PHyL v1.0

A parallel, flexible, and advanced software for hydrological and slope stability modeling at a regional scale

Chen, Guoding; Zhang, Ke; Wang, Sheng; Jia, Tianlong

DOI
10.1016/j.envsoft.2023.105882

Publication date
2023

Document Version
Final published version

Published in
Environmental Modelling and Software

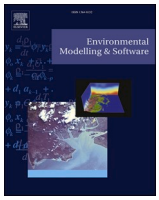
Citation (APA)
Chen, G., Zhang, K., Wang, S., & Jia, T. (2023). PHyL v1.0: A parallel, flexible, and advanced software for hydrological and slope stability modeling at a regional scale. *Environmental Modelling and Software*, *172*, Article 105882. https://doi.org/10.1016/j.envsoft.2023.105882

Environmental Modelling and Software 172 (2024) 105882

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft



PHyL v1.0: A parallel, flexible, and advanced software for hydrological and slope stability modeling at a regional scale

Guoding Chen [a,b], Ke Zhang [a,c,d,*], Sheng Wang [a,c], Tianlong Jia [b]

[a] The National Key Laboratory of Water Disaster Prevention, College of Hydrology and Water Resources, And Yangtze Institute for Conservation and Development, Hohai University, Nanjing, Jiangsu, 210024, China
[b] Delft University of Technology, Faculty of Civil Engineering and Geosciences, Department of Water Management, Stevinweg 1, 2628 CN, Delft, the Netherlands
[c] China Meteorological Administration Hydro-Meteorology Key Laboratory, Hohai University, Nanjing, Jiangsu, 210024, China
[d] Key Laboratory of Water Big Data Technology of Ministry of Water Resources, Hohai University, Nanjing, Jiangsu, 210024, China

ARTICLE INFO

Handling Editor: Daniel P Ames

Keywords:
Parallel computation
Hydrological-geotechnical models
3D slope stability model
Flood-landslide events

ABSTRACT

Physically-based hydrological-geotechnical modeling at large scales is difficult, especially due to the time-consuming nature of flow routing and 3D soil stability models. Although parallelization techniques are commonly used for each model individually, there is currently no concurrent parallelization strategy for both. This study proposed an open-source, **P**arallelized, and modular modeling software for regional **Hy**drologic processes and **L**andslides simulation and prediction (PHyL v1.0). It offers parallel computation in both hydrological and 3D slope stability modules, cross-scale modeling ability via a soil moisture downscaling method, and advanced input/output (I/O) and post-processing visualization. Additionally, PHyL v1.0 is flexible and extensible, making it compatible with all mainstream operating systems. We applied PHyL v1.0 in the Yuehe River Basin, where the computational efficiencies, parallel performance, parameter sensitivity analysis, and predictive capabilities were evaluated. The PHyL v1.0 is therefore appropriately used as an advanced software for high-resolution and complex simulations of regional floods and landslides.

1. Introduction

Rainfall-induced landslides are ubiquitous natural hazards that pose significant threats to human lives and infrastructure in hillslope environment (Hong et al., 2006; Wang et al., 2020; Zhang et al., 2016). The variations in soil strength are predominantly shaped by soil water content and pore water pressure (Lu et al., 2010). Thus, the triggering of slope instability is a multifaceted process influenced by the dynamic interplay between hydrological processes and the mechanical response of soil to hydrological loading (Fan et al., 2016; Lu and Godt, 2013). Moreover, the rainstorm, flood and landslide frequently occur in a cascading fashion, whereby a seemingly minor event could instigate a severe flood and/or landslide that poses a substantial threat to a community that is impacted (Zhang et al., 2016). The pressing demand for the predicting and evaluating hazards has spurred the creation of modeling tools that are explicated in a physically- and process-based approach, such as TRIGRS (Alvioli and Baum, 2016; Baum et al., 2008), GEOtop-FS (Simoni et al., 2008), iCRESTRIGRS (Zhang et al., 2016), and FSLAM (Guo et al., 2022). The integration of hydrological models of varying complexity within the framework of slope stability analysis has resulted in the generation of hydrological patterns and landslide susceptibility assessments. However, the computational efficiency of these models is highly dependent on the number of basic units, such as regular grids or triangulated irregular networks (Ivanov et al., 2004), as well as the model structures employed. This limitation presents a significant challenge in advancing the precision, scalability, and complexity of these models, including the incorporation of 3D slope geometry (Mergili et al., 2014b; Xie et al., 2006).

The swift advancement of computer science has led to the adoption of parallelization as a viable technique for enhancing computing efficiency, as opposed to serial computing (Asgari et al., 2022). A crucial method for parallel computing involves the division of a complex computational task into multiple independent loads, which can be allocated to several processors concurrently. In the context of watershed modeling, parallel simulation frequently involves the partitioning of the study area into multiple independent sub-basins. Each sub-basin is then

* Corresponding author. The National Key Laboratory of Water Disaster Prevention, College of Hydrology and Water Resources, And Yangtze Institute for Conservation and Development, Hohai University, Nanjing, Jiangsu, 210024, China.
E-mail address: kzhang@hhu.edu.cn (K. Zhang).

https://doi.org/10.1016/j.envsoft.2023.105882
Received 23 July 2023; Received in revised form 1 November 2023; Accepted 13 November 2023
Available online 2 December 2023
1364-8152/

regarded as a distinct task and processed by a separate thread, with the underlying assumption that inter-sub-basin communication is either absent or minimal (Vivoni et al., 2011). For example, Li et al. (2011) devised a dynamic parallel algorithm that facilitates the dynamic dispatching of subbasins to computing processes. Vivoni et al. (2011) utilized the channel network as a directed graph to partition subbasins. In addition to sub-basin decomposition, Liu et al. (2014) parallelized the distributed models at basic simulation-unit levels by employing a layered approach. Moreover, several advanced tools have been proposed to achieve the theoretical maximum speedup ratio (Liu et al., 2013; Wang et al., 2012; Xu et al., 2021). The thriving development of parallel computing in hydrology provides a foundation for more intricate and extensive simulation frameworks.

Despite recent advances in parallel computing techniques, the application of parallel design in slope stability models is not yet commonplace. Most hydrological-geotechnical frameworks utilize the infinite slope stability model (i.e., 1D model) as a submodule for landslide prediction (An et al., 2016; Aristizábal et al., 2016; He et al., 2016; Velásquez et al., 2020), and although these models do not incorporate parallel computing, they run efficiently at a regional scale since the 1D model is low-cost inherently. For larger scales, several successful cases employed the soil downscaling method to bridge the submodules (Leonarduzzi et al., 2021a, 2021b; Wang et al., 2020), thus avoiding the need for high-resolution hydrological modeling. To achieve higher efficiency without compromise, a small number of models have been proposed for parallel computing in the past decade. For instance, HIRESSS was provided to parallelize the Monte Carlo simulation to manage the parameters uncertainty (Rossi et al., 2013), and Alvioli and Baum (2016) improved and parallelized the TRIGRS using the Message Passing Interface (MPI) framework. However, they employed a procedure- and/or loop-oriented partitioning approach and do not utilize spatial domain decomposition. Mergili et al. (2014b) proposed r.slope.stability, a parallel 3D slope stability model aimed at reducing computation time for regional applications. Building upon this work, Chen et al. (2023) introduced enhancements by integrating a hydrological modeling core to enable monitoring of hydro-mechanical triggering. However, the increasing complexity of the model, diverse application scenarios, and input/output (I/O) processing constraints impose challenges on the overall model efficiency, as each submodule is subject to the "Bucket Effect". To date, a comprehensive parallel hydrological-geotechnical framework that enables parallelization for both hydrological and landslide modeling aspects remains elusive. To the best of our knowledge, this study presents the first such framework.

To fill this gap, this paper presents a new modeling software, PHyL v1.0. The model accelerates the hydrological-geotechnical modeling by embedding the parallel algorithms for both submodules. PHyL v1.0 has a modular structure that is both flexible and extensible, allowing for compatibility with all mainstream operating systems. The general structure of the PHyL v1.0 was introduced in Section 2, including basic submodules, parallel algorithms, I/O and visualization system. A case study was shown in Section 3 to evaluate the computational efficiencies, parallel performance, parameter sensitivity analysis, and the predictive capabilities of the floods and landslides, as presented in Section 4. We then discussed the results in Section 5 and drew conclusions in Section 6.

2. Model description

2.1. Software overview

PHyL v1.0 is an open-source software designed to model regional hydrologic processes and predict landslides. It is a modular software that is capable of parallel processing, and is an updated version of its predecessor, iHydroSlide3D v1.0 (Chen et al., 2023). The name of the software was also changed because the presented version boasts major modifications in its theoretical framework, structural organization, and utilization. These improvements include (1) changes to the hydrological routing scheme, (2) adjustments to programming languages and architectures, (3) the introduction of parallel computation, and (4) the incorporation of auxiliary runtime libraries. PHyL v1.0 is developed using Fortran, Python, and CMake, making it cross-platform compatible (Fig. 1). Further details regarding these features are provided below.

2.2. Model theory

PHyL v1.0 inherits the theoretical fundamentals from iHydroSlide3D v1.0 and comprises (i) a distributed hydrological model; (ii) a regional 3D slope stability model; and (iii) a soil moisture downscaling (SMD) method (Fig. 2). The software facilitates coupled hydrological and geotechnical modeling, with the SMD method enabling cross-scale running between the submodules.

2.2.1. CREST model

A physically-based hydrological model, the Coupled Routing and Excess STorage (CREST), was used as the hydrological core in PHyL v1.0 (Fig. 2a). The CREST model was jointly developed by NASA SERVIR project team (Wang et al., 2011) and University of Oklahoma (https://hydro.ou.edu, last access: December 23, 2014). It has been used as a flood detection toolset across the globe (Wu et al., 2012) and for operational near-real-time forecasting of flash floods in the United States and its territories (Gourley et al., 2017). Its flexible framework and data compatibility allow CREST to undergo continuous development, forming advanced models like CREST-iMAP (Li et al., 2021) and EF5 (Flamig et al., 2020a). CREST is also suitable for interdisciplinary projects, e.g., undertaking the simulation of the infiltration and soil water in landslide prediction (He et al., 2016; Zhang et al., 2016) and coupling with multiple remote sensing modules (Huang et al., 2022). Hydrological processes forced by gridded atmospheric data (P and ET in Fig. 2a) that are firstly partitioned into surface runoff and subsurface flow based on the surface characteristics and prior states:

$$= \begin{cases} W_m - W, i + P_{\text{soil}} \geq i_m \\ (W_m - W) + W_m \cdot \left[1 - \dfrac{i + P_{\text{soil}}}{i_m}\right]^{1+b}, i + P_{\text{soil}} < i_m \end{cases}, \quad (1)$$

$$R = P_{\text{soil}} - I, \quad (2)$$

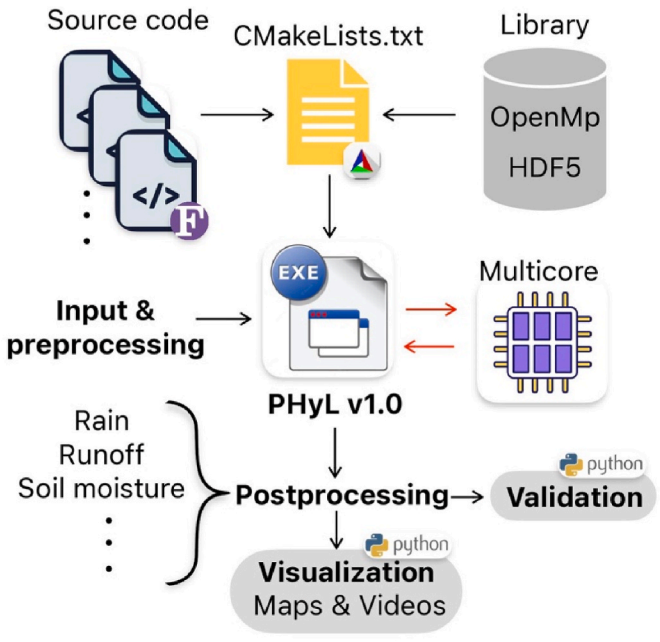


Fig. 1. Overview software framework of the PHyL v1.0.

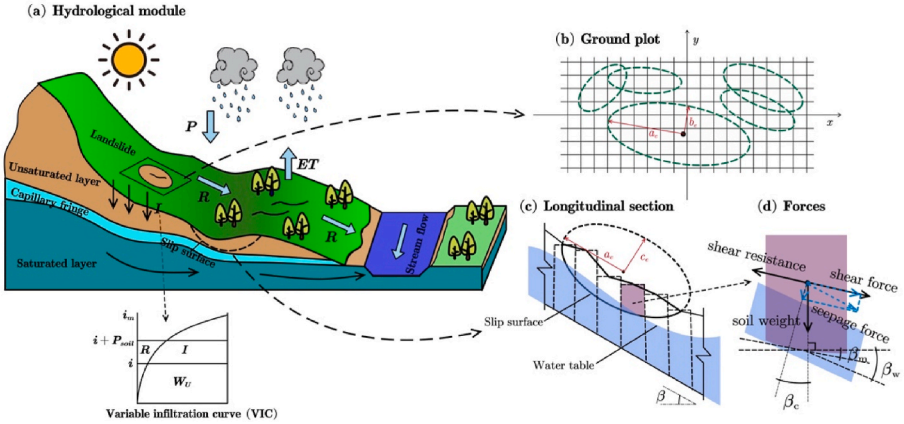


Fig. 2. Theoretical framework of the PHyL v1.0: (a) basic hydrological processes of the CREST and potential landslides at a hillslope; (b) random ellipsoid landslides within the GIS coordinate system; (c) features of a potential failure under a longitudinal section; (d) forces acting at a soil column considering the groundwater table.

where P_{soil} is the water reaching the soil surface; I is the infiltration water; R is the excess rain; W denote the water amount of a cell; W_m lies between the field capacity and the wilting point, and its distribution is determined by both topography and soil texture (Chen et al., 2023). This capability in PHyL v1.0 enables the consideration of soil properties and the thickness of the aeration zone in hydrological modeling. A relationship (i.e., the variable infiltration curve in Fig. 2a) (Ren-Jun 1992) between the infiltration capacity (i) and its maximum (i_m) of the soil was used to further separate the subsurface flow:

$$i = i_m\left[1-(1-a)^{\frac{1}{b}}\right], \tag{3}$$

where a is a fractional number of a grid cell and b is an empirical shape parameter.

Compared with the quasi-distributed approach in previous version (Chen et al., 2023), PHyL v1.0 calculates the runoff following a fully distributed linear reservoir routing scheme (Shen et al., 2017). For a general cell:

$$\text{Runoff} = \frac{\left[R(\text{out})_S + R(\text{out})_I\right]A_g + \sum R(\text{via})_S A_g(\text{depart})_S + \sum R(\text{via})_I A_g(\text{depart})_I}{\Delta t}, \tag{4}$$

where $R(\text{out})_{S,I}$ and A_g is the outgoing flow and grid area; $R(\text{via})_{S,I}$ is the overland flow and interflow that pass through the given grid cell to any downstream receptor grid cells; $A_g(\text{depart})_{S,I}$ is grid areas of the donor cells of $R(\text{via})_{S,I}$. This enhancement significantly enhances the prediction performance of channel and river flow but also increases the computational workload. More detailed theoretical bases of the CREST are discussed in Xue et al. (2013) and Shen et al. (2017).

2.2.2. Regional 3D slope stability model

Incorporating CREST into soil slope stability involves monitoring both the water and mechanical states within the soil (Chen et al., 2023). PHyL v1.0 employs a three-dimensional (3D) simulation approach to model regional landslides. The model was modified from an open-source code, r.slope.stability (Mergili et al., 2014a), by coupling with complete hydrologic processes (Chen et al., 2023). Embedded in a geographic information system (GIS), the model randomly generates a large number of ellipsoidal or truncated slip surfaces (see Fig. 2b and c). These ellipsoidal slip surfaces are defined by the geometric conditions (Fig. 2) of the centre, the length of the three principal semiaxes (a_e, b_e, and c_e), the inclination β, and the aspect α. The potential landslides are characterized by the above features and lie between the terrain surface and the bottom of the ellipsoids. The values of a_e, b_e, and c_e, and the location of the centre are generated separately for each ellipsoid using a simple pseudo-random algorithm based on the user-defined restriction (maximum and minimum values in Table 1). Note that when landslide records are limited, PHyL v1.0 can adopt an alternative method to estimate landslide depth by utilizing the wetting front concept linked to the infiltration process (He et al., 2016; Zhang et al., 2016). The total number of the random ellipsoids is determined by a parameter D_e (discussed in following sections). The encompassed soil elements are treated as pore space buckets, where the soil states are described by the CREST module. Hydrological dynamics that change states (e.g., groundwater table and soil moisture) in the soil mantle thereby impacts the mechanical state of the hillslope. The factor of safety (F_s) is used as an indicator in PHyL v1.0, arriving at the ratio between resisting and driving forces:

$$F_s = \frac{\sum[c_s \cdot A + (G^{'}\cos\beta_c + N_s - U)\tan\varphi^{'}]\cos\beta_m}{\sum(G^{'}\sin\beta_m + T_s)\cos\beta_m}, \tag{5}$$

where c_s is the cohesion; A is the projected slip surface of the considered cell; $G^{'}$ is the soil weight; $\varphi^{'}$ is the internal friction angle; β_c is the inclination of the slip surface at the considered soil column; β_m is the apparent dip of the slip surface in the direction of α; U is the pore pressure acting on the slip surface of each soil column; N_s and T_s are the contributions of the seepage force to the normal force and the shear force. Inter-column forces and external forces, such as seismic loading, are not considered. Solving F_s in the context of a GIS-based system and a coupled hydro-stability framework involves auxiliary strategies, such as coordinate transformation and geometric principles, which have been discussed in detail by Mergili et al. (2014b) and Chen et al. (2023). The model also provides estimations of landslide occurrence probability (P_f), landslide area (A_L), and landslide volume (V_L) (readers may refer to (Chen et al., 2023) for more detailed discussion). The utilization of 3D slope stability modeling within PHyL v1.0 offers a notable advantage in accommodating the spatial and vertical heterogeneity of soil properties and moisture conditions, which are vital considerations in the assessment of potential landslides.

2.2.3. Soil moisture downscaling method

Soil moisture downscaling (SMD) method serves as a bridge in PHyL v1.0 to tackle the contradictory requirements for applied resolutions within the hydro-stability system. The resolution used in large-scale hydrological modeling (e.g., at a 90-m grid scale or coarser) normally cannot support the slope stability analysis (normally at a finer scale than 90 m) (Wang et al., 2020). The side slope could be the main error source in the landslide prediction (Leonarduzzi et al., 2021a), and the spatial distribution of soil moisture in subgrid variability in topography is necessary for F_s estimation (Chen et al., 2023). The SMD was built upon the concept of using Topographic Wetness Index (TWI) (Beven et al.,

Table 1
Information pertaining to parameters in PHyL v1.0.

Parameters	Description	Unit	Ranges	Used value
K_{sat}[b]	Soil saturated hydraulic conductivity	mm/h	/	Distributed
W_m[b]	Soil water storage capacity	mm	/	Distributed
B[a]	The exponent of the variable infiltration curve	–	[0.05, 1.5]	0.65
IM[b]	Impervious area ratio	–	/	Distributed
$coeM$[a]	The overland runoff velocity coefficient	–	[1, 150]	100
$expM$[a]	The overland flow speed exponent	–	[0.1, 0.55]	0.45
$coeR$[a]	The flow speed ratio of channel to overland	–	[1, 3]	2
$coeS$[a]	The flow speed ratio of interflow to overland	–	[0.01, 1]	0.3
KS[a]	Overland reservoir discharge parameter	–	[0.001, 1]	0.6
KI[a]	Interflow reservoir discharge parameter	–	[0.001, 1]	0.25
c_s[b]	Soil cohesion	kPa	/	Distributed
γ_s[b]	The unit weight of dry soil	kN/m^3	/	Distributed
φ[b]	The angle of internal friction	°	/	Distributed
θ_s[b]	Soil saturated moisture content	m^3/m^3	/	Distributed
θ_r[b]	Soil residual moisture content	m^3/m^3	/	Distributed
α[b]	Inverse of air entry pressure for water saturated soil	kPa^{-1}	/	Distributed
n[b]	Pore size distribution parameter	–	/	Distributed
D_e[a]	Density sampling for random ellipsoids (potential landslides)	–	[10, 1000]	500
$a_{e,min}$[b]	Minimum of the major semiaxis of a random ellipse	m	/	50
$a_{e,max}$[b]	Maximum of the major semiaxis of a random ellipse	m	/	100
$b_{e,min}$[b]	Minimum of the minor semiaxis of a random ellipse	m	/	30
$b_{e,max}$[b]	Maximum of the minor semiaxis of a random ellipse	m	/	60
$c_{e,min}$[b]	Minimum of the landslide depth	m	/	1
$c_{e,max}$[b]	Maximum of the landslide depth	m	/	4
N_{sub}	Number of the sub-basin in a drainage basin	–	/	4
N_{tile}	Number of tiles decomposing the whole area	–	/	144
N_{Hthread}	Number of parallel processes for the hydrological module	–	/	4
N_{Lthread}	Number of parallel processes for the slope stability module	–	/	48

[a] the calibrated parameters;
[b] The parameters are determined by the input datasets, with the distributed values indicating their spatial heterogeneity; the remaining parameters are user-defined values. The assigned uniform values result from the calibration procedure.

1995) to redistribute soil saturation in space. In this work, we followed the expression in Wang et al. (2020), and the general expression can be written as:

$$SM_{fine} = f(SM_{coarse}, TWI, K_w, K_a, p, q, A_s), \quad (6)$$

where SM_{fine} and SM_{coarse} are the soil moisture for fine and coarse grid scales, respectively; K_w is the wetness coefficient; K_a is a function of slope and aspect for a grid cell; p and q are two line-regression coefficients for the relationship between TWI and K_w; A_s is the aspect at the fine resolution. The spatial pattern of the SM_{fine} is subsequently adopted in the slope stability module, thereby featuring PHyL v1.0 with cross-scale running ability in a large-scale application.

2.3. Parallel algorithm

Parallel algorithms were designed for PHyL v1.0 to address redundancies during three critical stages: (i) the routing scheme in CREST calculates the runoff for each cell by taking all the water passing through into account (Eq. (4)); (ii) the SMD is executed at each timestep; and (iii) each random ellipsoid solves the coordinate transformation, geometric derivation, and the series solution for F_s (Eq. (5)). The computational cost increases significantly for larger study areas since more grid cells are involved. PHyL v1.0 provides the strategy that the study area is split into a defined number of subbasins and tiles for hydrological processes and slope stability modeling, respectively. The code was programmed based on the OpenMP API and the column-major order in Fortran. Below is a detailed description of the algorithms.

2.3.1. Parallelization in CREST

- PHyL v1.0 first reads input parameters N_{sub} and N_{Hthread} (Algorithm 1). The term $N_{\text{sub}} > N_{\text{Hthread}}$ is recommended to efficiently exploit multi-core computers. The stream matrix is read as a basic input and is subsequently processed into a vector, $p_{\text{channel}} = [p_1, p_2, \cdots, p_n]^T$ (see Fig. 3e).
- The original basin is split by looping through all elements in p_{channel} until N_{sub} subbasins are created (Algorithm 1), and the cutoff channel pixels are marked as pour points (e.g., A, B, C, and Outlet in Fig. 3a). PHyL v1.0 thereby initializes the objective area as the set of subbasins ($S = [S_1, S_2, \cdots, S_{N_{\text{sub}}}]$) and channel pixels ($p_{\text{channel}}$).
- The model simulates subbasins separately for all hydrological processes after sending them to N_{Hthread} independent threads (Fig. 3d).
- The interaction only occurs when the routing water flows pass from the observing subbasin to another at each timestep. For example, in Fig. 3b, cell#1 and #3 accomplish the routing within S1 and S2, respectively. The interaction is detected by pour points A and B (denoted as p_1 and p_2 in Fig. 3e) for cell#2 ($S1 \rightarrow S2$) and #4 ($S2 \rightarrow S3$) and thereby forces the model to update the S and p_{channel}.
- The computations in a general timestep end with reconstructing the subbasins to the original basin.

It is worth noting that the watershed outflow location serves as the default pour point, and the calculation of pour points other than the default location requires Algorithm 1. In addition, the long, narrow watersheds may result in larger dimensions for channel vectors compared to compact watersheds, potentially affecting computational efficiency. However, it's worth noting that the subbasin generation process is a one-time computation performed before the main computational loop, making its impact on overall efficiency relatively modest. Apart from the main hydrological simulations, all other calculations are performed using one-dimensional column vectors, thereby fully exploiting the efficiency of the Fortran language.

Algorithm 1. Pseudo-code structure of the parallel hydrological module

Algorithm 1 Pseudo-code structure of the parallel hydrological module

```
procedure READ_PARALLEL_PARAMETERS(N_sub, N_Hthread)
    N_sub ← user-defined value
    N_Hthread ← user-defined value        \\N_sub ≥ N_Hthread
end procedure
procedure HYDROPARALLEL_PRE()
    A_subbasin ← A_basin/N_sub
    p_channel ← [p_1, p_2, ··· , p_n]^T
    A_residue ← A_basin
    while A_residue > A_subbasin do
        A_i ← Cal_upstream(p_i)        \\Calculate the upstream area
        if A_i = A_subbasin then
            S_i ← Subbasin_write(p_i)        \\Write S_1, S_2, ··· , S_{N_sub}
            A_residue ← A_basin − A_i
            p_channel ← p_residue        \\Update the searching p_channel
        else
            p_i ← p_{i+1}
        end if
    end while
    S ← [S_1, S_2, ··· , S_{N_sub}]        \\ Arrange all the subbasins
end procedure
procedure HYDROCAL_MAIN()
    for k ← 1 to n_t do        \\ Time step loop
        Initialize()
        do in parallel
            send S_i to Proc_i        \\ Send subbasins to processes
            S_i ← CREST_main(S_i)
        end do
        if Interaction detected = .true. then
            S_i ← Update(S_i)
            p_i ← Update(p_i)
        end if
        S ← Reconstruct(S_i)
    end for
end procedure
```

2.3.2. Parallelization in 3D regional slope stability model

- Similar to Algorithm 1, PHyL v1.0 reads the parallel and landslide parameters, N_{tile}, N_{Lthread}, and D_e. $N_{\text{tile}} > N_{\text{Lthread}}$ is also recommended (see in Algorithm 2).
- The original inputs (in the form of matrices) are split into N_{tile} tiles (Fig. 4a) by calculating the $N_{\text{pixel,tile}}$ for each. These tiles are subsequently reshaped into vectors and form a new set ($T_i = [\vec{T}_1, \vec{T}_2, \cdots, \vec{T}_{N_{\text{tile}}}]$).
- At each timestep, the simulated soil saturation is downscaled into fine resolution via Eq. (6). Then PHyL v1.0 separately calculates the slope stability for each tile by running the code in N_{Lthread} threads (Fig. 4a).
- A total of $N_{\text{ellipsoid}}$ random ellipsoids are computed for each tile. For each potential landslide, PHyL v1.0 minimizes the computation by restricting the main calculation to a smallest matrix (Fig. 4b) that is extended from the random ellipsoid centre and shape. All necessary matrices (such as DEM and slope maps) are also extracted and prepared for computation. The convergence is reached by pixel-level iteration between the overlapped matrices (Algorithm 2, line 25, and Fig. 4b).
- The final landslide susceptibility results are obtained by collecting and combining all the tiles (Algorithm 2, line 32).

Algorithm 2. Pseudo-code structure of the parallel 3D slope stability module

Algorithm 2 Pseudo-code structure of the parallel 3D slope stability module

```
1:  procedure READ_PARALLEL_PARAMETERS(N_tile, N_Lthread)
2:      N_tile ← user-defined value
3:      N_Lthread ← user-defined value        \\N_tile ≥ N_Lthread
4:      D_e ← user-defined value
5:  end procedure
6:  procedure LANDPARALLEL_PRE()
7:      N_pixel,tile ← N_pixel,tot/N_tile
8:      A_tile ← N_pixel,tile · Δx · Δy
9:      N_ellipsoid ← 16 · A_tile · D_e/(π · (a_e,min + a_e,max) · (b_e,min + b_e,max))
10:     [T_1, T_2, ··· , T_{N_tile}] ← MatrixToTile(N_pixel,tile, N_tile)
11:     T_i ← T⃗_i (i = 1 to N_tile)        \\ To one-dimensional vectors
12: end procedure
13: procedure LANDCAL_MAIN()
14:     Require: hydrological variables, soil properties, topography
15:     for k ← 1 to n_t do        \\ Time step loop
16:         S_M,fine ← Soil_downscale(S_M,coarse)
17:         Initialize()
18:         do in parallel
19:             send T_i to Proc_i        \\ Send tiles to processes
20:             for m ← 1 to N_ellipsoid do        \\ Potential landslide loop
21:                 (x_c, y_c, z_c) ← RandomCenterSample()        \\ Center
22:                 (a_e, b_e, c_e) ← RandomShapeSample()        \\ Shape
23:                 C_Matrix ← WindowExtend(center, shape)
24:                 (F_s,m, P_f,m, V_L,m, A_L,m) ← 3DSlopeStability()
25:                 if Ellipsoid overlapped = .true. then
26:                     F_s = min_{∀m∈N_ellipsoid} F_s,m
27:                     V_L = max_{∀m∈N_ellipsoid} V_L,m
28:                     A_L = max_{∀m∈N_ellipsoid} A_L,m
29:                 end if
30:             end for
31:         end do
32:         Maps(F_s, P_f, V_L, A_L) ← Reconstruct(T_i)
33:     end for
34: end procedure
```

2.4. Parameters required in PHyL v1.0 and sensitivity analysis

The modeling parameters in PHyL v1.0 are categorized as hydrological parameters, regional 3D slope stability parameters, and parallel parameters, as listed in Table 1. The hydrological module inherits most of parameters shared in the CREST family (Flamig et al., 2020b; Huang et al., 2022; Xue et al., 2013), which are either determined in a distributed manner by inputs or calibrated to uniform values. The parameters for the slope stability module are highly related to soil property inputs, while others are defined by the user to control the random sampling. The parallel parameters (N_{sub}, N_{Hthread}, N_{tile}, and N_{Lthread}), as

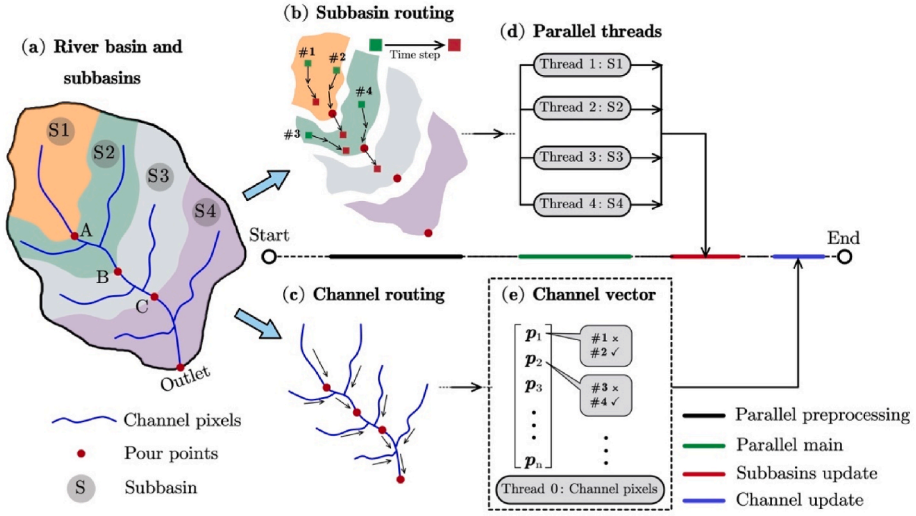


Fig. 3. Logical framework of the parallel procedures in the hydrological module. The procedure starts with the (a) decomposition of a drainage network into several subbasins; the model separately calculates the (b) subbasin routing and (c) channel routing; after the calculation of the (d) parallel threads at each timestep, the variables of the subbasins and (e) channel vector are updated sequentially.

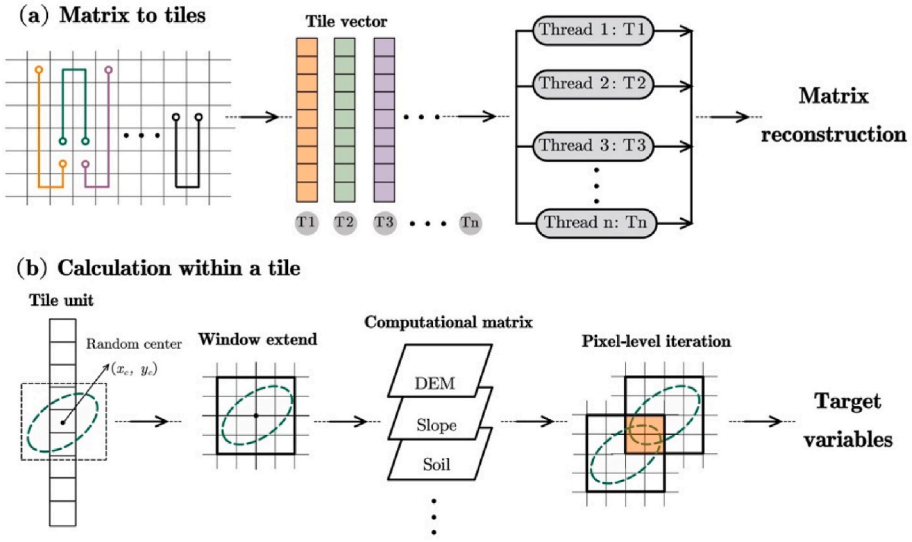


Fig. 4. Logical framework of the parallel procedures in the 3D slope stability module: (a) decomposition of the topographical matrices; (b) minimum computational window matrix that encompasses a potential failure.

introduced in Sect. 2.3, are freely set but are best suited for fully exploiting multi-core computers (e.g., divisibility relationships are recommended between N_{sub} and N_{Hthread}, N_{tile} and N_{Lthread}).

To investigate the sensitivity of the model parameters, a one-factor-at-a-time (OAT) approach was used, where each parameter was changed independently while holding all others constant. The results of the OAT analysis were compared using a sensitivity index (S_I), which was defined using central differences for the derivatives (Lenhart et al., 2002; Luo et al., 2008):

$$S_I = \frac{\Delta P}{\Delta I}\frac{I}{P(I)} = \frac{P(I+\Delta I) - P(I-\Delta I)}{2\Delta I}\frac{I}{P(I)}. \tag{7}$$

Here, P represents the dependent variable in the sensitivity analysis, while I is the input parameter. The absolute value of S_I indicates the degree of sensitivity of the target parameters for model predictions, with a larger absolute value indicating a higher sensitivity. In this study, the total runoff volume at the outlet and the regional unstable area were the target variables of interest.

2.5. Input/output (I/O) in software

HDF5 (Hierarchical Data Format 5) file format was employed as the I/O system of PHyL v1.0. HDF5 is a versatile file format that is widely used for storing and managing large and complex datasets (https://www.hdfgroup.org/). The PHyL v1.0 uses built-in compression algorithms, including the GZIP and SZIP algorithms, offering efficient storage and fast I/O performance without sacrificing data integrity. We also created a set of high-level wrapper subroutines instead of utilizing the original HDF5 Fortran API, supporting handier functions to use in PHyL v1.0.

PHyL v1.0 utilizes the same basic input datasets as iHydroSlide3D v1.0 (see Table 2 in Chen et al. (2023)). The outputs cover all the variables with respect to hydrology (e.g., precipitation, soil moisture, and runoff) and landslide susceptibility (F_s, P_f, A_L, and V_L). The results are saved in various groups within a single HDF5 file of interests, facilitating easy access and retrieval of relevant information. Note that applying PHyL v1.0 for cross-scale analysis yields landslide susceptibility results with a higher resolution than that of the hydrological processes. All parallel units can communicate with each other at each time step without immediate combination, and the combination occurs only when users wish to output specific variables. Additional detailed information for I/O can be found in the manual for PHyL v1.0, which is available in the "Software availability" section.

2.6. Visualization

PHyL v1.0 enhances the understanding of simulated outcomes by leveraging the Python libraries and tools. The visualization packages Matplotlib and OpenCV (CV) were adopted. The modeled spatial patterns of the aforementioned variables are plotted into figures and animations to monitor the evolutions of hydrological processes and slope stability states.

2.7. Model performance evaluation

The modeled discharge was validated using streamflow observations from the local gauge stations. PHyL v1.0 computes a range of statistical metrics, including the Nash–Sutcliffe coefficient of efficiency (NSCE), Pearson correlation coefficient (CC), and relative bias, to quantify the prediction performance of floods. To quantitatively evaluate the predictive capabilities of the model for landslides, a contingency table is constructed to calculate a series of indices: (1) True Positives (TP); (2) False Negatives (FN); (3) False Positives (FP); and (4) True Negatives (TN). The receiver operating characteristic (ROC) (Fawcett, 2006) curve

Table 2
Computational metrics.

Index	Expression	Range	Best value
NSCE	$1-\frac{\sum_{t=1}^{T}(Q_o^t-Q_S^t)^2}{\sum_{t=1}^{T}(Q_o^t-\overline{Q}_o)^2}$	$(-\infty,1)$	1
CC	$\frac{\text{cov}(O,S)}{\sqrt{\text{var}(O)\text{var}(S)}}$	$(-1,1)$	1
Bias	$\frac{\sum_{t=1}^{T}(S_i-O_i)}{\sum_{t=1}^{T}O_i}$	$(-\infty,+\infty)$	0
TPR	TP/(TP + FN)	(0,1)	1
FPR	FP/(FP + TN)	(0,1)	0
AUC	$\int_0^1 ROC$	(0,1)	1
Speedup ratio	T_s/T_p	$(1,+\infty)$	$+\infty$
Parallel efficiency	Speedup ratio/N	(0,1)	1

Note: the symbol "O" represents the observations, and "S" is the simulated values.
T_s is the serial execution time; T_p is the parallel execution time; N is the number of threads.

that consists of TPR and FPR (Table 2) pairs was applied to analyse the comparison between modeled slope failures and the landslide inventory database. The area under the ROC curve (AUC) further provided a global statistical accuracy indicator. In parallel aspects, the speedup ratio and efficiency were used to evaluate the effectiveness of parallel computing systems in PHyL v1.0. Table 2 lists all the metrics with more detailed information.

3. Case study: rainfall-induced floods and landslides at the catchment scale

3.1. The Yuehe River Basin and event description

We tested the PHyL v1.0 code in the Yuehe River Basin, located in Shaanxi Province, China (Fig. 5a). This basin encompasses a total area of 1100 km^2, with an elevation range of 270–2700 m. The topography of the basin is characterized by steep hills, gullies, and valleys, as reported by Zhang et al. (2019). Additionally, the soil type mainly featured clay loam and loam (see Fig. 5c) (Chen et al., 2023). The study area has a typical subtropical monsoon climate with an annual average precipitation range of 700–1100 mm. Approximately 80% of the precipitation falls from May to October. In July 2012, a rainstorm caused 54 landslides and floods in the study area, as depicted in Fig. 5b and d. Most of the observed landslides occurred on slopes ranging from 15° to 30° (see Fig. 5e). The locations of the failures and the discharge at the outlet were utilized for calibration since they were well-documented and served as a case for simulation.

3.2. Materials

An overview of the input data for PHyL v1.0 are provided in Table 3. Hourly precipitation data, based on observations from gauge stations, were provided by the China Meteorological Administration (CMA). The potential evapotranspiration (PET) data were derived from the Global Land Data Assimilation System (GLDAS). In this study, we utilized two different resolutions of digital elevation models (DEMs) obtained from the NASA Shuttle Radar Topography Mission (SRTM) Version 3.0 (SRTM3) and the Advanced Land Observing Satellite (ALOS). Specifically, we employed a 90-m resolution DEM for hydrological modeling and a 12.5-m resolution DEM for slope stability modeling. The DEMs were further used to derive other necessary surface properties such as slope, flow direction, and flow accumulation. The TWI data were also calculated using ArcHydro toolbox within ESRI ArcGIS. Soil texture was classified into the 12 United States Department of Agriculture (USDA) soil texture types, as described in the Harmonized World Soil Database (HWSD v1.2) (Wieder et al., 2014). A lookup table was employed by both the hydrological and landslide modules to facilitate this classification. For more detailed information, including data sources and resolution, please refer to Chen et al. (2023), which focuses on the same geographic region as this work.

Table 3
Summary of input datasets/maps.

Datasets type	Specific inputs
Topographic properties	Digital elevation model, flow direction, flow accumulation, topographic wetness index, river network, and topographic curvature
Soil texture	The Harmonized World Soil Database (HWSD v1.2)
Meteorological forcing	Observed precipitation and evapotranspiration data
Parameters	Dataset-derived and calibrated values
Calibration/ verification data	Investigation of floods and landslides within the study area

The simulation parameters utilized in this study are presented in Table 1. Distributed values pertain to parameters determined a priori from input datasets, while uniform values were manually calibrated for the hydrological module based on observed streamflow. The range of the random ellipsoid's maximum and minimum values in three dimensions was determined based on the landslide inventory. To ensure area convergence, D_e was evaluated in previous research by Mergili et al. (2014a) and subsequently set at 500. The selection of parallel parameters was guided by the examination of test results and hardware, as detailed in forthcoming sections.

3.3. Hardware

The code was executed on the high-performance computing system, DelftBlue, located at TU Delft (Centre, 2022). DelftBlue is specifically designed to address the computational requirements of complex problems in the fields of physics, mechanics, and dynamics. DelftBlue boasts an impressive configuration, with over 220 compute nodes, containing

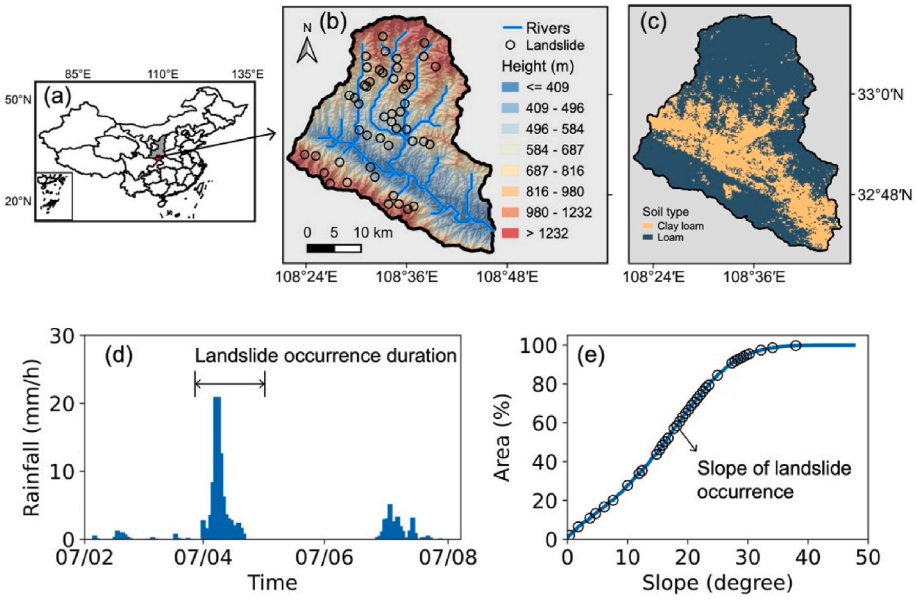


Fig. 5. (a) Location of the Yuehe River Basin in China; (b) catchment boundaries, elevation, main streams, and landslides; (c) Soil type; (d) the rainstorm during July 3rd to 4th, 2012; (e) slope distribution in the basin and the slope for landslide occurrence.

Table 4
Simulation scenarios in the Yuehe River Basin using PHyL v1.0.

Scenario	Parameters	Simulated duration	Output variables	File format
1. Run overview	D_e = 10, 50; all other parameters are listed in Table 1.	July 3rd to 4th	R, SM, W, F_s, P_f, V_L	ASC and HDF5
2. Parallel test	D_e = 10, 50; N_{sub} = 6, 12, 24, 48; N_{tile} = 20, 48, 96, 144; $N_{Hthread}$ and $N_{Lthread}$ = 1, 2, 4, 6, …, 48; all other parameters are listed in Table 1.	July 3rd to 4th	R, SM, W, F_s, P_f, V_L	HDF5
3. Sensitivity analysis	Calibration parameters are incremented by 10% within the interval; all other parameters are listed in Table 1.	July 2nd to 9th	$R_{tot}, \%(F_s < 1)$	HDF5
4. Case reproduction	All the parameters are listed in Table 1.	July 2nd to 9th	R, SM, W, F_s, P_f, V_L	HDF5

Note: R is the runoff; SM is the soil moisture; W is the water amount in a soil column; R_{tot} is the total runoff volume at the catchment outlet; $\%(F_s < 1)$ is the percentage of the unstable area.

more than 11,000 CPU cores. In addition, it incorporates a high-speed parallel storage subsystem based on the BeeGFS file system. For low-latency inter-node communication and high-throughput data transfer, all compute nodes and storage systems are interconnected with HDR100 InfiniBand technology. Each compute node is equipped with 2 Intel XEON E5-6248R 24C 3.0 GHz CPUs, providing a total of 48 cores and 192 GB of memory. A more detailed description can be found in DelftBlue Documentation.

3.4. Simulation scenarios

The PHyL v1.0 underwent testing in the Yuehe River Basin, wherein four distinct scenarios (see in Table 4) were employed to assess its efficacy. The first scenario aimed to evaluate the I/O system, specifically the advantages of employing the HDF5 file format. The second scenario constituted a parallel test that aimed to gauge the performance of multi-core computation by varying $N_{Hthread}$ and $N_{Lthread}$. The third scenario carried out a sensitivity analysis of parameters. Finally, scenario 4 comprised a comprehensive simulation that tried to reproduce flood and landslide events during a rainstorm (as detailed in Sect. 3.1) after the calibration of PHyL v1.0. The applied region was the Yuehe River Basin and the modeling timestep for all scenarios was 1 h.

4. Results

4.1. Overall performance

Table 5 presents an overview of the runtime and output storage for scenario 1. The runtime was measured separately for each submodule, and the disk storage occupied by outputs was recorded. The runtime for the slope-stability simulation generally exceeds that of the hydrological module. Notably, increasing the value of D_e significantly prolongs the runtime for landslide prediction without altering the I/O time, as the dimensions of the output matrix remain unchanged. The I/O time can dominate the total runtime (even up to ~90%) when applying a relatively small D_e and using the conventional ASC file format. However, the HDF5 format demonstrated its superiority by increasing the I/O speed by approximately tenfold. Regarding storage, the ASC format required 10 GB for one-day duration outputs (24 timesteps in total), while HDF5 only required about 20 MB, achieving a compression rate of over 550 times.

4.2. Performance of multi-core computation

Scenario 2 was conducted to evaluate the parallel performance of PHyL v1.0's hydrological and slope stability modules separately by varying the number of parallel processes (N) in a loop (Fig. 6). $N = 1$, representing serial computation, was used as a baseline for speedup analysis. In Fig. 6a, it can be observed that the speedup for the hydrological module did not consistently increase when the number of sub-processes (N_{sub}) is less than N. For $N_{sub} \geq N$, the speedup generally increases, albeit with some fluctuations, with N, peaking at a maximum value of ~8 (N_{sub} = 48 in Fig. 6a). Nevertheless, for all N_{sub} values tested, the parallel efficiency for the hydrological module drops rapidly as N increases, with the lowest efficiency falling below 0.2 (Fig. 6b). The slope stability module demonstrates superior parallel performance in terms of speedup and efficiency, as shown in Fig. 6c~f. Similar to the hydrological module, the speedup of the slope stability module is limited by the number of sub-processes (N_{tile}), with a maximum value of ~14 achieved for N_{tile} = 20 (Fig. 6c and e). For $N_{tile} \geq N$, a near-linear speedup growth was observed, with values ranging from 20 to 30. As N increases, the efficiency gradually decreases, resulting in a minimum range of 0.2–0.6 for N = 48 (Fig. 6d and f). A higher efficiency was obtained with larger N_{tile}, with values exceeding 0.6 for various parallel processes. The parameters D_e had no effect on the parallel performance but only influenced the overall execution time.

4.3. Sensitivity analysis of the parameters

Fig. 7 provides a comprehensive overview of the computed S_I for all calibrated parameters listed in Table 1. The analysis indicates that *coeR*, which represents channel flow parameters, had the most dominant influence on hydrologic runoff, followed by the overland flow speed parameters, *coeM* and *expM*. The impact of the interflow parameter *coeS* and the infiltration curve parameters (B) on runoff generation was comparatively lower. The partitioning of water from overland and interflow reservoirs to discharge for a general cell, as represented by *KS* and *KI*, exhibited minimal effects. The landslide parameter D_e, being effective only in the slope stability module, was found to be irrelevant to hydrological processes. For the regional unstable area, hydrological parameters were observed to have a slight effect on slope stability calculations, with B being relatively more sensitive, as demonstrated in Fig. 7b. However, the sensitivity of D_e greatly exceeds that of other

Table 5
Overview of the typical runtime and output storage of the PHyL v1.0.

Running case	Hydrological runtime (s)	Landslide runtime (s)	I/O time (s)	Total runtime (s)	Disk storage (MB)
D_e = 10 & I/O = ASC	20.5	99.2	1065.4	1185.1	10963
D_e = 10 & I/O = HDF5	19.5	99.7	97.1	216.3	19.9
D_e = 50 & I/O = ASC	19.9	498.6	998.1	1516.6	10963
D_e = 50 & I/O = HDF5	22.8	494.7	91.5	609.0	21.06

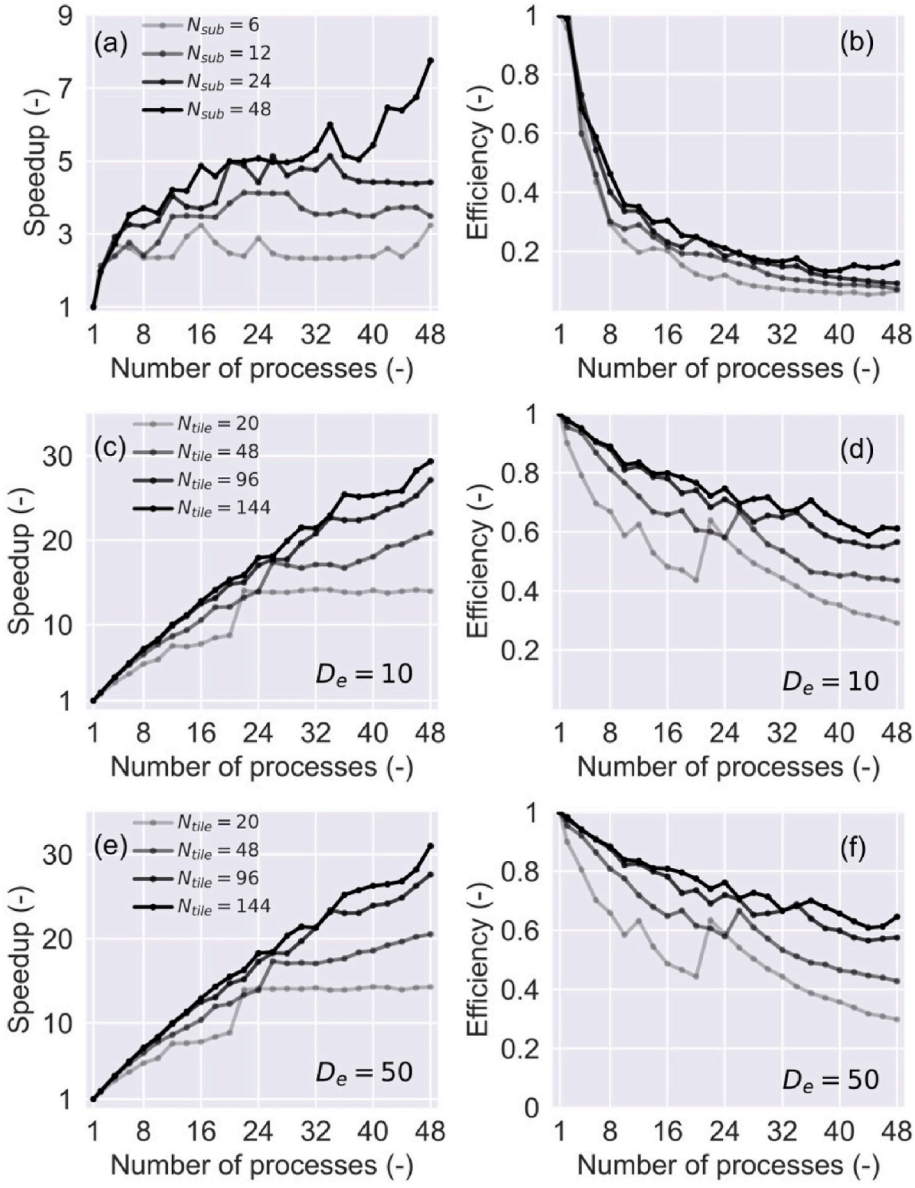


Fig. 6. Parallel speedup and efficiency plotted against the number of processes (N) for the hydrological module (a and b) and slope stability module (c ~ f), respectively. The number of the subbasin (N_{sub}), tiles (N_{tile}), and sampling density (D_e) are also tested. See the text for further explanations.

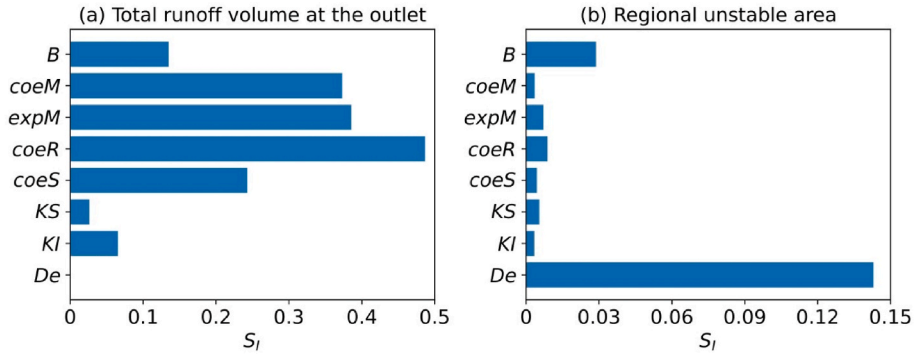


Fig. 7. Sensitivity analysis of the calibrated parameters in Table 1. The objectives are chosen as (a) runoff volume at the watershed outflow location and (b) regional unstable area.

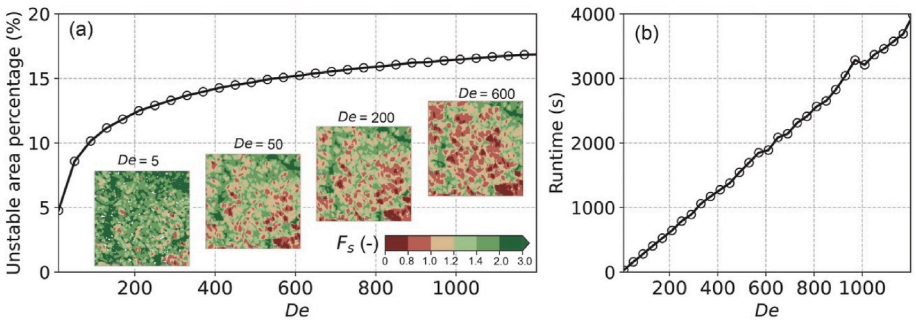


Fig. 8. Sensitivity analysis of the D_e related to (a) regional unstable area percentage and (b) runtime. The insets in (a) show the typical spatial pattern of F_s obtained with various D_e.

parameters, which was further tested in detail (Fig. 8). The increased D_e made the unstable area rapidly increase in the early stage and subsequently tend to converge (Fig. 8a), while too small values of D_e did not cover the entire area (e.g., the inset of D_e= 5 in Fig. 8a). Conversely, larger D_e values led to complete coverage through the overlap between random ellipsoids. In addition, a near linear dependency of the runtime on D_e was found in this work (Fig. 8b).

4.4. Outputs of the variables for regional analysis of hydrological processes and landslide susceptibility

PHyL v1.0 was employed as a simulator for hydrological processes and landslide modeling in the Yuehe River Basin. Specifically, the experimental setup of scenario 4 in Table 4 was manipulated to achieve this aim. The used parameters have been listed in Table 1. Visualization figures were produced using the built-in visualization functions and five representative moments were selected to encapsulate the rainfall peak.

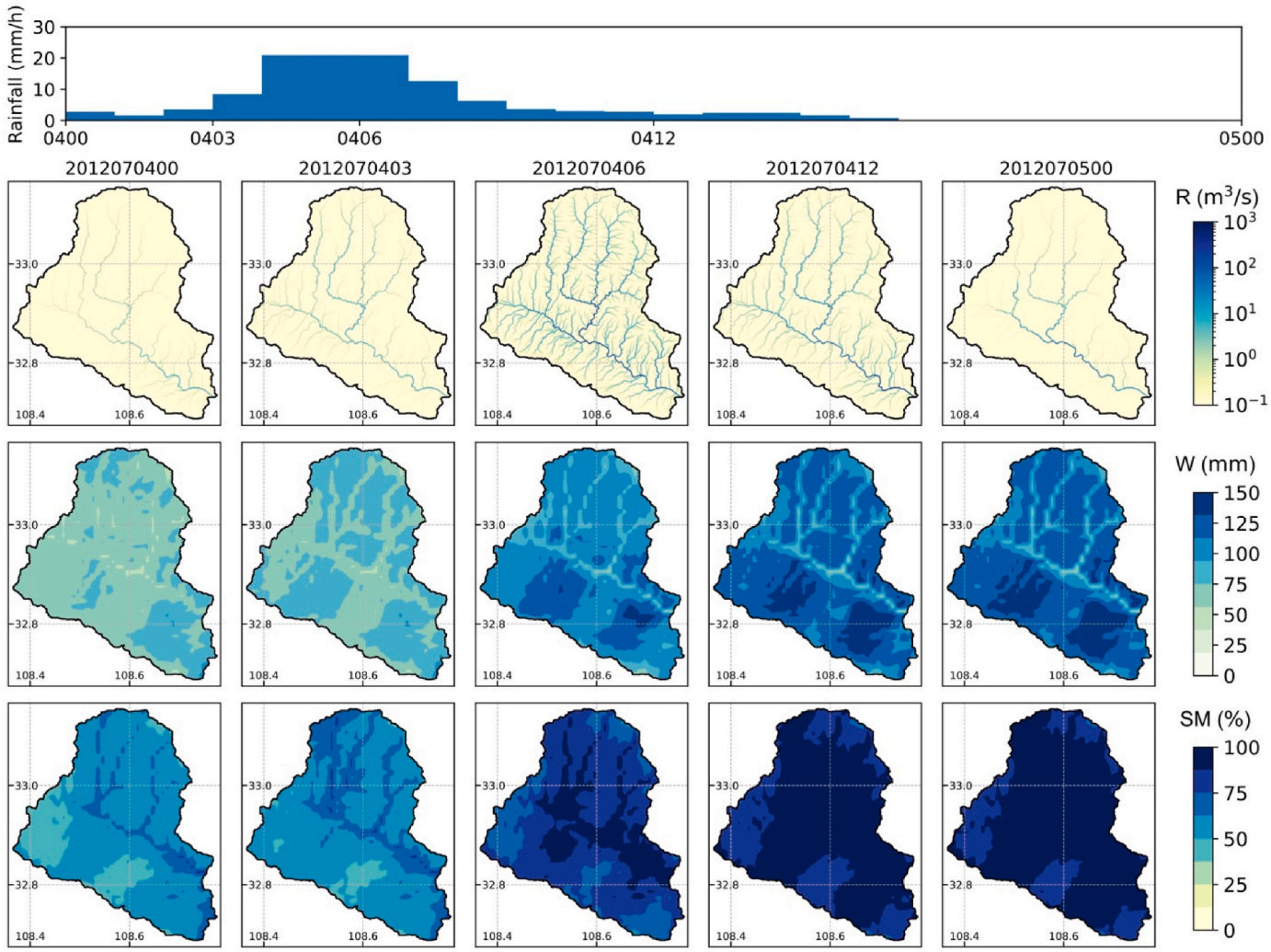


Fig. 9. Temporal evolution of the spatial pattern of the runoff (*R*), water amount filling the pore space (*W*), and the soil moisture (*SM*). Rainfall series presented were extracted from full records.

In Fig. 9, the typical spatial patterns of hydrological variables, including runoff, water amount in soil, and soil moisture were illustrated. Results reveal that during the initial stage of the rainstorm, runoff was primarily observed in the main river channels. Additionally, due to the antecedent soil moisture, the soil water amount was below 50 mm in most regions. The soil moisture exceeded the residual value, with relatively larger values found in mountain ridges as a result of the smaller water storage capacity. During the rain peak on Jul. 4 at 06:00, visual runoff was observed in almost all routing channels of the drainage network, reaching a maximum of approximately 1000 m^3/s. The excess portion of rainfall flowed downstream, resulting in lower soil water amount but higher soil moisture in ridge cells. At this moment, the soil moisture increased drastically across the region, with a small portion reaching the fully saturated soil state. Following the rain peak, the runoff gradually decreased, although the discharge at the main river did not significantly reduce. Similar patterns of *W* and *SM* were observed for the moments of Jul. 4 at 12:00 and Jul. 5 at 00:00, as the soil pore space was filled.

The susceptibility of landslides, as illustrated in Fig. 10, exhibited a strong correlation with soil moisture, which exerted a significant impact on suction stress. Prior to the occurrence of the rainstorm, the majority of the study area was deemed stable (as evidenced by $F_s > 1$ and $P_f = 0$), with the exception of a small portion that could have resulted from antecedent rainfall or inherent instability (i.e., unconditional unstable). However, these cells were calculated with low values of P_f and staggered most observed landslide locations in this case. The unstable region ($F_s < 1$) markedly expanded during the rainstorm, while the flat terrain remained stable. Similar behavior was also observed for P_f. Following the peak (Jul. 4 at 12:00), continued infiltration led to further expansion of the unstable region due to the loss of matric suction. The proportion of the area corresponding to $P_f > 0$ also increased, with a small section exceeding 60%. The slope stability patterns remained unchanged after the soil became fully saturated (since Jul. 4 at 12:00 in Fig. 10). Additionally, PHyL v1.0 provided landslide magnitude estimates for all unstable cells (referred to as V_L in this work), with values ranging from 4×10^4 to 4×10^5 m^3.

4.5. Model accuracy evaluation

The validation of model outputs was conducted using the performance metrics listed in Table 2. The modeled floods and landslides were compared with observations in the Yuehe River Basin (see Fig. 11). The comparison between modeled and observed discharge showed a satisfactory agreement, with values of NSEC = 0.72, CC = 0.86, and Bias = 4.0% (Fig. 11a). However, the faster flood receding process and slight fluctuation on Jul. 7 suggest possible uncertainties in the routing or flow concentration processes of the hydrological module in PHyL v1.0. A ROC analysis was conducted to assess the model's predictive capacity for landslides at the regional scale, with an AUC score of 0.74 (see Fig. 11b).

5. Discussion

The present study has evaluated the computational efficiencies, parallel performance, parameter sensitivity analysis, and predictive capabilities of the proposed software, PHyL v1.0. In comparison to other available parallel frameworks, as listed in Table 6, PHyL v1.0 stands out due to its ability to parallelize both hydrological and 3D slope stability modules, allowing for efficient operation at a broad scale. As a new version of its precursor, PHyL v1.0 has transitioned to a fully distributed linear reservoir routing method (Shen et al., 2017), departing from the earlier quasi-distributed approach. While PHyL v1.0 retains its foundational theoretical principles from iHydroSlide3D v1.0, the programming languages and architectures were reengineered to accommodate the newly introduced parallel algorithms and the software operating

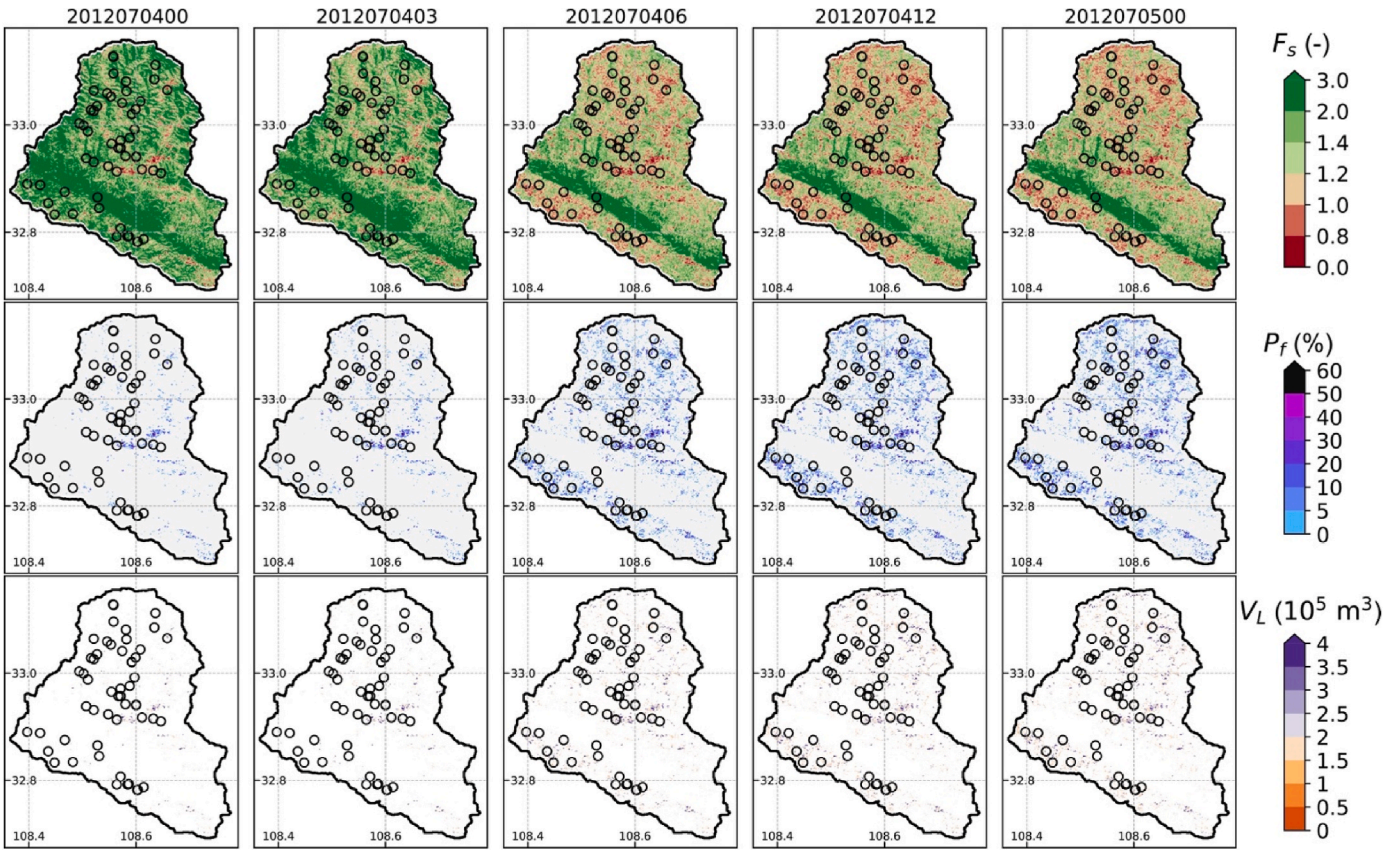


Fig. 10. Temporal evolution of the spatial pattern of the factor of safety (F_s), landslide occurrence probability (P_f), and landslide volume (V_L).

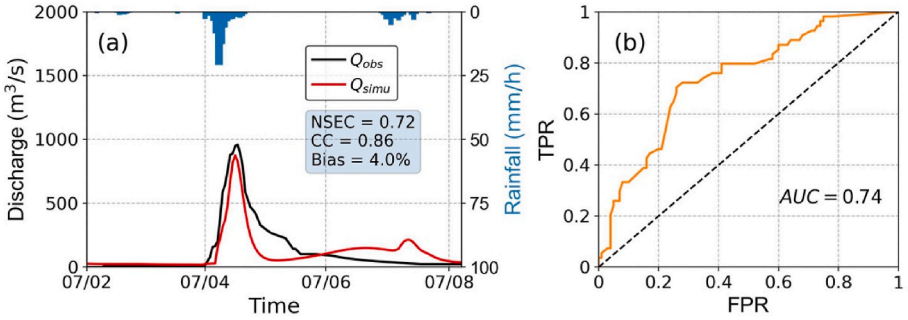


Fig. 11. Simulating performance of the PHyL v1.0: (a) modeled hydrographs against the observed streamflow; (b) ROC-AUC analysis relating unstable area to the observed landslides in the entire study area.

Table 6
A summary of parallel hydrological/landslide modeling framework.

Model/software	Key features	Protocol and language	Parallel unit	Speedup ratio	Efficiency
PHyL v1.0[a] (this study)	Parallelize both hydrological and 3D slope stability modules	OpenMP; Fortran and Python	Subbasin, Channel, and one-dimensional tile	~8/48; ~30/48	0.17; 0.63
HIRESSS[a] (Rossi et al., 2013)	Monte Carlo (MC) simulation; analytical hydrological solution; infinite slope	MPI; C++	MC task	52/512	0.10
TRIGRS[a] v2.1 (Alvioli and Baum, 2016)	Parallelize the file reading procedures and core computation subroutine	MPI; Fortran	Grid loop	~120/250	0.48
r.slope.stability[b] (Mergili et al., 2014a)	Probability calculation; 3D slope stability model	Python threading; C, Python, and R	Matrix tile	~8/42; ~23/42	0.19; 0.55
tRIBS[c] (Vivoni et al., 2011)	Triangulated Irregular Network (TIN)-based model; Load balancing for the simulation complexity	MPI; C++	subbasin	~70/512	0.14
SEIMS[c] (Zhu et al., 2019)	A modular and parallelized watershed modeling framework	OpenMP + MPI; C++ and Python	basic-unit, subbasin, and model	~18/40	0.45
APPA/CHESS[c] (Xu et al., 2021)	Automatic partition-based parallel algorithm	OpenMP; C++	subbasin and channel	~20/26	0.77

Note:
The speedup ratio is presented in the form of "speedup ratio/number of processes". Two speedup ratios of r.slope.stability refer to the small and large values of D_e, respectively.
[a] coupled hydrological-geotechnical model;
[b] pure slope stability model;
[c] pure hydrological model.

environment. This transition from previous commercial software environments (i.e., MATLAB) to an open-access, more accessible platform simplifies installation and offers modularity through separate subroutines for greater flexibility. Additionally, the I/O and visualization components facilitate rapid analysis and improved understanding of the coupled hydrological-geotechnical behaviour. As such, this model has the potential to address the key challenge of increasing demand for high-resolution, large-scale, and complicatedly coupled attempts in the community.

The computational burden of hydrological models is predominantly attributed to the flow routing process. The presented subbasin decomposition algorithm assumes that only one downstream receptor grid is assigned to the flow generated in an upstream grid (O'Callaghan and Mark, 1984). This criterion has been widely adopted as a parallel principle in many grid-based distributed hydrological models (Liu et al., 2013, 2014; Xu et al., 2021). PHyL v1.0 follows the dynamic parallelization concept in Li et al. (2011) and can flexibly generate any number of subtasks, resulting in increased computing acceleration when $N_{\text{sub}} > N_{\text{Hthread}}$, as opposed to the static parallel algorithm where a single computing thread handles only one subbasin (Kolditz et al., 2007; Vivoni et al., 2005). The PHyL v1.0 also includes a subbasin detection method (Fig. 3b and e) that reduces communication to the grid and minimizes idle time. While the efficiency of the hydrological module in PHyL v1.0 is not exceptional compared to other available parallel hydrological models (Table 6), this can be explained by the current low-cost version of CREST, which has a large irreducible serial component, and the absence of load balance or theoretical maximum speedup ratio (TMSR) considerations (Liu et al., 2013; Wang et al., 2012; Xu et al., 2021). In addition, PHyL v1.0 simplifies the physical foundation of soil water flow by avoiding the direct solving of Richards' equation (e.g., as done in GEOtop by Endrizzi et al. (2014) and Tufano et al. (2021)). Nevertheless, the proposed parallel scheme of PHyL v1.0 is poised for higher complexity models such as flood inundation (Li et al., 2021) and advanced solver for soil infiltration (Orgogozo et al., 2014; van Dam et al., 2008).

The simulation of slope stability is found to have a reduced dependency on logical relationships in comparison to the decomposition of river networks for parallelization purposes. By leveraging the capabilities of multi-processor computing environments, it becomes feasible to execute complex slope stability models over sizable areas while maintaining acceptable time constraints. The utilization of 3D geometry in PHyL v1.0 provides an improved depiction of the spatial heterogeneity between localized cells and their surrounding areas in relation to soil characteristics and surface topography, as evidenced by previous studies (Mergili et al., 2014a; Tran et al., 2018; Xie et al., 2004). The resulting outputs from this model, including F_s, P_f, and V_L, enable a comprehensive assessment of landslide susceptibility and hazard (Corominas et al., 2014; Vandromme et al., 2020). Notably, PHyL v1.0 exhibits superior performance over r.slope.stability (Mergili et al., 2014a) with respect to parallel efficiency in 3D slope stability modeling (0.63 versus 0.19 and 0.55). Additionally, PHyL v1.0 demonstrates consistent acceleration even when faced with varying sampling density values, and its runtime increases quasi-linearly with D_e. These outcomes can be attributed to the enhancements made to the parallel algorithm, which

include (i) the avoidance of null cells in tiles during processing, (ii) treating tiles as one-dimension arrays to accommodate the Fortran nature and irregular shape of the study area, and (iii) minimizing computation for each ellipsoid to a smallest window, thus saving RAM. However, the presence of invalid ellipsoids that intersect with ridge or valley lines due to the random process may hinder sampling efficiency. Future studies shall focus on considering the slope unit (Alvioli et al., 2016; Jacobs et al., 2020) as a parallel unit.

Another auxiliary strategy employed by the PHyL v1.0 involves utilizing the SMD method for cross-scale modeling in order to effectively address discrepancies in resolution requirements across various submodules. Prior research endeavours have employed SMD with great success, resulting in the significant expansion of landslide prediction capabilities (Leonarduzzi et al., 2021b; Wang et al., 2020). Our modular SMD approach also has the potential to be incorporated into other modeling systems, wherein soil moisture serves as the input parameter. Additionally, we posit that SMD may serve as an API to enable load balancing among submodules within a parallel framework via the adjustment of simulating resolution.

In the present work, we conducted a simple sensitivity analysis for the calibrated parameters. A similar summary was also provided by Huang et al. (2022) for CREST family models. Users can expedite the adjustment of parameters to match flood discharge by identifying the most sensitive ones. The results revealed that infiltration curve parameters (B) have a greater effect on soil stability than runoff-related parameters, which is not surprising given the association between landslide triggering mechanisms and soil water, and thus infiltration processes (Iverson, 2000; Lu and Godt, 2008). The still explicit slight effects could be attributed to the dynamic nature of subsurface hydrology and the complex interactions among processes in CREST. However, these parameters are highly correlated with surface flow velocity that could profoundly affect debris flows regarding propagation and/or deposit (Cui et al., 2014; Sidle et al., 2019), which is not considered in PHyL v1.0. The sensitivity of D_e was investigated as a parameter for achieving both result convergence and time efficiency. It is recommended that the test for D_e be repeated when applying PHyL v1.0 in a new region due to the potential variation in the dimensions of the ellipsoids (a_e, b_e, and c_e). While other soil property parameters such as c_s and φ' were found to be sensitive (Medina et al., 2021), they were not calibrated in this study. Instead, the PHyL v1.0 was utilized to determine all soil parameters by reading the soil map as input. The primary challenge that remains is to account for geotechnical uncertainty through the use of statistical models (Raia et al., 2014; Rossi et al., 2013).

6. Conclusions

This study presents PHyL v1.0, a novel parallel and coupled modeling framework designed for forecasting cascading flood-landslide disasters over large areas. The model integrates a fully distributed hydrological model and a 3D slope stability model, and incorporates a soil moisture downscaling method to address discrepancies in resolution requirements across submodules. The proposed parallel algorithms enable the simulation to be decomposed into any number of subtasks, thereby facilitating the exploitation of multi-processor computing environments. Additionally, PHyL v1.0 features auxiliary components such as I/O and visualization procedures, which make it a comprehensive modeling software. The computational efficiencies, parallel performance, parameter sensitivity analysis, and predictive capabilities of PHyL v1.0 were evaluated in the Yuehe River Basin. The flexible and extensible modular structure facilitates easy use and rapid development of PHyL v1.0 in multiple scenarios. Future iterations of PHyL v1.0 may focus on the development of more advanced solvers for flood inundation and soil infiltration, as well as the incorporation of statistical models to account for geotechnical uncertainty.

Software availability

Name of software: PHyL v1.0.
Developers: Guoding Chen, Ke Zhang, Sheng Wang.
Year first available: 2023.
Program language: Fortran, Python, and CMake.
Operating systems supported: Linux, macOS, and Windows.
Hardware required: basic computer with RAM $\geq$ 3 G and the number of CPU cores $\geq$ 2
Software or environment requirements: CMake $\geq$ 3.23, GNU compiler $\geq$ 11.2, OpenMP $\geq$ 4.5, HDF5 $\geq$ 1.10, and Python $\geq$ 3.8
Program size: 4.2 MB.
Availability: PHyL v1.0 can be freely distributed on GitHub (https://github.com/GuodingChen/PHyL_v1.0), where the user manual is also included.

CRediT authorship contribution statement

Guoding Chen: Conceptualization, Methodology, Software, Writing-original draft. **Ke Zhang:** Conceptualization, Supervision, Methodology, Validation. **Sheng Wang:** Methodology, Software. **Tianlong Jia:** Validation, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code and data have been made available in the manuscript with links.

Acknowledgement

This study was supported by the National Key Research and Development Program of China (2023YFC3006505), Fundamental Research Funds for the Central Universities of China (B230202009, B220204014, B220203051), National Natural Science Foundation of China (51879067), Natural Science Foundation of Jiangsu Province (BK20180022), and Six Talent Peaks Project in Jiangsu Province (NY-004).

References

Alvioli, M., Baum, R.L., 2016. Parallelization of the TRIGRS model for rainfall-induced landslides using the message passing interface. Environ. Model. Software 81, 122–135.
Alvioli, M., et al., 2016. Automatic delineation of geomorphological slope units with r.slopeunits v1. 0 and their optimization for landslide susceptibility modeling. Geosci. Model Dev. (GMD) 9 (11), 3975–3991.
An, H., et al., 2016. Development of time-variant landslide-prediction software considering three-dimensional subsurface unsaturated flow. Environ. Model. Software 85, 172–183.
Aristizábal, E., Vélez, J.I., Martínez, H.E., Jaboyedoff, M., 2016. SHIA_Landslide: a distributed conceptual and physically based model to forecast the temporal and spatial occurrence of shallow landslides triggered by rainfall in tropical and mountainous basins. Landslides 13, 497–517.
Asgari, M., Yang, W., Lindsay, J., Tolson, B., Dehnavi, M.M., 2022. A Review of Parallel Computing Applications in Calibrating Watershed Hydrologic Models. Environmental Modelling & Software, 105370.
Baum, R.L., Savage, W.Z., Godt, J.W., 2008. TRIGRS: a Fortran Program for Transient Rainfall Infiltration and Grid-Based Regional Slope-Stability Analysis. US Geological Survey Reston, VA, USA version 2.0.
Beven, K., Lamb, R., Quinn, P., Romanowicz, R., Freer, J., 1995. Topmodel. Computer Models of Watershed Hydrology, pp. 627–668.
Centre, D.H.P.C., 2022. DelftBlue Supercomputer (Phase 1).
Chen, G., Zhang, K., Wang, S., Xia, Y., Chao, L., 2023. iHydroSlide3D v1. 0: an advanced hydrological–geotechnical model for hydrological simulation and three-dimensional landslide prediction. Geosci. Model Dev. (GMD) 16 (10), 2915–2937.

Corominas, J., et al., 2014. Recommendations for the quantitative analysis of landslide risk. Bull. Eng. Geol. Environ. 73, 209–263.
Cui, P., Guo, C.-x., Zhou, J.-w., Hao, M.-h., Xu, F.-g., 2014. The mechanisms behind shallow failures in slopes comprised of landslide deposits. Eng. Geol. 180, 34–44.
Endrizzi, S., Gruber, S., Dall'Amico, M., Rigon, R., 2014. GEOtop 2.0: simulating the combined energy and water balance at and below the land surface accounting for soil freezing, snow cover and terrain effects. Geosci. Model Dev. (GMD) 7 (6), 2831–2857.
Fan, L., Lehmann, P., Or, D., 2016. Effects of soil spatial variability at the hillslope and catchment scales on characteristics of rainfall-induced landslides. Water Resour. Res. 52 (3), 1781–1799.
Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874.
Flamig, Z.L., Vergara, H., Gourley, J.J., 2020a. The ensemble framework for flash flood forecasting (EF5) v1.2: description and case study. Geosci. Model Dev. (GMD) 13 (10), 4943–4958. https://doi.org/10.5194/gmd-13-4943-2020.
Flamig, Z.L., Vergara, H., Gourley, J.J., 2020b. The ensemble framework for flash flood forecasting (EF5) v1.2: description and case study. Geosci. Model Dev. (GMD) 13 (10), 4943–4958. https://doi.org/10.5194/gmd-13-4943-2020.
Gourley, J.J., et al., 2017. The FLASH project: improving the tools for flash flood monitoring and prediction across the United States. Bull. Am. Meteorol. Soc. 98 (2), 361–372.
Guo, Z., Torra, O., Hürlimann, M., Abancó, C., Medina, V., 2022. FSLAM: a QGIS plugin for fast regional susceptibility assessment of rainfall-induced landslides. Environ. Model. Software 150, 105354.
He, X., et al., 2016. Development of a coupled hydrological-geotechnical framework for rainfall-induced landslides prediction. J. Hydrol. 543, 395–405. https://doi.org/10.1016/j.jhydrol.2016.10.016.
Hong, Y., Adler, R., Huffman, G., 2006. Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. Geophys. Res. Lett. 33 (22) https://doi.org/10.1029/2006GL028010.
Huang, Q., Long, D., Han, Z., Han, P., 2022. High-resolution satellite images combined with hydrological modeling derive river discharge for headwaters: a step toward discharge estimation in ungauged basins. Remote Sens. Environ. 277, 113030.
Ivanov, V.Y., Vivoni, E.R., Bras, R.L., Entekhabi, D., 2004. Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach. J. Hydrol. 298 (1–4), 80–111.
Iverson, R.M., 2000. Landslide triggering by rain infiltration. Water Resour. Res. 36 (7), 1897–1910.
Jacobs, L., et al., 2020. Regional susceptibility assessments with heterogeneous landslide information: slope unit-vs. pixel-based approach. Geomorphology 356, 107084.
Kolditz, O., et al., 2007. Development of a regional hydrologic soil model and application to the Beerze–Reusel drainage basin. Environ. Pollut. 148 (3), 855–866.
Lenhart, T., Eckhardt, K., Fohrer, N., Frede, H.-G., 2002. Comparison of two different approaches of sensitivity analysis. Phys. Chem. Earth, Parts A/B/C 27 (9–10), 645–654.
Leonarduzzi, E., Maxwell, R.M., Mirus, B.B., Molnar, P., 2021a. Numerical analysis of the effect of subgrid variability in a physically based hydrological model on runoff, soil moisture, and slope stability. Water Resour. Res. 57 (4), e2020WR027326.
Leonarduzzi, E., McArdell, B.W., Molnar, P., 2021b. Rainfall-induced shallow landslides and soil wetness: comparison of physically based and probabilistic predictions. Hydrol. Earth Syst. Sci. 25 (11), 5937–5950.
Li, T., Wang, G., Chen, J., Wang, H., 2011. Dynamic parallelization of hydrological model simulations. Environ. Model. Software 26 (12), 1736–1746.
Li, Z., et al., 2021. CREST-iMAP v1.0: a fully coupled hydrologic-hydraulic modeling framework dedicated to flood inundation mapping and prediction. Environ. Model. Software 141, 105051. https://doi.org/10.1016/j.envsoft.2021.105051.
Liu, J., Zhu, A.-X., Liu, Y., Zhu, T., Qin, C.-Z., 2014. A layered approach to parallel computing for spatially distributed hydrological modeling. Environ. Model. Software 51, 221–227.
Liu, J., Zhu, A.-X., Qin, C.-Z., 2013. Estimation of theoretical maximum speedup ratio for parallel computing of grid-based distributed hydrological models. Comput. Geosci. 60, 58–62.
Lu, N., Godt, J., 2008. Infinite slope stability under steady unsaturated seepage conditions. Water Resour. Res. 44 (11).
Lu, N., Godt, J.W., 2013. Hillslope Hydrology and Stability. Cambridge University Press.
Lu, N., Godt, J.W., Wu, D.T., 2010. A closed-form equation for effective stress in unsaturated soil. Water Resour. Res. 46 (5).
Luo, Y., Zhang, X., Liu, X., Ficklin, D., Zhang, M., 2008. Dynamic modeling of organophosphate pesticide load in surface water in the northern San Joaquin Valley watershed of California. Environ. Pollut. 156 (3), 1171–1181.
Medina, V., Hürlimann, M., Guo, Z., Lloret, A., Vaunat, J., 2021. Fast physically-based model for rainfall-induced landslide susceptibility assessment at regional scale. Catena 201, 105213.
Mergili, M., et al., 2014a. A strategy for GIS-based 3-D slope stability modelling over large areas. Geosci. Model Dev. (GMD) 7 (6), 2969–2982. https://doi.org/10.5194/gmd-7-2969-2014.
Mergili, M., Marchesini, I., Rossi, M., Guzzetti, F., Fellin, W., 2014b. Spatially distributed three-dimensional slope stability modelling in a raster GIS. Geomorphology 206, 178–195. https://doi.org/10.1016/j.geomorph.2013.10.008.
O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. Comput. Vis. Graph Image Process 28 (3), 323–344.
Orgogozo, L., et al., 2014. An open source massively parallel solver for Richards equation: mechanistic modelling of water fluxes at the watershed scale. Comput. Phys. Commun. 185 (12), 3358–3371.
Raia, S., et al., 2014. Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach. Geosci. Model Dev. (GMD) 7 (2), 495–514.
Ren-Jun, Z., 1992. The Xinanjiang model applied in China. J. Hydrol. 135 (1–4), 371–381.
Rossi, G., Catani, F., Leoni, L., Segoni, S., Tofani, V., 2013. HIRESSS: a physically based slope stability simulator for HPC applications. Nat. Hazards Earth Syst. Sci. 13 (1), 151–166.
Shen, X., Hong, Y., Zhang, K., Hao, Z., 2017. Refining a distributed linear reservoir routing method to improve performance of the CREST model. J. Hydrol. Eng. 22 (3), 04016061 https://doi.org/10.1061/(ASCE)HE.1943-5584.0001442.
Sidle, R.C., Greco, R., Bogaard, T., 2019. Overview of Landslide Hydrology. MDPI, p. 148.
Simoni, S., Zanotti, F., Bertoldi, G., Rigon, R., 2008. Modelling the probability of occurrence of shallow landslides and channelized debris flows using GEOtop-FS. Hydrol. Process.: Int. J. 22 (4), 532–545.
Tran, T.V., Alvioli, M., Lee, G., An, H.U., 2018. Three-dimensional, time-dependent modeling of rainfall-induced landslides over a digital landscape: a case study. Landslides 15 (6), 1071–1084. https://doi.org/10.1007/s10346-017-0931-7.
Tufano, R., Formetta, G., Calcaterra, D., De Vita, P., 2021. Hydrological control of soil thickness spatial variability on the initiation of rainfall-induced shallow landslides using a three-dimensional model. Landslides 18 (10), 3367–3380.
van Dam, J.C., Groenendijk, P., Hendriks, R.F., Kroes, J.G., 2008. Advances of modeling water flow in variably saturated soils with SWAP. Vadose Zone J. 7 (2), 640–653.
Vandromme, R., Thiery, Y., Bernardie, S., Sedan, O., 2020. ALICE (Assessment of Landslides Induced by Climatic Events): a single tool to integrate shallow and deep landslides for susceptibility and hazard assessment. Geomorphology 367, 107307.
Velásquez, N., Hoyos, C.D., Vélez, J.I., Zapata, E., 2020. Reconstructing the 2015 Salgar flash flood using radar retrievals and a conceptual modeling framework in an ungauged basin. Hydrol. Earth Syst. Sci. 24 (3), 1367–1392.
Vivoni, E.R., et al., 2011. Real-world hydrologic assessment of a fully-distributed hydrological model in a parallel computing environment. J. Hydrol. 409 (1–2), 483–496.
Vivoni, E.R., et al., 2005. Parallelization of a fully-distributed hydrologic model using sub-basin partitioning. Eos Trans. AGU 86, 52.
Wang, H., Zhou, Y., Fu, X., Gao, J., Wang, G., 2012. Maximum speedup ratio curve (MSC) in parallel computing of the binary-tree-based drainage network. Comput. Geosci. 38 (1), 127–135.
Wang, J., et al., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrol. Sci. J. 56 (1), 84–98. https://doi.org/10.1080/02626667.2010.543087.
Wang, S., Zhang, K., van Beek, L.P.H., Tian, X., Bogaard, T.A., 2020. Physically-based landslide prediction over a large region: scaling low-resolution hydrological model results for high-resolution slope stability assessment. Environ. Model. Software 124, 104607. https://doi.org/10.1016/j.envsoft.2019.104607.
Wieder, W.R., Boehnert, J., Bonan, G.B., Langseth, M., 2014. Regridded Harmonized World Soil Database V1. 2. ORNL DAAC.
Wu, H., Adler, R.F., Hong, Y., Tian, Y., Policelli, F., 2012. Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. J. Hydrometeorol. 13 (4), 1268–1284.
Xie, M., Esaki, T., Cai, M., 2004. A GIS-based method for locating the critical 3D slip surface in a slope. Comput. Geotech. 31 (4), 267–277.
Xie, M., Esaki, T., Qiu, C., Wang, C., 2006. Geographical information system-based computational implementation and application of spatial three-dimensional slope stability analysis. Comput. Geotech. 33 (4–5), 260–274.
Xu, Z., et al., 2021. An automatic partition-based parallel algorithm for grid-based distributed hydrological models. Environ. Model. Software 144, 105142.
Xue, X., et al., 2013. Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the wangchu basin of Bhutan: are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? J. Hydrol. 499, 91–99. https://doi.org/10.1016/j.jhydrol.2013.06.042.
Zhang, K., Wang, S., Bao, H., Zhao, X., 2019. Characteristics and influencing factors of rainfall-induced landslide and debris flow hazards in Shaanxi Province, China. Nat. Hazards Earth Syst. Sci. 19 (1), 93–105.
Zhang, K., et al., 2016. iCRESTRIGRS: a coupled modeling system for cascading flood-landslide disaster forecasting. Hydrol. Earth Syst. Sci. 20 (12), 5035–5048. https://doi.org/10.5194/hess-20-5035-2016.
Zhu, L.-J., Liu, J., Qin, C.-Z., Zhu, A.-X., 2019. A modular and parallelized watershed modeling framework. Environ. Model. Software 122, 104526.