\pi}{2}\right)\right] \tag{8}$$

where day (-) is the day number starting from first of January, and hour (-) is the hour number starting from midnight; $\hat{e}_p$ (LT^{-1}) (Table 1), is the annual average daily potential evapotranspiration rate. Note that this parameter can vary with catchment/period upon calibration.

There are physical solution constraints that need to be enforced when numerically (and therefore inexactly) obtaining the solution to a flow system. We introduce these constraints where they appear in the equations. For instance, here, depending on the $\hat{e}_p$ value (which is a calibrated parameter), it is possible for e_a to become larger than the stored water available for evapotranspiration, that is, static storage S_s. Thus, we enforce:

$$E_a = \begin{cases} e_a, & e_a \le S_s \\ S_s, & e_a > S_s \end{cases} \tag{9}$$

The mobile portion of surface storage, or the "kinematic" storage, S_k given by Equation 6, is routed on the surface and out of the HSU with overland flow flux q_o (LT^{-1}):

$$q_o = \frac{\upsilon S_k}{c_s} \tag{10}$$

where overland flow velocity is obtained from the diffusion wave approximation to the St Venant (shallow water) equations:

$$v = \pm \frac{R_H^{2/3}}{N} \left(\pm \left[\sin(\beta) - \cos(\beta) \frac{dS_k}{d\eta} \right] \right)^{1/2} \tag{11}$$

where N ($TL^{-1/3}$) is the Manning's roughness coefficient, which is either the $\hat{n}_{hs}$ or $\hat{n}_{ch}$ parameter (Table 1), depending on whether the HSU in question is a hillslope or channel, respectively. R_H (L) is the hydraulic radius defined as the ratio of cross sectional flow area to the wetted perimeter. For channel HSUs, assuming rectangular channel, $R_H = \frac{wS_k}{2S_k + w}$, where w (L) is the channel width; for hillslope HSUs $R_H = S_k$. Our DEM data is at 2m resolution, which happens to be a reasonable width for channels in Upper Ashop, thus we set the channel width to $w = c_s$ in this case.

Note that the $\pm$ signs are necessary to ensure real roots for the velocity term. As was mentioned in the introduction, physically the $\pm$ represent the direction of propagation (see Lighthill and Whitham (1955) for more information). The first term in Equation 11 represents the topography-driven flux which would generate flow only in the downslope direction, while the second term is the diffusion flux, which would dominate only in flatter areas and could result in upslope (when $v < 0$) as well as downslope ($v > 0$) flow. This requires derivation of both upslope and downslope FDMs, which are discussed in Section 2.3.1. Obtaining $\frac{dS_k}{d\eta}$ for each HSU requires derivation of a diffusion matrix which is discussed separately in Section 2.3.2.

Here, the physical solution constraint requires that the flux out of surface units does not exceed the available kinematic surface storage, that is, $q_o \leq S_k$, at all times (which can be violated depending on the Manning's roughness coefficient value). Thus:

$$Q_o^{\pm} = \begin{cases} q_o, & |q_o| \leq S_k \\ \frac{q_o}{|q_o|} S_k, & |q_o| > S_k \end{cases} \tag{12}$$

Flux into the surface store from other HSUs, that is, Q_{o_i} (LT^{-1}) in Figure 1, stems from the interaction between different HSUs and is described in Section 2.3.1.

2.2.2. In the Unsaturated Zone

Q_x (LT^{-1}) in Figure 1 is the infiltration flux from the surface to the unsaturated zone. Currently, it represents the case where there is no limit to the infiltration rate, and all mobile (kinematic) surface water (Equation 6) infiltrates until the subsurface is saturated, which is suited for our study catchment that is dominated by saturation excess overland flow. Infiltration excess overland flow could easily be incorporated for other catchment types, but it is not pursued here. To define the infiltration flux, we first need to know the volume (per unit area) available to receive infiltration in the subsurface. We define this as a storage deficit:

$$D = \hat{H}_{\max} - S_w - S_u \tag{13}$$

$\hat{H}_{\max}$ (L) is the catchment average maximum subsurface storage (Table 1). After calculating the subsurface storage deficit, D (L), the infiltration flux is given by:

$$Q_x = \begin{cases} S_k, & S_k \leq D \\ D, & S_k > D \end{cases} \tag{14}$$

Traditionally in Dynamic-TOPMODEL, the process of percolation of water through the unsaturated zone and to the water table is represented by a linear reservoir model where vertical flux from the unsaturated zone to the saturated zone is taken to be $\frac{1}{t_d} \frac{S_u}{H_u}$, that is, the ratio of the available unsaturated zone storage, S_u, to the thickness of the unsaturated zone delayed by a mean residence time, t_d (TL^{-1}), where H_u (L) is defined as:

$$H_u = \hat{H}_{\max} - S_w \tag{15}$$

Comparing the $\frac{1}{t_d}\frac{S_u}{H_u}$ approximation to single phase Darcy's Law, which approximates vertical flux as $K\frac{d\psi}{d\xi}$, where ψ (L) is vertical hydraulic head in ξ direction and K (LT^{-1}) is hydraulic conductivity in ξ direction, shows that the $\frac{S_u}{H_u}$ is equivalent to an average hydraulic head across the thickness of the unsaturated zone (i.e., H_u), and t_d represents an inverse of a depth-averaged vertical hydraulic conductivity (averaged across H_u).

Given that the saturated zone's hydraulic conductivity is represented by Equation 16 (expressed as transmissivity, which is a depth integrated conductivity), it is clear that allowing an independent parameter t_d to govern the unsaturated zone's hydraulic conductivity is equivalent to assuming two unrelated conductivity profiles for the same soil. Since soil conductivity profile is often defined as a function of moisture content alone (Boll et al., 1998; Frankenberger et al., 1999; Walter et al., 2002) such a two-profile representation is somewhat arbitrary. A better approach would recognize that both saturated and unsaturated hydraulic conductivity are constrained by the same physical soil properties with their difference controlled by moisture content.

Having said that, relating hydraulic conductivity to soil moisture content generally requires information on residual water content and empirical shape parameters describing the soil water retention curve (Van Genuchten, 1980), thus introducing at least three additional uncertain parameters to be calibrated. This problem can be worse in specific catchment types such as peatlands, which may require a dual porosity function to differentiate between macropore flow and flow through peat matrix (Holden, 2009; Rezanezhad et al., 2016). Thus, any potential gain from improved, moisture content dependent hydraulic conductivity representation may be lost to equifinality in the additional uncertain parameters (Perrin et al., 2001). Here, we suggest that a better (single-profile) representation of hydraulic conductivity is possible without resorting to additional parameters.

Using the hydraulic conductivity profile for the saturated zone to derive the unsaturated zone profile. GMD-TOPMODEL uses a general power-law saturated transmissivity profile suggested by Duan and Miller (1997), and also Iorgulescu and Musy (1997) (as opposed to dynamic-TOPMODEL which assumes an exponential profile, as was discussed in Section 2.1):

$$T = \hat{T}_{\max}\left(\frac{S_w}{\hat{H}_{\max}}\right)^d \tag{16}$$

where $\hat{T}_{\max}$ (LT^{-1}) (Table 1) is the maximum lateral transmissivity (when soil is fully saturated), $\hat{H}_{\max}$ is the average maximum subsurface storage, and d (L) (Table 1), is the power law exponent which controls the rate of decay of transmissivity with depth. For the unsaturated zone, we propose a vertical head-based Darcy's flux, for two-phase flow, but when one of the phases is assumed inviscid (i.e., air in this case). Under such assumptions, the vertical flow in the unsaturated zone can be approximated by (L. Wang et al., 2009; Smith et al., 2006; Yang et al., 2000):

$$q_v = K\left(-\frac{d\psi}{d\xi} + 1\right) \tag{17}$$

We retain the original TOPMODEL's vertical hydraulic gradient in the unsaturated zone $\frac{d\psi}{d\xi} \approx \frac{S_u}{H_u}$, that is, as an effective hydraulic head across the thickness of the unsaturated zone, H_u. Since a mean hydraulic gradient over thickness H_u is assumed, the corresponding mean hydraulic conductivity needs to be calculated from Equation 16 and the following integral:

$$\overline{K} = \frac{1}{H_u}\int_{H_u}^{0}\frac{dT}{d\xi}.d\xi = \left(\frac{\hat{T}_{\max} - T}{H_u}\right) \tag{18}$$

But the above integral is valid under "saturated" soil conditions (because T (LT^{-1}) in Equation 16 is the saturated transmissivity). Thus we assume that deviation of depth-averaged "unsaturated" hydraulic conductivity from that of saturated conductivity is proportional to the deviation of the unsaturated layer from fully saturated conditions which is $\frac{S_u}{H_u}$:

$$\overline{K} = \left(\frac{\hat{T}_{\max} - T}{H_u}\right)\left(\frac{S_u}{H_u}\right) \tag{19}$$

which leads to the following vertical flux from unsaturated zone to saturated zone:

$$q_v = \overline{K}\left(-\frac{S_u}{H_u}\cos(\beta)+1\right) \tag{20}$$

The $\cos(\beta)$ term adjusts the unsaturated zone thickness H_u, which is measured in a surface normal direction, to account for vertical unsaturated zone flow, which travels a longer distance than the surface normal distance to reach the water table.

Another solution constraint is required here, because depending on the $\hat{T}_{\max}$ value (Table 1), q_v may be larger than the available S_u, which requires the following solution constraint to be enforced:

$$Q_v = \begin{cases} q_v, & q_v \le S_u \\ S_u, & q_v > S_u \end{cases} \tag{21}$$

In defining Q_u in Figure 1, we note that depending on the balance of input and output fluxes to the unsaturated zone, during numerical solution its storage may exceed the available storage in this zone. This mimics the situation where this store is full, but the extra storage cannot simply be discarded (issue of manually "zeroing the fluxes" Clark & Kavetski, 2010). Thus, the excess unsaturated zone storage is added to the surface as surface excess. This flux is given by:

$$Q_u = \begin{cases} S_u - H_u, & S_u > H_u \\ 0, & S_u \le H_u \end{cases} \tag{22}$$

2.2.3. In the Saturated Zone

In the saturated zone, the base flow along the slope α is given by the Boussinesq equation (Childs, 1971):

$$q_b = \frac{T}{c_s}\left(\sin(\alpha)-\cos(\alpha)\frac{dS_w}{d\eta}\right) \tag{23}$$

where the first term represents flow due to hydraulic gradient and the second term flow due to diffusion, and T is given by Equation 16. Here again if diffusion dominates, it can generate flow in the upslope direction, which will be routed using an upslope FDM (see Section 2.3.1). How to obtain $\frac{dS_w}{d\eta}$ for each HSU is discussed separately in Section 2.3.2. Physical solution constraint requires:

$$Q_b^{\pm} = \begin{cases} q_b, & |q_b| \le S_w \\ \frac{q_b}{|q_b|}S_w, & |q_b| > S_w \end{cases} \tag{24}$$

Flux into the subsurface store from other HSUs, that is, Q_{b_i} (LT^{-1}) in Figure 1 is described in Section 2.3.1.

Similar to Q_u (Equation 22), in defining Q_w in Figure 1, we note that depending on the balance of input and output fluxes to the saturated zone, its storage may exceed the available storage in this zone. Thus, the excess unsaturated zone storage is added to the surface as surface excess. This flux is given by:

$$Q_w = \begin{cases} S_w - \hat{H}_{\max}, & S_w > \hat{H}_{\max} \\ 0, & S_w \le \hat{H}_{\max} \end{cases} \tag{25}$$

Note that threshold functions such as Equation 25 are impossible to integrate due to presence of singularities (Kavetski & Kuczera, 2007). Thus, for numerical implementation of the physical solution constraints in

Equations 5, 9, 12, 14, 21, 22, 24 and 25, we use a continuous differentiable hyperbolic tangent function of the form below to evaluate the step-wise fluxes in those equations:

$$cond. = \frac{1}{2}\left(1 + \tanh\left(\frac{F - F_0}{\varepsilon}\right)\right) \tag{26}$$

where ε (-) is machine precision, set equal to 10^{-64} in our case. For condition $F > F_0$, the above function produces ones for HSUs where the condition is met and zeros for all other HSUs, this enables vectorized and therefore faster computations in MATLAB.

2.3. Interactions Between HSUs

2.3.1. Bulk Flow Distribution Matrices

The M8 multiple flow directional algorithm of Quinn et al. (1991) leads to a weighting matrix that assigns fractional flow from each DEM cell to its eight neighboring cells (see Section 2.1.2). Upslope neighbors, with elevation higher than the central (target) cell receive no flow. The remaining neighbors receive flow proportional to the slope in those directions. Thus for N cells, M8 generates an $N \times N$ sparse matrix of the fractional flows from each cell to all other cells. A computationally efficient MATLAB algorithm for generating this sparse matrix is available in TopoToolbox of Schwanghart and Scherler (2014).

Metcalfe et al. (2015) used the above fractional flow matrix to derive a downslope FDM for overland flow, $\mathbf{W}_o^+$ (+ denotes downslope and $_o$ overland flow) specific to the HSU discretization being used. For any given HSU classification, $\mathbf{W}_o^+$ contains information regarding the portion of downslope surface fluxes from each HSU entering other units. Given that many (or some, depending on HSU sizes) cells in each HSU will only have neighbors that are in the same HSU, most of the flow is recirculated within the same HSU, and only some is transferred to other units. This is referred to as the "recycling" property of the FDM, which accounts for the time each flow parcel spends within a given HSU. Metcalfe et al. (2015) used $\mathbf{W}_o^+$ to also distribute subsurface flow in the downslope direction, based on the assumption that phreatic surface is always parallel to the ground surface. Since this assumption has been relaxed in GMD-TOPMODEL, here, using the eight-directional subsurface gradients introduced in Section 2.1, we are able to derive a separate FDM for the subsurface in the downslope direction, that is, $\mathbf{W}_b^+$ (b referring to baseflow). Note that the derivation of $\mathbf{W}_b^+$ is exactly the same as $\mathbf{W}_o^+$ except that it uses subsurface gradients instead of surface gradients.

As mentioned in the previous section, it is possible for water to flow against the topographic surface gradient if the diffusion term in Equations 11 and 24 exceeds the gradient term (reflecting a dominance of pressure head over elevation head). In such situations, it is not appropriate to distribute the flow using a downslope FDM. To obtain an upslope FDM for the overland flow, that is, $\mathbf{W}_o^-$, we first invert the DEM, reversing the upslope-downslope relationships between adjacent DEM cells, such that the FDM routes the flow in the upslope direction. However, care is required because in the upslope direction cells with larger gradients relative to the central cell must receive less flow. Therefore, to obtain an upslope FDM for the subsurface, that is, $\mathbf{W}_b^-$, we use the inverted DEM and the same ΔZ_{ref} discussed in Section 2.1 to obtain new phreatic surface gradients in upslope direction, from which we then calculate a $\mathbf{W}_b^-$ in the same way as $\mathbf{W}_o^-$.

Note that FDMs also contain some boundary condition information. In particular, $\mathbf{W}_o^+$ and $\mathbf{W}_o^-$ are modified such that channel HSUs give flow only to other channel HSUs, meaning that once water reaches the channel network it remains within the channel network, unless it infiltrates to the subsurface through a dry channel segment. We do not impose the same boundary condition for subsurface channel HSUs, because surface flows are focused by the banks of streams and rivers while subsurface flows do not experience the same topographic focusing.

Overland flow fluxes from other units into each HSU, Q_{o_i} (LT^{-1}) ($_o$ denotes overland flow and $_i$ denotes influx), are calculated using $\mathbf{W}_o^+$, $\mathbf{W}_o^-$ and the fluxes out of each HSU, Q_o (Equation 12). The overall surface flux into each HSU, Q_{o_i} in Figure 1, is calculated using:

$$Q_{o_i} = \frac{1}{A_T}\left[\mathbf{W}_o^+ * \left(A \;\; Q_o^+\right) + \mathbf{W}_o^- * \left(A \;\; Q_o^-\right)\right] \tag{27}$$

where Q_o^+ and Q_o^- are the positive (downslope) and negative (upslope) branches of the surface flux in Equation 12 (note that $\mathbf{W}_o^+$ and $\mathbf{W}_o^-$ are matrices of size $n_c \times n_c$ whereas Q_o^+ and Q_o^- are vectors of size n_c). A_T is the total catchment area, which is sum of all HSU areas. A is the vector of surface areas of each HSU.

Traditionally, for a DEM of resolution c_s meters, A is taken to be the plan area equal to $n_{\text{cell}} \times c_s^2$ where n_{cell} is the number of DEM cells in a given HSU. An alternative way of calculating HSU areas, which we advocate, is to first estimate DEM cell areas along their slope and then aggregate HSU areas, such that cells with steeper slope have larger surface areas (proportional to their slope) than flat cells.

The plan area method is appealing because it is cheaper to numerically evaluate once n_{cell} is known, and because input rainfall to each DEM cell is always per unit plan area, no matter the cell slopes. However, this is an incomplete account of the role of surface areas, because they also affect the time each water parcel spends in a cell, the total volume of water it can hold; and subsequently, the inflow/outflow rate from/to other cells (as is evident in Equation 27), and total actual evapotranspiration, meaning that more accurate estimates of cell surface areas can be important, especially in landscapes with significant spatial variability in topographic characteristics, over length scales $\geq c_s$. Thus for HSU #k we calculate surface area using:

$$A_k = \sum_{l=1}^{n_{\text{cell}}} \frac{c_s^2}{\cos\left(\overline{\beta}_l\right)} \tag{28}$$

where $\overline{\beta}_l$ (°) is the linearly weighted mean of downslope surface gradients for DEM cell l, and n_{cell} (-) is the number of DEM cells in HSU #k. To ensure that the effect of such area treatment on rainfall input to each HSU is minimal, each HSU receives a fraction of total rain based on the fraction of the total planform catchment area it occupies.

Similarly to Equation 27, base flow distribution between HSUs is carried out using $\mathbf{W}_b^+$, $\mathbf{W}_b^-$ and the fluxes out of each HSU, Q_b (Equation 24):

$$Q_{b_i} = \frac{1}{A_T}\left[\mathbf{W}_b^+ * \left(A \;\; Q_b^+\right) + \mathbf{W}_b^- * \left(A \;\; Q_b^-\right)\right] \tag{29}$$

2.3.2. Bulk Diffusion Matrix

To calculate the $\frac{dS_k}{d\eta}$ and $\frac{dS_w}{d\eta}$ terms in Equations 11 and 23, we note that both these equations are written in bulk form for a given HSU where every quantity appearing in them is averaged across the HSU area (including the topographic and phreatic gradients β and α, which are averaged in all eight directions and for all DEM cells in a given HSU). Thus the only values for the $\frac{dS_k}{d\eta}$ and $\frac{dS_w}{d\eta}$ that are commensurate with the HSU based formulation is the bulk storage gradient, that is, average spatial gradient of storage for a given HSU, or, mean of $\frac{dS}{d\eta}$ (dropping the subscript to mean either of $_k$ or $_w$) in all directions and for DEM cells that are in one HSU.

Similar to FDM derivation, most (or some, depending on HSU size) DEM cells within HSU#k are adjacent only to the cells that are within the same HSU, which by definition have the same storage value, therefore their storage gradient is zero. For DEM cells in boundaries of HSU #k with HSU #j, the gradient would be $\frac{S_k - S_j}{c_s}$. Thus, the bulk storage gradient would be the weighted average of $\frac{dS}{d\eta}$ for DEM cells in HSU#k. For HSUs #k and #j, this weight would be the total number of DEM cells in HSU #k that are in contact with HSU #j, divided by total number of cells in HSU #k (so that weights sum to one). For internal cells that neighbor only the same HSU cells, the weight would be: total number of HSU cells minus those with different HSU neighbors, divided by total number of cells in HSU #k.

As such, the bulk diffusion matrix, $\overline{\mathbf{D}}_M$, which is calculated once before running the model, is a matrix that contains the weights with which the bulk gradient $\frac{dS}{d\eta}$ should be calculated from the individual HSU storage values (whether it is S_k or S_w). For HSU#k:

$$\frac{dS}{d\eta}\Big|_k = \overline{\mathbf{D}}_M * \left(\frac{\mathbf{S} - S_k}{c_s}\right) \tag{30}$$

where $\mathbf{S}$ is the vector of storage values for all HSUs. When implementing in MATLAB, we vectorize this operation such that it is performed for all HSUs at once during the solution, rather than in a one-at-a-time manner (i.e., in a loop) as Equation 30 denotes.

3. Adaptive Multistep Numerical Solution

3.1. Casting the Governing Equations in ODE Form

When solving numerical problems, it is necessary to distinguish between dependent and independent variables. In this case, space and time are independent variables, all other variables are dependent variables. Among the dependent variables, we chose the three storage S_x, S_u, and S_w (see Figure 1) as Primary Dependent Variables (PDVs), or state variables, because the state of the system can be known once these three PDVs are known. The mathematical problem can then be described by three coupled mass conservation equations, which can be represented by Partial Differential Equations (PDEs) in space and time. Spatial discretization into HSUs further reduces the governing PDEs to coupled sets of ODEs in time only. These ODEs are:

$$\begin{cases} dS_x/dt &= \dot{R} - E_a \quad - Q_x + Q_u + Q_w + Q_{o_i} - Q_o \\ dS_u/dt &= Q_x - Q_v - Q_u \\ dS_w/dt &= Q_{b_i} - Q_b + Q_v - Q_w \end{cases} \tag{31}$$

Note the coupling of S_x with S_u through Q_x and Q_u, coupling of S_u with S_w through Q_v, and coupling of S_x with S_w through Q_w; also the coupling of HSUs through Q_{o_i} and Q_{b_i} terms. We initialize the system using:

$$\begin{cases} S_x = 0, & t = 0 \\ S_u = 0, & t = 0 \\ S_w = \hat{H}_{\max}, & t = 0 \end{cases} \tag{32}$$

but allow a few days of "spin up" period at the beginning of the record to allow the system to re-balance; during this period model's performance/prediction is not considered.

The modeled total output discharge (Q (LT^{-1})) consists of a base flow component (Q_b (LT^{-1})) and a overland flow component (Q_o (LT^{-1})). At any given time, surface flow reaching the outlet HSU, the outlet DEM cell, will leave the system with a rate given by Equation 12. Subsurface contribution to output hydrograph at the outlet cell is calculated via Equation 24. Thus, modeled output hydrograph is given by $Q(t) = Q_b^*(t) + Q_o^*(t)$, where * denotes outlet cell (HSU).

Finally, to minimize the effects of unknown antecedent conditions on our calibration, we use a "spin up" period (as shown in Figure 3b1) where that period of the record is not included when evaluating model performance.

3.2. Solution Using MATLAB's ode15s

We integrate Equation 31 in the time domain using the Method of Lines approach in which the spatially discretized PDEs (now ODEs in time) are solved simultaneously using any ODE-solver of choice. We use MATLAB's adaptive multistep stiff ODE-solver, ode15s, which solves initial value problems governed by first-order systems of ODEs. ode15s uses the so-called Numerical Differentiation Formulae (NDF), which are a modified version of Backward Differentiation Formulae (BDF) associated with Gear's method (Shampine & Reichelt, 1997). Both NDF and BDF methods are available in ode15s. In our simulations, we use NDF rather than BDF, as they are typically more efficient.

The Jacobian matrix contains information regarding the interdependencies of the components of the ODE set Equation 31 on one another. Since the Jacobian matrix of complex ODE sets, such as Equation 31, is always time variable, in MATLAB it is possible to provide the ODE-solver with a Jacobian "pattern" as a sparse matrix of zeros and ones to indicate where the Jacobian matrix is nonzero and thus needs to be evaluated. Specifying the Jacobian pattern a-priori can lead to savings in computation time because the solver avoids unnecessarily evaluating the full matrix. Note that HSUs only interact through the S_x and S_w variables. The sparsity pattern of this interaction is determined by GMD-TOPMODEL's catchment-specific FDMs, that is, $\mathbf{W}_o^+$, $\mathbf{W}_o^-$, $\mathbf{W}_b^+$ and $\mathbf{W}_b^-$, as well as the diffusion matrix $\overline{\mathbf{D}}_M$.

Table 2
Differences Between Dynamic-TOPMODEL and Generalized Multistep Dynamic-TOPMODEL

Feature	Dynamic-TOPMODEL	GMD-TOPMODEL
Time-stepping	Fixed	Adaptive
Channel routing	None	Diffusion wave
Frame of reference	Horizontal	Along the slope
Phreatic surface slope	Fixed to surface slope	Independent
Overland flow velocity	Fixed	Depth-dependent
Soil hydraulic conductivity profile	Exponential	General power law
Continuous hydraulic conductivity profile	No	Yes
Mechanism for constraining spatially distributed information	None	Isobasins

ode15s uses a variable timestepping scheme which changes the step-size according to accuracy requirements of the problem being solved, which are known to the solver using relative and absolute tolerance values (*RelTol* and *AbsTol*). *RelTol* controls the number of correct digits relative to the true solution, in the computed solution and *AbsTol* controls the absolute difference between the computed solution and the true solution. At each step, the error e in component i of the solution y must satisfy:

$$|e_i| \leq \max(RelTol * |y_i|, AbsTol) \tag{33}$$

In each timestep, and based on the above criterion, the solver decides whether the timestep needs to be reduced further, or, if a larger timestep will be sufficient (to save computation time). We use MATLAB's default *RelTol* (-) and *AbsTol* (-) values of 10^{-3} and 10^{-6}, respectively. For more information regarding how the solver works, see Shampine and Thompson (2001) and Shampine and Reichelt (1997).

Effectively, irrespective of the validation or forcing data resolution, ode15s uses a variable timestep such that the solution in time is always accurate at least to six digits (*AbsTol*); meaning that all stores of all HSUs are solved in a fully-coupled manner. In other words, continually through time, timesteps are chosen such to ensure all solution components (HSUs and their stores) simultaneously satisfy their respective mass conservation statement with a uniform and fixed accuracy.

Note that due to its variable step nature, ode15s will require rainfall intensity values ($\dot{R}$ in Equation 31) at times that lie between the available observed rainfall data points. For this reason, during solution, the solver uses cubic interpolation to obtain rainfall intensity values depending on the solver's time.

Having described all the model features, for tractability we summarize the differences between Dynamic-TOPMODEL and GMD-TOPMODEL in Table 2.

4. Case Study Description

4.1. The Catchment

The case study catchment of the River Ashop (Figure 2) is a 9.02 km^2 headwater system in the Peak District of northern England, ranging from 625 to 315 m altitude with around 1,550 mm of precipitation per year (Pawson et al., 2012). The catchment is dominated by blanket peat cover. Blanket peatlands are characterized by organic-rich deposits, typically 1–3 m thick, that form over largely impermeable substrates. The accumulated peat tends to cloak the underlying topography. Generally, these rain-fed peatland systems are dominated by shallow water tables for most of the year (Evans et al., 1999), a high saturated hydraulic conductivity in the upper parts of the soil profile, but a very low saturated hydraulic conductivity the rest of the soil profile (Holden & Burt, 2003a), such that waterlogging of the deeper layers is maintained year round. Saturation-excess overland flow tends to dominate the runoff response in blanket peatlands (Holden & Burt, 2003b). However, where the blanket peatland has been degraded, gully erosion can develop, and this erosion is thought to be particularly severe in the Ashop system with a high density of steep-sided channels that have incised into the peatland and frequent remnant hagg systems forming complex terrain (Pawson et al., 2012). Thus, water tables across the Ashop system can be

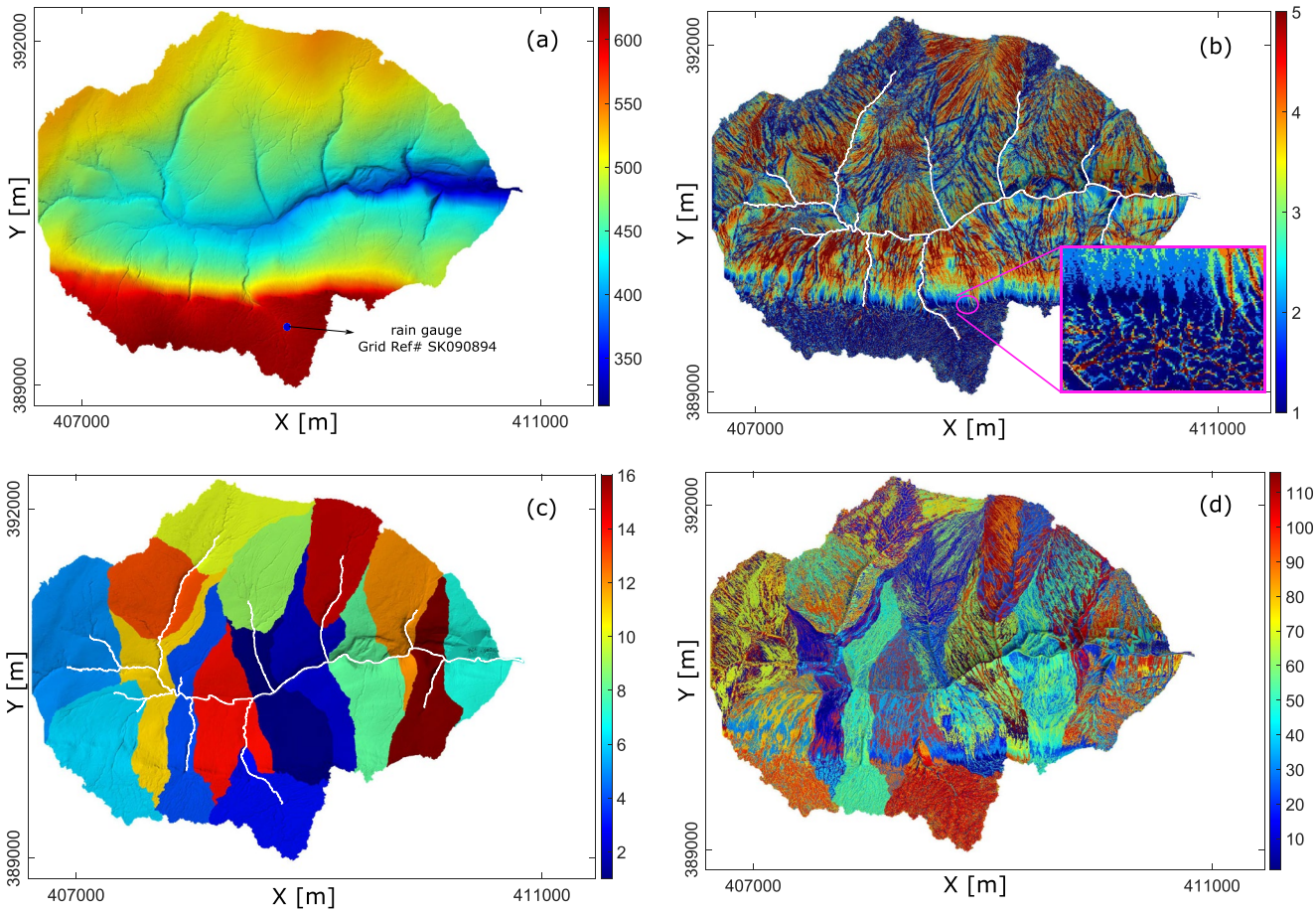

Figure 2. Upper Ashop catchment maps: (a) elevation (in meters); (b) Topographic Index values, from Equation 4, grouped into five bins ($n_c = 5$); (c) break up of catchment into smaller "iso-basins" each (roughly) 500,000 m^2 in area; (d) number of Hydrologically Similar Units that results from the combination of panels (b and c) groupings.

considerably deeper than in more intact blanket peatlands (Allott et al., 2009). However, these catchment features mean that we have a good test case for our model, aligning with our revised model assumptions.

All simulations in our study are for a 3 month period starting from 1 October 2014. Rainfall data were collected from a tipping bucket raingauge recording tips per timestep at 10 min resolution. 10 min resolution discharge data was estimated from a rated section of the channel where stage was measured using a pressure transducer. The rating curve was constructed from 10 salt dilution gaugings across a discharge range of 0.086–2.10 m^3/s (30% of peak discharge in the study period) with excellent agreement to a power law stage-discharge relationship ($r^2 = 0.997$—more details in Section 5.3).

4.2. Discretization Into HSUs

Traditionally, similarity based hydrological models group DEM cells into HSUs based on a catchment-specific TI of similarity, the most common form of which is Equation 1. This grouping is controlled by specifying the number of similarity groups (classes) n_c. Once n_c is specified, cells are grouped into HSUs either based on their TI rank, that is, the $\frac{N}{n_c}$ smallest index values are grouped together as HSU#1, through to the $\frac{N}{n_c}$ largest which form HSU#n_c; or based on predefined $n_c + 1$ bin edges. See for example, Figure 2b for rank-based grouping and $n_c = 5$. This grouping assumes that cells with similar values of γ tend to exhibit similar rainfall-runoff (RR) behavior independent of their specific spatial location, such that numerical solution of the representative units, that is, HSUs, would be sufficient to reproduce the catchment's response to rainfall.

By definition, however, when grouping is carried out solely based on γ values, local spatial information is lost, because similar γ values can exist anywhere within the catchment (the HSUs are often patchy and scattered

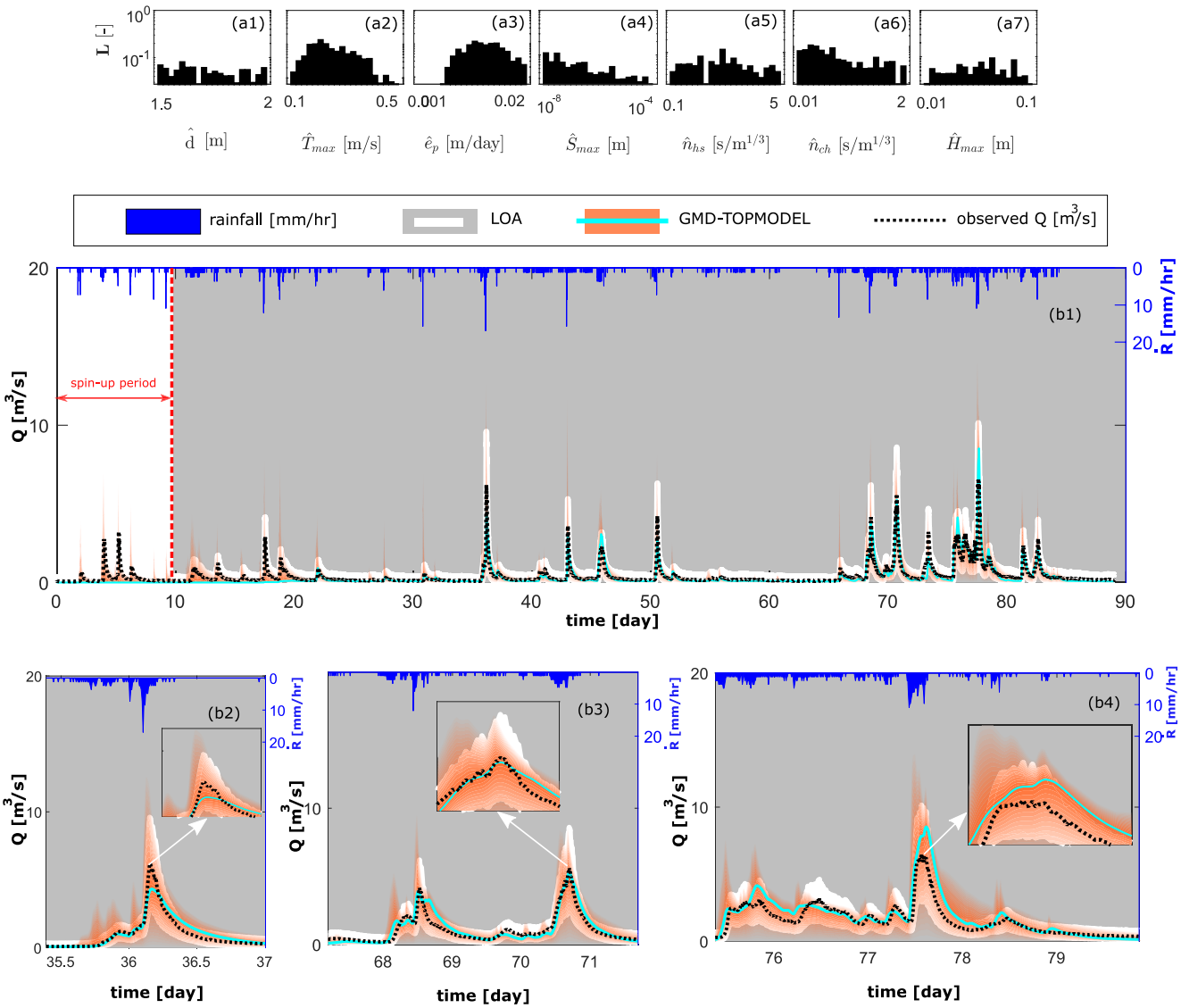

Figure 3. (a1–a7) Posterior parameter distributions for the base case, that is, with 16 iso-basins and 5 Topographic Index (TI) classes. (b1–b4) Corresponding discharge predictions where orange clouds are all behavioral predictions, and cyan lines indicate the best model prediction when compared to the observed discharge (black dotted lines), and according to the four performance metrics in Table 4. *Y*-axes are days starting from 1 October 2014.

across the catchment as in Figure 2b). Increasing n_c can improve spatial representation, but not in a systematic manner, for example, even setting n_c equal to the number of DEM cells does not guarantee convergence to a fully spatially distributed representation wherein each DEM cell is an HSU (in fact convergence is very unlikely). This is because various slope-area pairs (see Equation 4) can produce duplicate γ values (the likelihood of which increases with domain size). This can be a problem in that it would be difficult to reliably assess a model's sensitivity to the resolution of spatial discretization when it is not asymptotic to the fully spatially distributed case.

Additionally, given that similarity based discretization is predicated on the tendency to saturation according to non-unique (thus non-location-specific) topographic characteristics, where there are other location-specific factors (such as land cover, or rainfall, etc.) that may alter the catchment hydrograph response, similarity based discretization alone lacks flexibility in constraining or incorporating the location-specific information.

Consider, for example, a storm cell of size smaller than a catchment discretized based only on similarity of its DEM cells' γ values. Assuming spatially uniform rainfall across the entire catchment, will lead to both spatial and temporal approximation errors. Even if input rainfall to each HSU is specified as a separate time-dependent

boundary condition, the rainfall signal is still spatially diluted (dispersed) because HSUs will contain cells in (potentially) widely distant locations. One solution would be to include the rainfall in the similarity index, as suggested, for example, by Coxon et al. (2019). Given the non-stationary nature of the storm cell and its time-dependent rainfall rate, such a similarity index will be variable with time, meaning that the HSU grouping needs to be performed in each timestep. But the CPU time it takes to process DEM data is prohibitive, even for small catchments (which is why HSU formation is usually a pre-processed task). Also note that, as shown by Gao et al. (2018), this problem will not occur for a fully spatially distributed model, but the downside is considerably longer runtimes.

As such, to provide more flexibility, overlaying the γ based discretization, we devise an additional discretization which allows more flexible incorporation of spatial information, and in a way that it approaches the fully spatially distributed case upon successive refinements. Inspired by Lindsay (2018), but using a different algorithm, we discretize the catchment into smaller drainage areas. The size of these areas are controlled by a target area value A_{tr}. Based on A_{tr} our algorithm finds roughly similar size "iso-basins." To do this, it first narrows the search by finding the list of DEM cells with upslope areas $\geq A_{tr}$; then it loops through the list to find those cells with upslope areas equal to $A_{tr}\pm$ 5% tolerance, it then updates the "global" iso-basins map after each iteration of the loop. From the list, it then eliminates cells that are sat within an iso-basin to further narrow the search. At the end of the loop, if there are parts of the catchment that have not been assigned an iso-basin, the tolerance value is relaxed until no unassigned cells remain. Note that the smaller the A_{tr} and/or the larger the catchment the longer this process will take. However, this is a pre-processing task that only needs to be performed once per catchment/A_{tr} value. In our case, the iso-basin discretization process took roughly 2 hr on a PC with a 3.1 GHz CPU.

In simple terms, this modification gives users the flexibility to better constrain the model such that its HSUs are not allowed to send/receive mass (water) to/from DEM cells that are "too" far away, where this distance is set by adjusting A_{tr}. Others (e.g., Takeuchi et al., 1999) have chosen to discretize the catchment using square blocks of specified width, we believe the iso-basin approach is superior because it honors flowpaths. Note that iso-basins can be used to constrain any type of spatially distributed information. Note that through the use of iso-basins the model can be used to study land-cover change or soil types without further modification. However, for studying spatially variable rainfall additional modifications are needed which were not required for our study. Users interested in including spatially variable rainfall should define different rainfall zones within the model and assign IDs to "sets" of iso-basins that sit within each zone. Within the ODE function, separate rainfall time-interpolation for each zone needs to take place (because, being time-continuous the ODE solver's time will fall in between two rainfall data points; hence the need for interpolation).

Figure 2c, shows the iso-basin IDs for the case of A_{tr} = 500,000 m^2. Figure 2d, shows the 112 HSUs that result from combining (b) and (c), that is, each TI class in each iso-basin is assigned a different ID. Note that some of the 112 HSUs are channel reaches, which are defined as separate HSUs to be used in the channel routing. To define channel reaches, we first identify a channel network based on a channel initiation threshold area A_{ch}. The white lines in Figures 2b and 2c are the channel map at A_{ch} = 40, 000 m^2, that is, all DEM cells with $\geq$40,000 m^2 of drainage area are taken to be channels. The channel network is then broken up at the channel intersections into smaller reaches, but also at points where a channel segment spans across two neighboring iso-basins. The catchment outlet (i.e., the DEM cell with the largest upslope contributing area) is defined as a separate HSU, which is the HSU from which the hydrograph (surface + subsurface runoff) is extracted. In this case, the combinaiton of A_{ch} together with A_{tr} resulted in 38 different channel reaches each assigned a separate HSU ID.

Note that the channel network discretization can be completely independent of the iso-basins, and in the finest discretization limit, each DEM cell within the channel network can be defined as a separate HSU. This provides further flexibility in modeling the detailed channel network if needed, and where sufficient computational resources are available.

5. Model Calibration

5.1. Uncertainty Estimation Framework

Recall the uncertain input parameters of GMD-TOPMODEL listed in Table 1. These seven parameters are hereafter referred to as a parameter-set. We chose to use the GLUE methodology of Beven and Binley (1992) for our calibration. However, there are other "formal" Bayesian approaches available to choose from. For more discussion

Table 3
Calibration Parameter Ranges

Parameter (unit)	Run1	Run2	Run3
$\hat{d}$ (m)	1–10	1–5	1.5–2
$\hat{T}_{max}$ (m s^{-1})	10^{-6}–1	0.01–1	0.1–0.5
$\hat{e}_p$ (m day^{-1})	10^{-8}–0.1	10^{-6}–0.1	0.001–0.02
$\hat{S}_{max}$ (m)	10^{-8}–0.1	10^{-8}–0.01	10^{-8}–10^{-4}
$\hat{n}_{hs}$ (s m$^{-\frac{1}{3}}$)	0.001–10	0.1–10	0.1–5
$\hat{n}_{ch}$ (s m$^{-\frac{1}{3}}$)	0.001–10	0.01–5	0.01–2
$\hat{H}_{max}$ (m)	10^{-8}–1	10^{-4}–0.1	0.01–0.1

on pros and cons of each approach readers are referred to Clark et al. (2011) and the references therein. We assume a uniform prior "likelihood" distribution for all parameters. We then run 20,000 simulations with parameter-sets sampled randomly from this uniform prior distribution using a Latin Hypercube method. We then retain only the "behavioral" sets to generate the posterior distribution. Behavioral parameter-sets are those which result in modeled discharges that are within the "LOA" of the observed discharge, for 95% of the record, as well as fully within our Base Flow Limit (BFL). If the model identifies N_b number of behavioral predictions, then there will be N_b number of associated parameter-sets. Posterior distribution of individual parameters can then be obtained by constructing a histogram where the y-axis denotes the number of behavioral parameters in each histogram bin.

5.2. Parameter Ranges

In choosing the parameter ranges, we took a similar approach to Lane and Milledge (2013). We start with a wide range and sample the parameter space 5,000 times (Run1). We run the model and plot parameter distributions weighted based only on their performance metrics (NSE, PTE, and PME), regardless of whether they satisfy LOA (see Section 5.3) and BFL. This is only to get a sense for roughly where in the selected range the model is likely to concentrate its best predictions. Based on visual inspection of Run1 parameter distributions, we conservatively narrow each parameter range and re-sample another 5,000 parameter-sets (Run2). We repeat the process and further narrow the parameter ranges. Once we obtain a reasonably narrow range, we sample each parameter 20,000 times and perform the final and main run (Run3). On Run3, we impose the LOA and BFL to obtain parameter distributions and retain only the behavioral sets when assessing model performance. Parameter ranges for each run are shown in Table 3.

5.3. Limits of Acceptability

LOA are upper and lower bound limits on the measured discharge within which the model predictions must lie to be deemed "behavioral" and thus accepted. Note that, if available, it is better to use more than one flow gauge to improve the reliability of LOA (i.e., make LOA a harder criterion to satisfy), but we felt that a single gauge was sufficient for the purpose of demonstrating our model. In calculating the LOA we use the "voting point" likelihood method of H. K. McMillan and Westerberg (2015) to account for both aleatory and epistemic errors by randomly sampling single segment rating curves of the form $q = bS^a$, where q (LT^{-1}) is the discharge, S (L) is the stage, and b (-) and a (-) are coefficients of the rating curve. The choice of single segment curves is based on the 10 available gauging points (10 stage and 10 discharge) from the study site, with the official rating curve $q = 9.7S^{3.22}$, showing that a single segment power-law curve fits the data very well. We randomly sample a and b values from a uniform prior distribution 10,000 times. We sampled b from the 5–15 range, and a from the 2–6 range.

H. K. McMillan and Westerberg (2015) represent aleatory errors using a logistic distribution with scale parameter σ a function of normalized discharge (normalized by mean flow) fitted using an exponential relationship, with the location parameter μ set to zero. For our gauging data set however, we found that a nonzero location factor was necessary. We thus fit both σ and μ to the normalized discharge also using a power-law relationship. Further, H. K. McMillan and Westerberg (2015) assumed stage measurement errors are negligible relative to discharge measurement errors. However, here we chose to also account for stage measurement errors using the method provided by (Petersen-Øverleir & Reitan, 2005). Following Keeland et al. (1997), and making the conservative assumption that stage measurement accuracy was low for pressure transducers of the type used at our sites, we assume the standard deviation of error in the stage measurement was $\Omega = \pm 7.5$ mm. We then propagate this stage error through the official rating curve using a standard frequentist inferential methodology (Petersen-Øverleir & Reitan, 2005) to calculate 95% uncertainty bounds on the discharge as a consequence of stage measurement error alone:

$$dq = \pm 1.96 \times b.a.\Omega.S^{m-1} \tag{34}$$

dq is then the contribution stage measurement errors to the LOA and added to the rating curve + discharge measurement error contributions calculated earlier.

Note that there are other errors (e.g., input rainfall errors, sensor drift) that we have not been able to quantify and include in our LOA. Consequently, we have sought to allow for the above errors by retaining models even when their predictions fall outside LOA for up to 5% of the observed data points.

5.4. Base Flow Limit

Since performance metrics (e.g., Nash Sutcliffe) are not necessarily a reliable measure of model internal fluxes (Khatami et al., 2019), we impose an extra condition to reflect field observations in UK upland blanket peatlands which is discharge response to rainfall is dominated by overland flow (e.g., Daniels et al., 2008; Holden & Burt, 2002, 2003b). Thus, the BFL requires that <30% of predicted outlet discharge can be composed of base flow.

6. Results

From the 20,000 simulations performed during calibration, GMD-TOPMODEL produced 549 behavioral discharge predictions. The combination of 549 behavioral sets led to a set of parameter distributions (one for each uncertain parameter). These distributions are shown in Figures 3a1–3a7 as 20-bin histograms, where *x*-axes are parameter values and the *y*-axes are the normalized (by the total) number of behavioral predictions in each bin. The *y*-axes range from 0 to 1, and are plotted in log scale. These distributions provide insight into the internal processes of the model, and are particularly useful when comparing different catchments/interventions by comparing and analyzing shifts in their distributions (e.g., Goudarzi et al., 2021).

Figure 3b1 shows the entire 3-month record, and Figures 3b2–3b4 the three largest storms, visualizing both predicted and observed discharge (dotted black line). The collection of all behavioral predictions is shown as the orange cloud. The cyan line shows the best performer according to our performance metrics in Table 4. To ensure the best prediction is chosen such that it captures all stages of the hydrograph accurately, we use both high-flow weighted and low-flow weighted metrics as shown in Table 4. The best prediction is that with the highest performance across all four metrics, where all metrics are given equal importance.

Note that throughout the rest of this paper, when assessing model performance, we are only concerned with the single behavioral prediction that best matches the single available observation (and not the mean of performances of all behavioral predictions, nor the performance of the mean of all behavioral predictions). The reasons for this are explained in the discussion section. In this case, the identified best prediction shows good performance across all metrics: $NSE_H = 0.84$ (-), $NSE_L = 0.81$ (-), $RMSE_H = 0.29$ (m^3/s), and $RMSE_L = 0.135$ (m^3/s).

We also test the model's sensitivity to spatial discretization into HSUs, both in terms of the number of iso-basin and number of TI classes. In order to do this, using the results presented above as the "base case," we consider two scenarios: (1) when spatial discretization is changed (refined and coarsened) while the model is allowed to freely pick its behavioral parameters from the same pool of 20,000 prior parameters discussed in Section 5.2; and (2) when behavioral parameter-sets are fixed to those of the base case and only spatial discretization is changed. Scenario (1) assesses whether refining spatial discretization leads to improved model performance; equally whether coarser discretization results in degradation in performance. Scenario (2) assesses whether parameter-sets obtained via calibration at a certain spatial resolution, are applicable at coarser and/or finer resolutions.

Within each of the above scenarios, we separately apply the model when its number of iso-basins is halved and then when it is doubled. Similarly, when its number of TI classes is (approximately) halved and when it is doubled. This makes a total of five cases within each scenario. Figure 4 shows the parameter distributions for all five cases in scenario (1). Note that parameter distributions for all five cases in scenario (2) are by definition the same as base case (black color). When the model is allowed to change its parameter distributions (scenario (1)), visually, very significant variations in distributions are detectable when changing spatial resolution, whether it is changing number of iso-basins or TI classes.

Table 4
Performance Metrics

Performance metric (unit)	Equation
NSE_H (-)	$1-\frac{\sum(Q_o-Q_p)^2}{\sum(Q_o-\overline{Q})^2}$
NSE_L (-)	$1-\frac{\sum(\log(Q_o+0.1)-\log(Q_p+0.1))^2}{\sum(\log(Q_o+0.1)-\overline{Q})^2}$
$RMSE_H$ (m^3/s)	$\sqrt{\frac{\sum(Q_o-Q_p)^2}{N_o}}$
$RMSE_L$ (m^3/s)	$\sqrt{\frac{\sum(\log(Q_o+0.1)-\log(Q_p+0.1))^2}{N_o}}$

Note. *H* denotes high-flow weighted, and *L* low-flow weighted. In all cases Q_o is observed discharge, Q_p is predicted discharge. For NSE_H, $\overline{Q} = mean(Q_o)$ and for NSE_L, $\overline{Q} = mean(\log(Q_o + 0.1))$. Note that we cap low flows at 0.1 m^3/s; roughly 4.5% of the observed flow range was below this level.

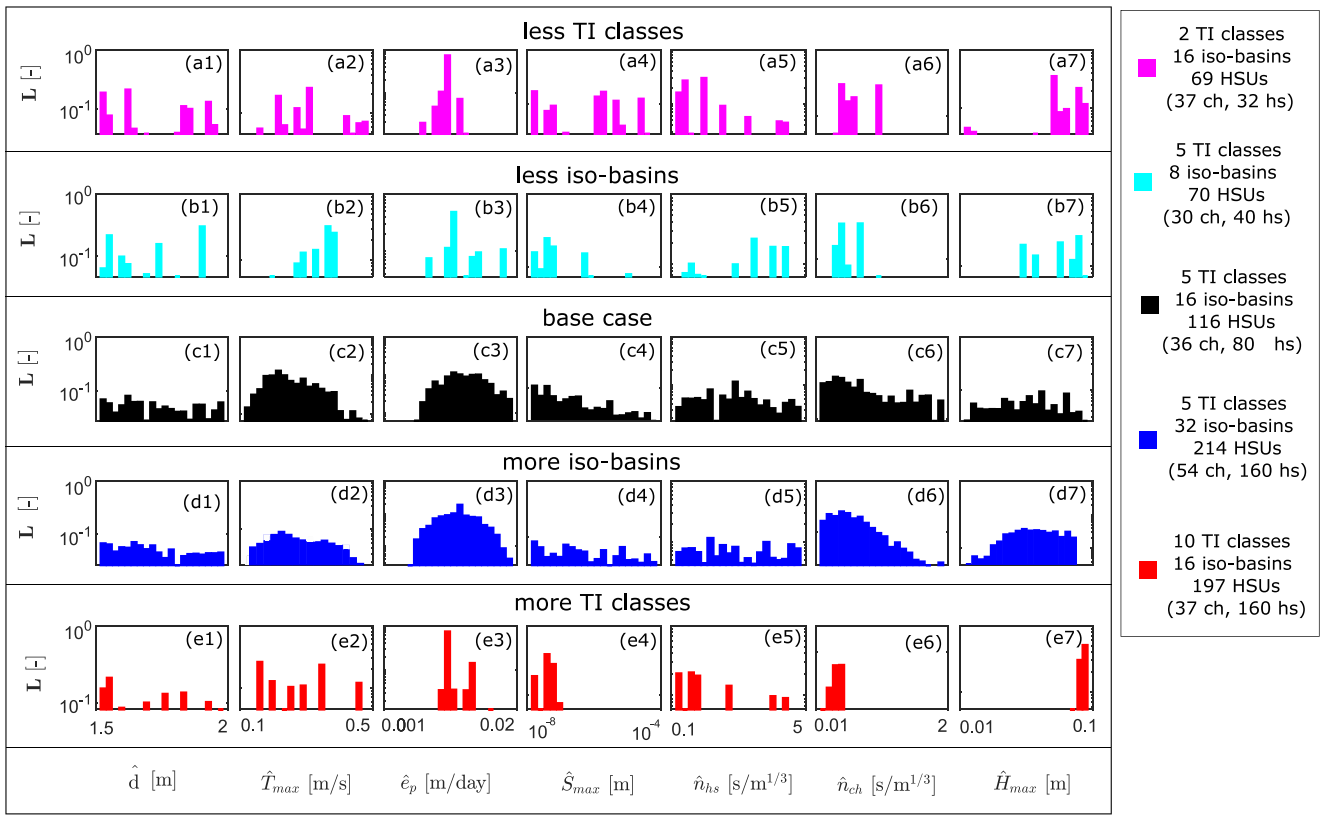

Figure 4. Parameter distributions for the five cases in scenario (1): changing spatial resolution while allowing the model to choose its parameters via calibration. The top two rows, a1–a7, and b1–b7, represent the coarsening of TI and iso-basins, respectively, and relative to the base case, that is, c1–c7. In the same fashion, d1–d7, and e1–e7, represent the refining of iso-basins and TI classes, respectively, and relative to c1–c7 (the base case).

To assess the impact of changes in distributions on model performance, we monitor best model prediction in each case, as shown in Figure 5. In this figure, scenario (1) is shown as solid squares and scenario (2) as hollow triangles. Each panel shows values for the four performance metrics in Table 4, for the best prediction. When parameter distributions are allowed to change in order to produce the best prediction possible (scenario (1)—squares), there is a relatively small variation in performance, independent of whether spatial discretization was refined or coarsened. Refining iso-basins leads to a slight improvement in performance across all four metrics. But both refining and coarsening of the TI classes seem to result in degradation in performance across all metrics. This may suggest that there is a specific number of TI classes that best conforms to the catchment-period pair in question. We note more on this in the discussion section.

When parameter distributions are fixed to those of the base case (scenario (2)—triangles), any change in resolution of spatial discretization leads to worse performance across the board. However, on a relative scale, the model is much less sensitive to changes in its number of iso-basins, than it is to changes in its number of TI classes. For the number of iso-basins, although both refinement and coarsening lead to performance degradation, but it was less so if finer iso-basins are used than when coarser ones are used. The reverse seems to be the case for the number TI of classes, where increasing the number leads to more performance degradation than is the case when fewer TI classes are used.

Finally, Figure 6 shows both the individual runtimes (solids circles + error bars) and the total (for the full 20,000 simulations—solid squares) runtime of the model in each case in scenario (1). It shows a clear positive correlation between the number of HSUs and runtime as is expected. Mean simulation times range from 37.5 to 918.2 s, and the total runtimes for the 20,000 simulations range from under one day to over 1 week, using parallel computing with 12 CPUs in parallel.

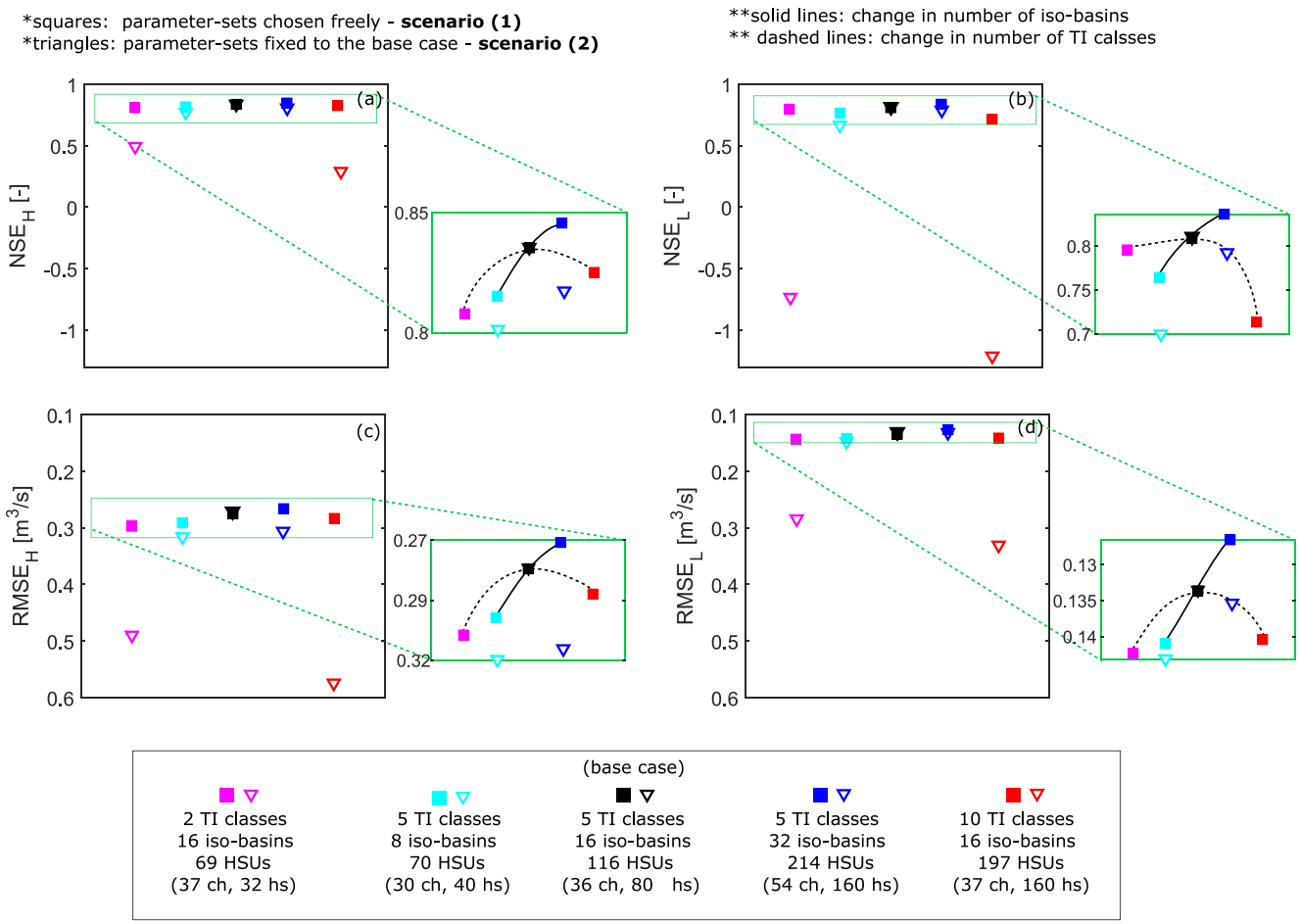

Figure 5. Tracking model performance with changing spatial resolution for both scenarios (1) and (2). Performance metrics are given in Table 4. "ch" denotes channel Hydrologically Similar Units and "hs" hillslope. Panel (a) shows the NSE values for high flows and for the different TI/iso-basin resolutions as denoted by the legend; similarly, panel (b) shows the NSE values for the low flows; panel (c) and (d) show the RMSE values for high and low flows, respectively.

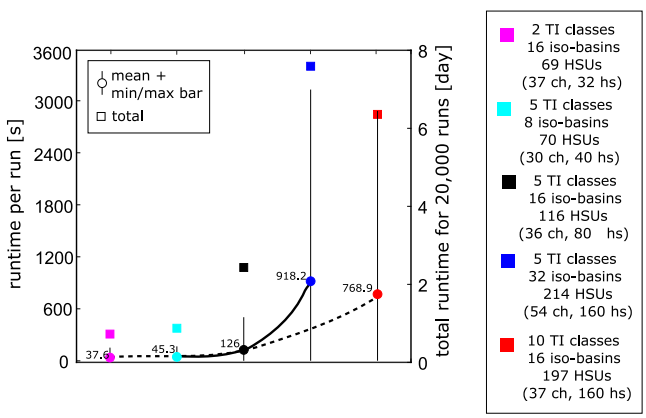

Figure 6. Model runtime performance for scenario (1).

7. Discussion

7.1. Performance Evaluation Framework

In relation to the modeling results and their interpretation, following Clark et al. (2011) we note that a RR model is a "hypothesis" about catchment function which encompasses a description of the dominant natural processes involved, and how they combine to produce a response to rainfall. Here, we test hypotheses about spatial discretization, rather than model structure (Renard et al., 2010), and instead assume that our chosen model structure mimics the catchment in question well (based on the catchment description in Section 4 and model assumptions in Section 2.1). But it does subject the findings to unknown degrees of model-related "epistemic" uncertainties (Beven, 2016) due to incomplete knowledge of the full set, or the dynamics, of catchment-specific processes and/or how to model them with sufficient accuracy or detail.

There are also model input parameter uncertainties that arise, partly from the general scarcity of field measurements of such parameters, but mainly due to scaling (commensurability) issues (Beven, 2016), where the time/spatial scale of model parameters are different from those measured in field. Parameter uncertainties are often accounted for by calibrating to observed discharge

data, but these data are not free of uncertainties themselves (see e.g., H. McMillan et al., 2012). In our case, for example, we define LOA (Beven, 2018) that represent the estimated precision with which we have been able to measure stage and convert it to discharge, meaning that our observed discharge record, Q_o, could have been any other time-series within LOA, yet the difference is undetectable by the combination of our equipment (e.g., pressure transducers) and methods (e.g., rating curve). Such imprecision makes it unlikely that Q_o will be exactly the same as the "true" catchment discharge record.

Given uncertain model parameters and observations, when calibrating a numerical RR model, we essentially allow model parameters to vary within the uncertainty of the observations (LOA), equivalent to numerically producing all (or as many as) possible realizations of true catchment discharge within the LOA. In this context, insofar as our hypothesis about model structure is true, each behavioral model prediction is a representation of different realizations of true catchment discharge. Given the single observation at hand (i.e., Q_o), it is not obvious how the relative performance of multiple behavioral predictions should be assessed.

One option is to compare all behavioral predictions to Q_o and assign higher likelihood of being the true catchment discharge depending on their goodness-of-fit to Q_o. This is equivalent to assuming that the single observation approximates the true catchment discharge with negligible error, and on this basis, worse predictions should be given less importance (likelihood). However, this runs the risk of punishing model predictions with worse fits to Q_o, even when in reality Q_o might be in error; the probability of which is not negligible according to the LOA.

Therefore, in the absence of a better alternative, we assess model performance in terms of its ability to reproduce the single available observation, Q_o. We examine goodness-of-fit of the best model prediction to Q_o as an indication that the model might also be reproducing other realizations (for which we do not have measurements) with comparable accuracy. As an aside, this should only be the case when assessing performance against past (available) data. When/if applying the model to make future predictions, or predictions in ungauged basins, "all" behavioral predictions should be considered, since they all equally represent one possible realization of true catchment discharge (insofar as our model hypothesis is true).

7.2. Scale Dependence of Model Parameters

Regarding hypotheses testing, here we are testing sensitivity to spatial discretization in terms of TI classification and iso-basins. Our underlying hypothesis in both cases was that increasing resolution of discretization would result in improved "realism," and lead to better model performance, or at least not worse. Degradation in model performance falsifies the hypothesis.

For iso-basins, we find that increasing their resolution, while allowing the model to adjust its parameters, improves model performance in both low-flows and high-flows, albeit not dramatically (around 2%–3% in NSE and 6%–7% in RMSE, see blue squares in Figure 5), confirming the hypothesis that finer iso-basins increases the "realism" of the model. Iso-basins have a well constrained functionality within the model: they limit the interference between storage dynamics for widely separated, but topographically similar, parts of the catchment (see Section 4.2). As such, a systematic behavior (i.e., refining iso-basins = better performance, and vice versa) can reasonably be expected. However, when the model parameters are fixed to the base case, both refining and coarsening of iso-basins slightly worsen performance (less so for refinement than coarsening; compare cyan and blue triangles). This suggests limited model sensitivity to changes in iso-basin resolution. However, parameter distributions (Figure 4) remain much more similar to the base case in response to iso-basin refinement than coarsening. Since Figure 5 only considers the best model prediction, but parameter distributions offer a wider view, we interpret these results as: if iso-basins are kept $\leq$ the size at which the model was calibrated, parameters are more likely to remain independent of the iso-basin size.

For TI classes, reducing their resolution worsens model performance across the board and relative to the base case (pink squares and triangles in Figure 5), consistent with our hypothesis. However, contrary to our hypothesis, increasing the resolution of TI classes worsens model performance, whether the model is allowed to adjust its parameters or not (see red squares and triangles). We note that TI classification assigns a priori "saturation tendencies" to different parts of the catchment. Assuming that there exists a map of saturation tendencies that best conforms to the catchment-period pair in question, arbitrarily reducing or increasing the resolution of TI classification, would both result in deviation from the true saturation tendencies. This would explain the non-systematic behavior, where both refining and coarsening TI classification worsens model performance. This underlines the

importance of better constraining the TI classification; as is also highlighted by the much higher sensitivity of model performance to changes in TI resolution than iso-basin resolution (compare red/pink triangles to cyan/blue triangles).

One way to improve TI classification might be to define the number of classes, n_c, as an additional calibrated parameter. The downside is increased dimensionality of the parameter space, and the necessity to redefine TI classes before each simulation. In particular, the CPU-time required to redefine TI classes is prohibitive especially within a Monte Carlo type calibration procedure. An alternative might be to manually tune the TI map such that the highest saturation tendency class roughly conforms to the river network (though this requires a catchment-specific definition of "river," that is, also consistent across the scales of interest). This would introduce some physical constraint to an otherwise arbitrary TI classification, where a value for n_c is subjectively assigned (as is the common-practice, and has been the case in our study).

Put together in more general terms, our results suggest that: (a) calibrating a model with an arbitrary TI classification, would likely make its parameter distributions inapplicable under alternative TI classification scenarios, even for the same catchment; and (b) if TI classification is held constant, and as long as smaller/equal iso-basins are used, parameter distributions are likely to remain independent of discretization.

Of course, these results are obtained for a constant catchment topography. However, it is unlikely that the discretization issues discussed here, which severely affect model parameters in the simplest case (constant topography), will be less important under changing topography; for example, when upscaling field-scale observations of land-use change to catchment scale. There, the main concern would be whether parameters calibrated to a field-scale experimental catchment could be applied at a larger scale. Our results suggest that, at least a good starting point would be to use iso-basins of the size $\leq$ the experimental catchment where the land-use change parameters are obtained via calibration; and also to ensure continuity of TI classification across the two scales (field and catchment scale).

One possible way to do so could be a top-down approach, that is, by first defining a global TI map at the largest scale of interest, from which the field-scale TI map is then carved out and used to obtain land-use change parameters (although this would be more difficult, but not impossible, where the experimental site is not part of the larger catchment of interest). This would enforce some continuity into the TI map going from field scale to catchment scale (i.e., "saturation tendency" values assigned at field-scale will continue to mean the same when put back into the catchment wide TI map). However, whether this will alleviate some of the parameter transferability issues across scales requires more rigorous assessments using GMD-TOPMODEL, which is the focus of ongoing research.

7.3. GMD-TOPMODEL's Runtime

Finally, for comparison of GMD-TOPMODEL's runtime with that of Metcalfe et al.'s (2015) Dynamic-TOPMODEL, upon which our model is built, Astagneau et al. (2021) reports Dynamic-TOPMODEL to have average run-times of around 40 s, for a 10-year record at daily timesteps (i.e., 3,650 steps), but using only 3 HSUs. They used a PC with 1.80 GHz CPU speed. For the same number of HSUs and length of record, GMD-TOPMODEL runs in under 4s, but using a PC with 3.1 GHz CPU; this translates into a roughly five fold improvement in runtime due to adaptive timestepping. We acknowledge that we are comparing the two models under different setups (catchments/HSUs), but these models' runtimes are predominantly controlled by the record-length and the total number of HSUs, rather than the catchment to which they are being applied, and we give context on both of those quantities when we make our comparison. However, the five-fold improvement that we report should be considered a guide value.

Such runtime improvements allow the model to be run for considerably more HSUs (e.g., 69 HSUs 23 times that of Astagneau et al., 2021), and/or longer records (12,960 timesteps around three times that of Astagneau et al., 2021), but to retain comparable runtimes (37.5s for GMD-TOPMODEL vs. 40s for Dynamic-TOPMODEL Astagneau et al., 2021).

8. Conclusions

Motivated by a current lack of adaptive timestep RR models, despite the known problems of fixed timestep schemes, we have shown how model equations can be reformulated in a constraint-handling form, which is necessary in order to use the existing packaged adaptive solvers, such as MATLAB ODE Suite. The same approach can be taken for other RR models, and using other available ODE solver packages. Under such formulation, model state variables and fluxes, as well as the dependencies of other variables on them, are clearly defined, making it easier to add/remove variables/fluxes depending on the application. Not only that, but adaptive schemes tend to be much faster to run. Our results suggest at least a five-fold improvement in runtimes compared to a fixed timestep setup.

In developing GMD-TOPMODEL, from its predecessor Dynamic-TOPMODEL, we have applied a series of modifications, which based on first principles, will generalize our model making it applicable to a wider range of landscapes than before. In particular, relaxing water-table parallelism to the ground surface, better slope treatment, and inclusion of diffusion, will make GMD-TOPMODEL applicable to deeper soils, and to both gentler and steeper catchments (relative to Dynamic-TOPMODEL). Also, the additional iso-basin discretization layer devised on top of the TI discretization, allows more flexibly constraining spatial information, which would be necessary in upscaling studies, or when modeling spatially distributed interventions, such as land-cover change. Note that in such cases it is important to tailor iso-basin discretization to the intended application (e.g., to capture different land cover types), rather than purely as limiting water movement in the catchment (as has been the case in this paper), because iso-basins can do both simultaneously.

For a 9 km^2 peatland catchment, we calibrated the model within a GLUE framework with LOA defined priori to running simulations. We then tested the sensitivity of GMD-TOPMODEL to TI discretization and iso-basins. Our results suggest that calibrated parameters are highly sensitive to discretization, in particular with respect to TI, even for the same catchment. Care is therefore required when porting calibrated parameters between sites, for example, upscaling studies; from our study we infer that, in such cases, parameters will be more readily transferable between sites where (a) the continuity of TI classification is preserved across the intended scales (e.g., by taking a top-down approach to TI classification); and (b) iso-basin sizes are restricted to the size of the smallest catchment that is being upscaled.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The DEM, rainfall and discharge data, together with the MATLAB source code of GMD-TOPMODEL are available in Goudarzi (2022).

Acknowledgments
This research was conducted as part of the NERC PROTECT-NFM Project (NE/R004587). Special thanks to Emma Shuttleworth, Tim Allott, and Martin Evans (at University of Manchester) and Mike Pilkington and Tom Spencer (at Moors for the Future Partnership) for their efforts in collecting, curating and checking the rainfall, stage and discharge data used here.

References

Allott, T. E. H., Evans, M. G., Lindsay, J. B., Agnew, C., Freer, J. E., Jones, A., & Parnell, M. (2009). Water tables in Peak District blanket peatlands (Moors for the Future Reports).

Arenas-Bautista, M. C., Arboleda-Obando, P. F., Duque-Gardeazabal, N., Saavedra-Cifuentes, E., & Donado, L. D. (2018). Hydrological modeling in tropical regions via TopModel. Study case: Central sector of the Middle Magdalena Valley-Colombia. *Preprints*, 2018070210. https://doi.org/10.20944/preprints201807.0210.v1

Astagneau, P. C., Thirel, G., Delaigue, O., Guillaume, J. H., Parajka, J., Brauer, C. C., et al. (2021). Hydrology modelling R packages: A unified analysis of models and practicalities from a user perspective. *Hydrology and Earth System Sciences*, *25*(7), 3937–3973.

Beven, K. (2016). Facets of uncertainty: Epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, *61*(9), 1652–1665. https://doi.org/10.1080/02626667.2015.1031761

Beven, K., & Binley, A. (1992). The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, *6*(3), 279–298. https://doi.org/10.1002/hyp.3360060305

Beven, K., & Freer, J. (2001). A dynamic TOPMODEL. *Hydrological Processes*, *15*(10), 1993–2011. https://doi.org/10.1002/hyp.252

Beven, K. J. (2011). *Rainfall-runoff modelling: The primer* (2nd ed.). John Wiley & Sons.

Beven, K. J. (2018). On hypothesis testing in hydrology: Why falsification of models is still a really good idea. *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1278. https://doi.org/10.1002/wat2.1278

Beven, K. J., & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal*, *24*(1), 43–69. https://doi.org/10.1080/02626667909491834

Beven, K. J., Kirkby, M. J., Freer, J. E., & Lamb, R. (2021). A history of TOPMODEL. *Hydrology and Earth System Sciences*, *25*(2), 527–549. https://doi.org/10.5194/hess-25-527-2021

Boll, J., Brooks, E. S., Campbell, C. R., Stockle, C. O., Young, S. K., Hammel, J. E., & McDaniel, P. A. (1998). Progress toward development of a GIS based water quality management tool for small rural watersheds: Modification and application of a distributed model. In *ASAE Annual International Meeting in Orlando, July* (pp. 12–16).

Borga, M., Dalla Fontana, G., & Cazorzi, F. (2002). Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index. *Journal of Hydrology*, *268*(1–4), 56–71. https://doi.org/10.1016/s0022-1694(02)00118-x

Childs, E. C. (1971). Drainage of groundwater resting on a sloping bed. *Water Resources Research*, *7*(5), 1256–1263. https://doi.org/10.1029/wr007i005p01256

Chirico, G. B., Grayson, R. B., & Western, A. W. (2003). On the computation of the quasi-dynamic wetness index with multiple-flow-direction algorithms. *Water Resources Research*, *39*(5), 1115. https://doi.org/10.1029/2002wr001754

Clark, M. P., & Kavetski, D. (2010). Ancient numerical daemons of conceptual hydrological modeling: 1. Fidelity and efficiency of timestepping schemes. *Water Resources Research*, *46*(10), W10510. https://doi.org/10.1029/2009wr008894

Clark, M. P., Kavetski, D., & Fenicia, F. (2011). Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research*, *47*(9), W09301. https://doi.org/10.1029/2010wr009827

Coxon, G., Freer, J., Lane, R., Dunne, T., Knoben, W. J., Howden, N. J., et al. (2019). DECIPHeR v1: Dynamic fluxEs and connectivity for predictions of HydRology. *Geoscientific Model Development*, *12*(6), 2285–2306. https://doi.org/10.5194/gmd-12-2285-2019

Daniels, S. M., Agnew, C. T., Allott, T. E. H., & Evans, M. G. (2008). Water-table variability and runoff generation in an eroded peatland, South Pennines, UK. *Journal of Hydrology*, *361*(1–2), 214–226. https://doi.org/10.1016/j.jhydrol.2008.07.042

Duan, J., & Miller, N. L. (1997). A generalized power function for the subsurface transmissivity profile in TOPMODEL. *Water Resources Research*, *33*(11), 2559–2562. https://doi.org/10.1029/97wr02186

Evans, M. G., Burt, T. P., Holden, J., & Adamson, J. (1999). Runoff generation and water table fluctuations in blanket peat: Evidence from UK data spanning the dry summer of 1995. *Journal of Hydrology*, *221*(3–4), 141–160. https://doi.org/10.1016/s0022-1694(99)00085-2

Frankenberger, J. R., Brooks, E. S., Walter, M. T., Walter, M. F., & Steenhuis, T. S. (1999). A GIS-based variable source area hydrology model. *Hydrological Processes*, *13*(6), 805–822. https://doi.org/10.1002/(sici)1099-1085(19990430)13:6<805::aid-hyp754>3.0.co;2-m

Freer, J. E., McMillan, H., McDonnell, J. J., & Beven, K. J. (2004). Constraining dynamic TOPMODEL responses for imprecise water-table information using fuzzy rule based performance measures. *Journal of Hydrology*, *291*(3–4), 254–277. https://doi.org/10.1016/j.jhydrol.2003.12.037

Freeze, R. A., & Harlan, R. L. (1969). Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, *9*(3), 237–258. https://doi.org/10.1016/0022-1694(69)90020-1

Fu, X., Luo, L., Pan, M., Yu, Z., Tang, Y., & Ding, Y. (2018). Evaluation of TOPMODEL-based land surface–atmosphere transfer scheme (TOPLATS) through a soil moisture simulation. *Earth Interactions*, *22*(15), 1–19. https://doi.org/10.1175/ei-d-17-0037.1

Gao, J., Holden, J., & Kirkby, M. (2015). A distributed TOPMODEL for modelling impacts of land-cover change on river flow in upland peatland catchments. *Hydrological Processes*, *29*(13), 2867–2879. https://doi.org/10.1002/hyp.10408

Gao, J., Kirkby, M., & Holden, J. (2018). The effect of interactions between rainfall patterns and land-cover change on flood peaks in upland peatlands. *Journal of Hydrology*, *567*, 546–559. https://doi.org/10.1016/j.jhydrol.2018.10.039

Gil, E. G., & Tobón, C. (2016). Hydrological modelling with TOPMODEL of Chingaza páramo, Colombia. *Revista Facultad Nacional de Agronomía Medellín*, *69*(2), 7919–7933. https://doi.org/10.15446/rfna.v69n2.59137

Goudarzi, S., (2022). Generalised multistep dynamic TOPMODEL in MATLAB [Software and Data]. Retrieved from https://zenodo.org/badge/latestdoi/514527559

Goudarzi, S., Milledge, D. G., Holden, J., Evans, M. G., Allott, T. E., Shuttleworth, E. L., et al. (2021). Blanket peat restoration: Numerical study of the underlying processes delivering natural flood management benefits. *Water Resources Research*, *57*(4), e2020WR029209. https://doi.org/10.1029/2020wr029209

Harman, C., & Sivapalan, M. (2009). A similarity framework to assess controls on shallow subsurface flow dynamics in hillslopes. *Water Resources Research*, *45*(1), W01417. https://doi.org/10.1029/2008wr007067

Harman, C. J., Reeves, D. M., Baeumer, B., & Sivapalan, M. (2010). A subordinated kinematic wave equation for heavy-tailed flow responses from heterogeneous hillslopes. *Journal of Geophysical Research*, *115*(F1), F00A08. https://doi.org/10.1029/2009jf001273

Hjerdt, K. N., McDonnell, J. J., Seibert, J., & Rodhe, A. (2004). A new topographic index to quantify downslope controls on local drainage. *Water Resources Research*, *40*(5), W05602. https://doi.org/10.1029/2004wr003130

Holden, J. (2009). Flow through macropores of different size classes in blanket peat. *Journal of Hydrology*, *364*(3–4), 342–348. https://doi.org/10.1016/j.jhydrol.2008.11.010

Holden, J., & Burt, T. P. (2002). Infiltration, runoff and sediment production in blanket peat catchments: Implications of field rainfall simulation experiments. *Hydrological Processes*, *16*(13), 2537–2557. https://doi.org/10.1002/hyp.1014

Holden, J., & Burt, T. P. (2003a). Hydraulic conductivity in upland blanket peat; measurement and variability. *Hydrological Processes*, *17*(6), 1227–1237. https://doi.org/10.1002/hyp.1182

Holden, J., & Burt, T. P. (2003b). Runoff production in blanket peat covered catchments. *Water Resources Research*, *39*(7), 1191. https://doi.org/10.1029/2002wr001956

Iorgulescu, I., & Musy, A. (1997). Generalization of TOPMODEL for a power law transmissivity profile. *Hydrological Processes*, *11*(9), 1353–1355. https://doi.org/10.1002/(sici)1099-1085(199707)11:9<1353::aid-hyp585>3.0.co;2-u

Jeziorska, J., & Niedzielski, T. (2018). Applicability of TOPMODEL in the mountainous catchments in the upper Nysa Kłodzka River basin (SW Poland). *Acta Geophysica*, *66*(2), 203–222. https://doi.org/10.1007/s11600-018-0121-6

Kahaner, D., Moler, C., & Nash, S. (1989). *Numerical methods and software*. Prentice Hall.

Kavetski, D., & Clark, M. P. (2010). Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, *46*(10), W10511. https://doi.org/10.1029/2009wr008896

Kavetski, D., Fenicia, F., & Clark, M. P. (2011). Impact of temporal data resolution on parameter inference and model identification in conceptual hydrological modeling: Insights from an experimental catchment. *Water Resources Research*, *47*(5), W05501. https://doi.org/10.1029/2010wr009525

Kavetski, D., & Kuczera, G. (2007). Model smoothing strategies to remove microscale discontinuities and spurious secondary optima in objective functions in hydrological calibration. *Water Resources Research*, *43*(3), W03411. https://doi.org/10.1029/2006wr005195

Kavetski, D., Kuczera, G., & Franks, S. W. (2003). Semidistributed hydrological modeling: A "saturation path" perspective on TOPMODEL and VIC. *Water Resources Research*, *39*(9), 1246. https://doi.org/10.1029/2003wr002122

Keeland, B. D., Dowd, J. F., & Hardegree, W. S. (1997). Use of inexpensive pressure transducers for measuring water levels in wells. *Wetlands Ecology and Management*, *5*(2), 121–129. https://doi.org/10.1023/a:1008203625917

Khatami, S., Peel, M. C., Peterson, T. J., & Western, A. W. (2019). Equifinality and flux mapping: A new approach to model evaluation and process representation under uncertainty. *Water Resources Research*, *55*(11), 8922–8941. https://doi.org/10.1029/2018wr023750

Kirkby, M. (1975). Hydrograph modelling strategies. In *Peel R., Chisholm M., Haggert P., Processes in physical and human geography* (pp. 69–90). Heineman.

Kirkby, M. J. (1997). TOPMODEL: A personal view. *Hydrological Processes*, *11*(9), 1087–1097. https://doi.org/10.1002/(sici)1099-1085(199707)11:9<1087::aid-hyp546>3.0.co;2-p

La Follette, P. T., Teuling, A. J., Addor, N., Clark, M., Jansen, K., & Melsen, L. A. (2021). Numerical daemons of hydrological models are summoned by extreme precipitation. *Hydrology and Earth System Sciences*, *25*(10), 5425–5446. https://doi.org/10.5194/hess-25-5425-2021

Lane, S. N., & Milledge, D. G. (2013). Impacts of upland open drains upon runoff generation: A numerical assessment of catchment-scale impacts. *Hydrological Processes*, *27*(12), 1701–1726. https://doi.org/10.1002/hyp.9285

Lanni, C., McDonnell, J. J., & Rigon, R. (2011). On the relative role of upslope and downslope topography for describing water flow path and storage dynamics: A theoretical analysis. *Hydrological Processes*, *25*(25), 3909–3923. https://doi.org/10.1002/hyp.8263

Li, W., Lin, K., Zhao, T., Lan, T., Chen, X., Du, H., & Chen, H. (2019). Risk assessment and sensitivity analysis of flash floods in ungauged basins using coupled hydrologic and hydrodynamic models. *Journal of Hydrology*, *572*, 108–120. https://doi.org/10.1016/j.jhydrol.2019.03.002

Lighthill, M. J., & Whitham, G. B. (1955). On kinematic waves I. Flood movement in long rivers. *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences*, *229*(1178), 281–316.

Lindsay, J. B. (2018). *WhiteboxTools user manual* (p. 20). Geomorphometry and Hydrogeomatics Research Group, University of Guelph.

Loague, K. (2010). *Rainfall-runoff modelling: Selection, introduction and commentary*. IAHS Press.

McMillan, H., Krueger, T., & Freer, J. (2012). Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality. *Hydrological Processes*, *26*(26), 4078–4111. https://doi.org/10.1002/hyp.9384

McMillan, H. K., & Westerberg, I. K. (2015). Rating curve estimation under epistemic uncertainty. *Hydrological Processes*, *29*(7), 1873–1882. https://doi.org/10.1002/hyp.10419

Metcalfe, P., Beven, K., & Freer, J. (2015). Dynamic TOPMODEL: A new implementation in R and its sensitivity to time and space steps. *Environmental Modelling & Software*, *72*, 155–172. https://doi.org/10.1016/j.envsoft.2015.06.010

Montgomery, D. R., & Dietrich, W. E. (1994). A physically based model for the topographic control on shallow landsliding. *Water Resources Research*, *30*(4), 1153–1171. https://doi.org/10.1029/93wr02979

Montgomery, D. R., & Dietrich, W. E. (2002). Runoff generation in a steep, soil-mantled landscape. *Water Resources Research*, *38*(9), 7–1. https://doi.org/10.1029/2001wr000822

Mukae, K., Miwa, K., Okazawa, H., Yamamoto, T., & Inoue, T. (2018). Application of TOPMODELs for assessment of ecosystem services: Regulating service in agricultural and forest watershed. *International Journal of Environmental and Rural Development*, *9*(1), 43–50. https://doi.org/10.32115/ijerd.9.1_43

Page, T., Beven, K. J., Freer, J., & Neal, C. (2007). Modelling the chloride signal at Plynlimon, Wales, using a modified dynamic TOPMODEL incorporating conservative chemical mixing (with uncertainty). *Hydrological Processes: International Journal*, *21*(3), 292–307. https://doi.org/10.1002/hyp.6186

Park, D., Fan, H. J., Zhu, J. J., Oh, S. E., Um, M. J., & Jung, K. (2019). Evaluation of reliable digital elevation model resolution for TOPMODEL in two mountainous watersheds, South Korea. *Applied Sciences*, *9*(18), 3690. https://doi.org/10.3390/app9183690

Pawson, R. R., Evans, M. G., & Allott, T. E. H. A. (2012). Fluvial carbon flux from headwater peatland streams: Significance of particulate carbon flux. *Earth Surface Processes and Landforms*, *37*(11), 1203–1212. https://doi.org/10.1002/esp.3257

Perrin, C., Michel, C., & Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, *242*(3–4), 275–301. https://doi.org/10.1016/s0022-1694(00)00393-0

Petersen-Øverleir, A., & Reitan, T. (2005). Uncertainty in flood discharges from urban and small rural catchments due to inaccurate head measurement. *Hydrology Research*, *36*(3), 245–257. https://doi.org/10.2166/nh.2005.0018

Petzold, L. (1983). Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations. *SIAM Journal on Scientific and Statistical Computing*, *4*(1), 136–148. https://doi.org/10.1137/0904010

Ponce, V. M. (1989). *Engineering hydrology: Principles and practices* (Vol. 640). Prentice Hall.

Quinn, P. F. B. J., Beven, K., Chevallier, P., & Planchon, O. (1991). The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, *5*(1), 59–79. https://doi.org/10.1002/hyp.3360050106

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*(5), W05521. https://doi.org/10.1029/2009wr008328

Rezanezhad, F., Price, J. S., Quinton, W. L., Lennartz, B., Milojevic, T., & Van Cappellen, P. (2016). Structure of peat soils and implications for water storage, flow and solute transport: A review update for geochemists. *Chemical Geology*, *429*, 75–84. https://doi.org/10.1016/j.chemgeo.2016.03.010

Rogelis, M. C., Werner, M., Obregón, N., & Wright, N. (2016). Hydrological model assessment for flood early warning in a tropical high mountain basin. *Hydrology and Earth System Sciences Discussions*, 1–36. https://doi.org/10.5194/hess-2016-30

Schoups, G., Vrugt, J. A., Fenicia, F., & Van De Giesen, N. C. (2010). Corruption of accuracy and efficiency of Markov chain Monte Carlo simulation by inaccurate numerical implementation of conceptual hydrologic models. *Water Resources Research*, *46*(10), W10530. https://doi.org/10.1029/2009wr008648

Schwanghart, W., & Scherler, D. (2014). TopoToolbox 2–MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surface Dynamics*, *2*(1), 1–7. https://doi.org/10.5194/esurf-2-1-2014

Shampine, L. F. (1994). *Numerical solution of ordinary differential equations*. Chapman & Hall.

Shampine, L. F., & Reichelt, M. W. (1997). The MATLAB ode suite. *SIAM Journal on Scientific Computing*, *18*(1), 1–22. https://doi.org/10.1137/s1064827594276424

Shampine, L. F., & Thompson, S. (2001). Solving odes in MATLAB. *Applied Numerical Mathematics*, *37*(4), 441–458. https://doi.org/10.1016/s0168-9274(00)00055-6

Smith, R. N. B., Blyth, E. M., Finch, J. W., Goodchild, S., Hall, R. L., & Madry, S. (2006). Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system. *Meteorological Applications: A Journal of Forecasting, Practical Applications, Training Techniques and Modelling*, *13*(2), 89–109. https://doi.org/10.1017/s1350482705002069

Speight, J. G. (1980). The role of topography in controlling throughflow generation: A discussion. *Earth Surface Processes*, *5*(2), 187–191. https://doi.org/10.1002/esp.3760050209

Takeuchi, K., Ao, T., & Ishidaira, H. (1999). Introduction of block-wise use of TOPMODEL and Muskingum-Cunge method for the hydroenvironmental simulation of a large ungauged basin. *Hydrological Sciences Journal*, *44*(4), 633–646. https://doi.org/10.1080/02626669909492258

Van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, *44*(5), 892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x

Walter, M. T., Steenhuis, T. S., Mehta, V. K., Thongs, D., Zion, M., & Schneiderman, E. (2002). Refined conceptualization of TOPMODEL for shallow subsurface flows. *Hydrological Processes*, *16*(10), 2041–2046. https://doi.org/10.1002/hyp.5030

Wang, J., Zhang, J., Wang, G., Song, X., Yang, X., & Wang, Y. (2020). Ensemble flood simulation for the typical catchment in humid climatic zone by using multiple hydrological models. *Proceedings of the International Association of Hydrological Sciences*, *383*, 213–222. https://doi.org/10.5194/piahs-383-213-2020

Wang, L., Koike, T., Yang, D., & Yang, K. (2009). Improving the hydrology of the Simple Biosphere Model 2 and its evaluation within the framework of a distributed hydrological model. *Hydrological Sciences Journal*, *54*(6), 989–1006. https://doi.org/10.1623/hysj.54.6.989

Xue, L., Yang, F., Yang, C., Wei, G., Li, W., & He, X. (2018). Hydrological simulation and uncertainty analysis using the improved TOPMODEL in the arid Manas River basin, China. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-017-18982-8

Yang, D., Herath, S., & Musiake, K. (2000). Comparison of different distributed hydrological models for characterization of catchment spatial variability. *Hydrological Processes*, *14*(3), 403–416. https://doi.org/10.1002/(sici)1099-1085(20000228)14:3<403::aid-hyp945>3.0.co;2-3

Zhang, Z., Zimmermann, N. E., Kaplan, J. O., & Poulter, B. (2016). Modeling spatiotemporal dynamics of global wetlands: Comprehensive evaluation of a new sub-grid TOPMODEL parameterization and uncertainties. *Biogeosciences*, *13*(5), 1387–1408. https://doi.org/10.5194/bg-13-1387-2016