Supplementary Material for

Multi-Machine Learning Ensemble Regionalization of Hydrological Parameters for Enhances Flood Prediction in Ungauged Mountainous Catchments

Kai Li, Linmao Guo, Genxu Wang*, Jihui Gao*, Xiangyang Sun, Peng Huang, Jinlong Li, Jiapei Ma, Xinyu Zhang

State Key Laboratory of Hydraulics and Mountain River Engineering, College of Water Resource and Hydropower, Sichuan University, Chengdu, 610000, China

*Corresponding author: Genxu Wang (wanggx@scu.edu.cn) and Jihui Gao (jgao@scu.edu.cn).

The Topography-based Subsurface Storm Flow Hydrological Model (Top-SSF model)

The Topography-based Subsurface Storm Flow Hydrological Model (Top-SSF model) is a process-based model developed to simulate the hydrological response of mountainous catchments, with a particular emphasis on flash flood. The model structure (Fig. S1) and its key components are detailed in the subsequent sections.

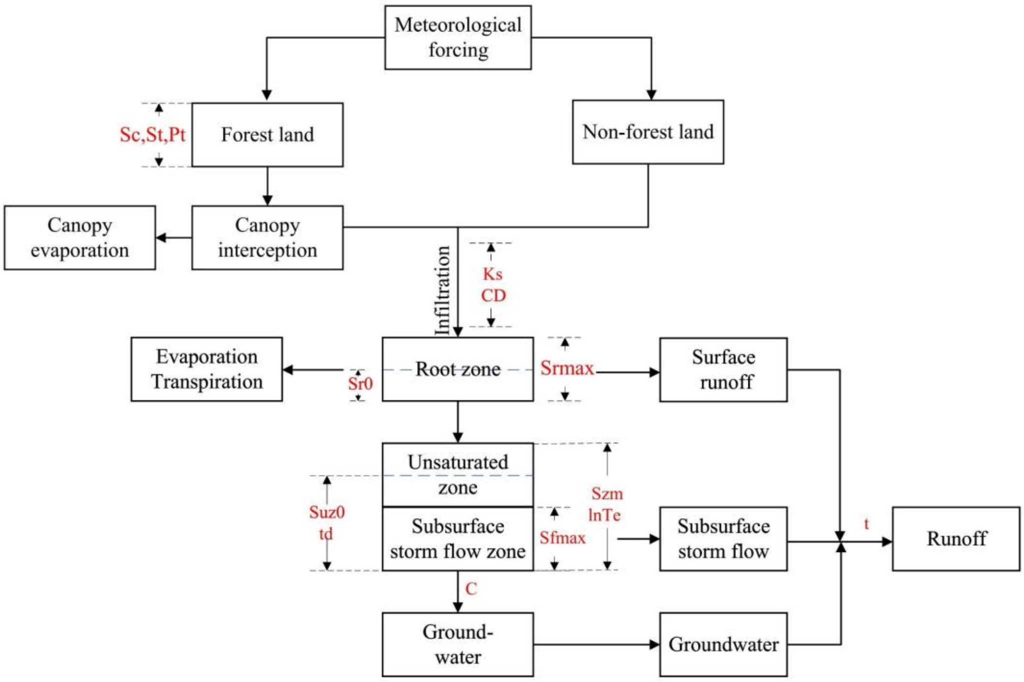

Fig.S1. Schematic diagram of the Top-SSF model structure

1. Canopy Interception

Canopy interception is calculated based on measured rainfall data and forest cover characteristics. The process is divided into three distinct phases: canopy wetting, canopy saturation, and canopy drying. In the Top-SSF model, the 1995 Gash model (Gash et al., 1995) was modified and used as the canopy interception module. The improved parts are as follows.

During the canopy humidification period, (1) the total interception equation for calculating the rainfall events was converted to the hourly canopy interception equation (Eq. 3), and (2) the total trunk runoff equation for calculating rainfall events was converted to the hourly trunk runoff equation (Eq. 4).

$$P_g' = -(\overline{R}/\overline{E})S_c \ln(1 - \overline{R}/\overline{E}) \quad (1)$$

$$P_g'' = \overline{R}/(\overline{R} - \overline{E})(S_t/P_t) + P_g' \quad (2)$$

$$I(t) = \begin{cases} cP_g(t) & (P_g(t) < P_g') \\ cP_g' + c\overline{E}(P_g(t) - P_g')/\overline{R} + S_t & (P_g(t) \geq P_g'') \\ cP_g' + cP_t(1 - \overline{E}/\overline{R})(P_g(t) - P_g') + c\overline{E}(P_g(t) - P_g')/\overline{R} & (P_g(t) > P_g', P_g(t) < P_g'') \end{cases} \quad (3)$$

$$SF(t) = \begin{cases} 0 & (P_g(t) < P_g') \\ cP_t(1 - \overline{E}/\overline{R})(P_g(t) - P_g') - cS_t & (P_g(t) \geq P_g'') \\ 0 & (P_g(t) > P_g', P_g(t) < P_g'') \end{cases} \quad (4)$$

where: P_g' is the minimum rainfall required for the canopy to reach saturation (mm); $\overline{R}$ is the average rainfall intensity (mm/h); $\overline{E}$ is the average potential evaporation rate of the canopy (mm/h); S_c is the canopy storage capacity (mm); P_g'' is the minimum rainfall needed in the trunk to reach saturation (mm); S_t is the trunk storage capacity (mm); P_t is the trunk runoff coefficient (%); $I(t)$ is the canopy interception (mm); $P_g(t)$ is the rainfall (mm); $SF(t)$ is the trunk runoff (mm); and c is the forest canopy closure (%), which is equal to the forest cover.

During the canopy saturation period, canopy interception and trunk interception are equal to zero, and canopy evaporation can be estimated as potential evapotranspiration using the Penman–Monteith equation (Rutter et al., 1971).

During the canopy dry period, the original Gash model assumes that when the canopy is completely dry, the drying time exceeds 8 hours (Gash et al., 1995). In the Top-SSF model, Eq. 9 was used to calculate the hourly canopy evaporation:

$$E(t) = E_p(t)\left(\frac{C_h(t)}{Sc}\right) \quad (5)$$

where $E(t)$ for actual canopy evaporation (mm); $C_h(t)$ is the depth of water held on canopy at time t (mm).

2. Soil Infiltration

In this study, infiltration is simulated using the Green-Ampt model. When surface ponding occurs, the infiltration rate is determined by solving the Green-Ampt equation iteratively, for which the Newton-Raphson method is employed. The infiltration rate (f_{in}) is given by：

$$f_{in} = -\frac{Ks(CD+F_{satrt})}{Szm(1-\exp(F_{satrt}/Szm))} \quad (5)$$

where, f_{in} is the infiltration rate (m/h)； Ks is surface hydraulic conductivity (m/h); CD is capillary drive (m); F_{satrt} is the initial cumulative infiltration (m); Szm is the maximum water storage capacity in the unsaturated zone (m).

3. Runoff Generation and Storage Dynamics

3.1. Soil Evaporation

$$E_a = E_{pt}(1 - \frac{Srz}{Srmax}) \quad (7)$$

where, E_a is the Actual soil evapotranspiration (m); E_{pt} is the potential evapotranspiration (m); Srz is the root zone water deficit (m); $Srmax$ is the maximum water storage capacity of the root zone (m).

3.2. Overland Flow

Overland flow in the Top-SSF model consists of saturation-excess and infiltration-excess components.

Saturation-excess flow: Occurs when groundwater table depth $S_i \leq 0$ at computational cell i:

$$r_{s,i} = max\{Suz_i - \max(S_i, 0), 0\} \quad (8)$$

where， $r_{s,i}$ is the depth of saturation excess overland flow generated at cell i (m)； Suz_i is the soil water storage in the unsaturated zone, at cell i (m)； S_i is the groundwater table depth at cell i (m).

Infiltration-excess flow: Activated when rainfall intensity exceeds soil infiltration capacity.

3.3. Subsurface storm flow

Water deficit in subsurface storm flow zone ($S_{sf,i}$) is determined by topographic controls:

$$S_{sf,i} = S_{fmax} - \frac{\left(\frac{a}{tan\,\beta}\right)A_i}{\int_A \left(\frac{a}{tan\,\beta}\right) dA_i}(S_{fmax} - \overline{S}_{sf}) \quad (9)$$

where, $S_{sf,i}$ is the water deficit in the subsurface storm flow zone at cell i (m); S_{fmax} is the maximum subsurface storm flow zone deficit (m); $\frac{a}{tan\,\beta}$ is the subsurface topographic index (-); $\overline{S}_{sf}$ is the average water deficit in the subsurface storm flow zone (m);A_i is the percentage of the catchment area occupied by cell i (%).

The unsaturated zone recharges the subsurface storm flow zone：

$$r_{v,i} = \frac{Suz_i}{S_i t_d} \quad (10)$$

where, $r_{v,i}$ is the depth of unsaturated zone recharges the subsurface storm flow zone at cell i (m); t_d is the unsaturated zone time delay per unit storage deficit (h/m).

The depth of storm subsurface flow generated at computational cell i , $r_{sf,i}$ is given by:

$$r_{sf,i} = q_{sf0}(1 - S_{sf,i}/S_{fmax}) \quad (11)$$

where, $r_{sf,i}$ is the depth of storm subsurface flow at cell i (m); q_{sf0} is initial subsurface storm flow (m); $S_{sf,i}$ is the water storage deficit in the storm subsurface flow zone at cell i (m).

The subsurface storm flow recharges the groundwater：

$$r_{g,i} = \min\left(C\left(S_{famx} - S_{sf,i}\right), S_i\right) \quad (12)$$

where, $r_{g,i}$ is the subsurface storm flow recharge groundwater at (m); C is the transfer coefficient (m^2/h).

The average water deficit of subsurface storm flow zone ($\overline{S}_{sf}$) and the average depth of groundwater ($\overline{S}_g$) in the catchment are updated as follows:

$$\Delta\overline{S}_{sf}/\Delta t = -\sum_{i=1}^{M} r_{v,i}\,\mathrm{A}_i + \sum_{i=1}^{M} r_{sf,i}\,\mathrm{A}_i + \sum_{i=1}^{M} r_{g,i}\,\mathrm{A}_i \quad (13)$$

$$\Delta\overline{S}_g/\Delta t = -\sum_{i=1}^{M} r_{g,i}\,\mathrm{A}_i + r_b \quad (14)$$

where, $\Delta\overline{S}_{sf}$ is the change in the average subsurface storm flow zone (m); M is the total number of computational cells; $\Delta\overline{S}_g$ is the change in the average groundwater level (m);Δt is the time step (h);

3.4. Groundwater Flow

The depth of groundwater discharge is calculate as：

$$r_b = e^{\mathrm{lnTe}-\lambda-\overline{S}_g/Szm} \qquad (15)$$

where，r_b is depth of groundwater discharge (m);$lnTe$ is the log of the areal average of $T0$ (m^2/h); is the catchment average topographic index；$\overline{S}_g$ is the catchment average groundwater table depth (m).

4. Flow Routing

Catchment response time calculation：

$$T_{c,j} = t\sum_{k=1}^{j}\left(\frac{0.87L_{ch,n}}{1000S_{ch,n}}\right)^{0.385} \qquad (16)$$

where N is the number of river subsections within the catchment; $L_{ch,n}$ is the length of the river channel (km); $S_{ch,n}$ is the slope of the river segment (m·m^{-1}); and t is the time-correction coefficient (-).

For any simulation time step t, the proportion of the catchment area contributing to the flow at the outlet is determined. If the simulation time t is greater than or equal to the time of concentration for the catchment, $T_{c,N}$ (i.e., the time of concentration from the most hydrologically distant point), then the entire catchment area is assumed to be contributing. Otherwise, If the simulation time t is less than $T_{c,N}$, the catchment is partially contributing. The proportion of the catchment area, contributing to the outlet flow at time t is calculated by linear interpolation between isochrones:

$$AR_t = ACH_{j-1} + \frac{t-T_{c,j-1}}{T_{c,j}-T_{c,j-1}}(ACH_j - ACH_{j-1}) \qquad (17)$$

where, AR_t is the proportion of the catchment area contributing to outlet flow at time t (%);$T_{c,j}$ and $T_{c,j-1}$ are the travel times defining the boundaries of the j-th and $(j-1)$-th isochrones, respectively (h); $ACH_j$and $ACH_{j-1}$are the cumulative proportions of the total catchment area enclosed by the j-th and $(j-1)$-th isochrones, respectively (%).

References

Gash, J.H.C., Lloyd, C.R.Lachaud, G. (1995). Estimating sparse forest rainfall interception with an analytical model. Journal of Hydrology, 170(1): 79-86. https://doi.org/10.1016/0022-1694(95)02697-N

Rutter, A.J., Kershaw, K.A., Robins, P.C.Morton, A.J. (1971). A predictive model of rainfall interception in forests, 1. Derivation of the model from observations in a plantation of Corsican pine. Agricultural Meteorology, 9: 367-384. https://doi.org/10.1016/0002-1571(71)90034-3