A Unified Scheme for Modeling Saturation and Infiltration Excess Runoff

Yuanqi Hong[1], Guta Wakbulcho Abeshu[2], Hong-Yi Li[1], Dingbao Wang[3], Mengqi Zhao[4], Thomas Wild[5], Günter Blöschl[6], and L. Ruby Leung[2]

[1]Department of Civil and Environmental Engineering, University of Houston, Houston, 77204, USA
[2]Atmospheric, Climate, & Earth Sciences Division, Pacific Northwest National Laboratory, Richland, 99354, USA
[3]Department of Civil, Environmental and Construction Engineering, University of Central Florida, Orlando, 32816, USA
[4]Earth and Environmental Directorate, Pacific Northwest National Laboratory, Richland, 99354, USA
[5]Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, 20740, USA
[6]Institute of Hydraulic and Water Resources Engineering, Vienna University of Technology, Vienna, 1040, Austria

Correspondence to: Hong-Yi Li (hongyili.jadison@gmail.com)

1 Mathematical derivation of URSSIE

This section provides a detailed, step-by-step derivation of the equations presented in the main manuscript for the newly developed URSSIE runoff generation scheme (Section 2). Note that, in this supplementary section, we include some equations essentially the same as those in the main text for better completeness. The following subsections will detail each equation's derivation, key assumptions, and mathematical techniques employed.

1.1 Spatial distribution of soil water storage capacity

We begin our derivation with cumulative distribution function (CDF) of soil water storage capacity proposed by Wang (2018), i.e., Eq. (S1):

$$F(C) = 1 - \frac{1}{a} + \frac{C+(1-a)S_b}{a\sqrt{(C+S_b)^2 - 2aS_bC}} \tag{S1}$$

Where C is the point-scale water storage capacity (m) supported on a positive semi-infinite interval ($C \geq 0$), a is the shape parameter (-) defined within the interval (0,2], and S_b is the maximum value of average storage capacity (m) over the catchment.

1.2 Spatial distribution of soil infiltration capacity

Building on Eq. (S1), we next derive a spatial distribution function of soil infiltration capacity at the catchment scale. We assume the one-to-one relationship between the point-scale infiltration capacity f_c (m/s), and the soil water deficit D (in meters), can be expressed as a power-law type formula:

$$f_c = m_k \left(\frac{D}{S_b}\right)^n \tag{S2}$$

The coefficient m_k sets the upper limit of infiltration capacity (m/s) under extreme rainfall conditions. The dimensionless exponent, n, reflects the spatial heterogeneity of landscape conditions within a catchment. We rearrange Eq. (S2), to explicitly express water deficit D in terms of f_c as:

$$D = f_c^{\frac{1}{n}} m_k^{-\frac{1}{n}} S_b \tag{S3}$$

Meanwhile, D can also be determined by the difference between the total soil storage capacity C (in meters) and the initial storage C_0 (in meters):

$$D = C - C_0 \tag{S4}$$

By equating Eq. (S3) to Eq. (S4) and solving for C, we obtain:

$$C = C_0 + f_c^{\frac{1}{n}} m_k^{-\frac{1}{n}} S_b \tag{S5}$$

Substituting Eq. (S5) into Eq. (S1), we obtain the CDF of infiltration capacity $H(f_c)$:

$$H(f_c) = 1 - \frac{1}{a} + \frac{C_0 + f_c^{\frac{1}{n}} m_k^{-\frac{1}{n}} S_b + (1-a) S_b}{a\sqrt{\left(C_0 + f_c^{\frac{1}{n}} m_k^{-\frac{1}{n}} S_b + S_b\right)^2 - 2aS_b\left(C_0 + f_c^{\frac{1}{n}} m_k^{-\frac{1}{n}} S_b\right)}} \tag{S6}$$

$H(f_c)$ can be used to delineate the fraction of the catchment area that is saturated when evaluated at the onset of a hydrological event. At the initial time step, if a portion of the catchment is already saturated, i.e., $D = 0$, we get $f_c = 0$. Substituting $f_c = 0$ into Eq. (S6) yields α_0, which represents the fraction of the catchment that is fully saturated at the beginning of the event. On the CDF plot, $H(f_c) \leq \alpha_0$ indicates the antecedent saturation subzone ($A_{0\sim\alpha_0}$), while $H(f_c) > \alpha_0$ identifies areas that are unsaturated. Depending on the rainfall intensity i (m/s), for the unsaturated areas, actual infiltration rate, f (m/s), can be expressed as:

$$f = \begin{cases} i, & if\ i \leq f_c \\ f_c, & if\ i > f_c \end{cases} \tag{S7}$$

The infiltration capacity f_c, is not static but rather evolves during and after a rainfall event, necessitating regular updates to reflect the dynamic soil-water interactions. The initial soil water deficit, D_0, is established at time (in seconds) $t = 0$, marking the onset of the event. As rainfall continues, the cumulative infiltration F (in meters), increases, signifying the absorption of water by the soil. To accurately track the changing soil moisture conditions, f_c is updated by the following relationship:

$$f = m_k \left(\frac{D_0 - F}{S_b}\right)^n \tag{S8}$$

At time $t = 0$, the initial infiltration rate f_0 and the cumulative infiltration F_0 (in meters) are given by:

$$f_0 = m_k \left(\frac{D_0}{S_b}\right)^n \tag{S9}$$

$$F_0 = 0 \tag{S10}$$

At saturation time t_s (in seconds), which represents the time lapse from the onset of the event to when the soil is saturated, the following can be summarized with regard to f and F:

$$f(t = t_s) = 0 \tag{S11}$$

$$F(t = t_s) = D_0 \tag{S12}$$

To calculate the change in the infiltration rate in the unsaturated subzones, we can take the derivative of Eq. (S8). However, since the cumulative infiltration F is a function of time t (in seconds), we need to apply the chain rule to account for this dependency:

$$\frac{df}{dt} = \frac{df}{dF}\frac{dF}{dt} \tag{S13}$$

From Eq. (S8), we get df/dF as:

$$\frac{df}{dF} = \left(\frac{D_0 - F}{S_b}\right)^{n-1}\left(-\frac{m_k n}{S_b}\right) \tag{S14}$$

dF/dt is the rate of infiltration, which is f itself. Therefore, substituting it into Eq. (S14) gives the change of infiltration rate with time as:

$$\frac{df}{dt} = f\left(\frac{D_0 - F}{S_b}\right)^{n-1}\left(-\frac{m_k n}{S_b}\right) \tag{S15}$$

Eq. (S15) shows how the rate of change of the infiltration rate with respect to time depends on both the rate of infiltration itself and how close the cumulative infiltration, F is to saturate the soil's storage capacity, S_b. As F increases, this derivative typically becomes more negative, indicating a decreasing infiltration rate over time as the soil approaches saturation. Rearranging Eq. (S8) and raising both sides to the power of $n - 1$, we obtain:

$$\left(\frac{D_0 - F}{S_b}\right)^{n-1} = \left(\frac{f}{m_k}\right)^{\frac{n-1}{n}} \tag{S16}$$

Substituting Eq. (S16) into Eq. (S15) and reorganizing them, we get:

$$f^{\frac{1-2n}{n}}\,df = -\frac{n}{S_b} m_k^{\frac{1}{n}}\,dt \tag{S17}$$

Integrating both sides of Eq. (S17), from $t = 0 \; to \; t$ for the right-hand side and $f = f_0 \; to \; f$ for the left-hand side, we get:

$$f = \left(f_0^{\frac{1-n}{n}} + \frac{n-1}{S_b} m_k^{\frac{1}{n}}\, t\right)^{\frac{n}{1-n}} \tag{S18}$$

Recall that f_0 corresponds to the infiltration rate at t = 0 and is given by Eq. (S9). By substituting Eq. (S9) into Eq. (S18), we obtain a complete expression for the infiltration rate as a function of time:

$$f = \left(m_k^{\frac{1-n}{n}}\left(\frac{D_0}{S_b}\right)^{1-n} + \frac{n-1}{S_b} m_k^{\frac{1}{n}}\, t\right)^{\frac{n}{1-n}} \tag{S19}$$

If the initial infiltration capacity of a spatial point at a timestep is lower than rainfall intensity i, infiltration excess occurs right at the beginning of the timestep. The saturation time under this condition, denoted as t_{s1}, represents the interval from the start of the time step to the moment when the soil becomes fully saturated. At saturation, the infiltration rate f reduces to zero. Substituting $f = 0$ and $t = t_{s1}$ into Eq. (S19), we can derive the saturation time t_{s1} for this condition as follows:

$$t_{s1} = \frac{D_0{}^{1-n} S_b{}^n}{m_k(1-n)} \tag{S20}$$

If the initial infiltration capacity is greater than the rainfall intensity i, the onset of infiltration excess occurs only when f_c decreases to i. At this point, soil reaches a critical water deficit state, D_i (in meters), which corresponds to conditions where $f_c = i$. This critical state can be mathematically expressed as follows:

$$f = i = m_k (D_i/S_b)^n \tag{S21}$$

Solving for the critical water deficit state D_i, we get:

$$D_i = \left(\frac{i}{m_k}\right)^{\frac{1}{n}} S_b \tag{S22}$$

Eqs. (S21) and (S22) establish the conditions under which the soil's ability to absorb water matches the ongoing rainfall intensity, marking the critical threshold at which infiltration capacity begins to limit the rate of rainfall absorption. The critical point corresponds to α^*:

$$\alpha^* = 1 - \frac{1}{a} + \frac{D_i + C_0 + (1-a)S_b}{a\sqrt{(D_i + C_0 + S_b)^2 - 2aS_b(D_i + C_0)}} \tag{S23}$$

Between the onset of an event and reaching the critical water deficit state D_i, infiltration excess runoff does not commence immediately; rather, it requires some time for water to accumulate, which is referred to as the ponding time, t_p (in seconds). t_p represents the duration from the start of the rainfall event to the beginning of infiltration excess. It is the necessary time for enough water to accumulate on the surface, exceeding the soil's infiltration capacity reduced to the level of the rainfall intensity i. t_p can be calculated as:

$$t_p = \frac{D_0 - D_i}{i} \tag{S24}$$

Once ponding begins, the soil continues to absorb water until it reaches saturation. The saturation time, denoted as t_{s2} (in seconds), includes the ponding time t_p and additional time required for the soil to become fully saturated after ponding starts. This total saturation time can be computed as:

$$t_{s2} = t_{si} + t_p \tag{S25}$$

Where t_{si} is the saturation time corresponding to D_i and can be obtained by swapping D_0 by D_i in Eq. (S20):

$$t_{si} = \frac{D_i{}^{1-n} S_b{}^n}{m_k(1-n)} \tag{S26}$$

The relationship between the saturation time for the critical point (t_{si}) and the timestep of interest Δt (in seconds) determines the infiltration excess behavior in different subzones of the catchment. We can consider two distinct conditions based on this relationship.

i) *Condition-I* ($t_{si} > \Delta t$): Under this condition, the infiltration excess starts at the beginning of Δt in the $A_{\alpha_0 \sim \alpha_1}$ subzone. In the $A_{\alpha_1 \sim \alpha_2}$ subzone, infiltration excess starts at the beginning of some portion of the subzone, while the infiltration excess starts after ponding time at the rest.

ii) *Condition-II* ($t_{si} < \Delta t$): Infiltration excess starts at the beginning of Δt over part of the $A_{\alpha_0 \sim \alpha_1}$ subzone, while the infiltration excess starts after ponding time over the rest. In the $A_{\alpha_1 \sim \alpha_2}$ subzone, infiltration excess starts after ponding time.

1.3 Partitioning the spatial unit into subzones based on infiltration dynamics

The dynamic nature of soil water deficit results in the formation of four distinct spatial subzones within a given spatial unit of interest (i.e., catchment or grid), each characterized by different infiltration capacities and responses to rainfall (Fig. 1). Here we describe how these subzones are distinguished:

a) *Antecedent saturation subzone* ($A_{0\sim\alpha_0}$): Infiltration capacity is zero, and all the rainfall becomes saturation excess runoff. The boundary α_0 for this subzone is determined by substituting $f_c = 0$ into Eq. (S6):

$$\alpha_0 = 1 - \frac{1}{a} + \frac{C_0 + (1-a)S_b}{a\sqrt{(C_0 + S_b)^2 - 2aS_bC_0}} \quad \text{(S27)}$$

b) *Event-induced saturation subzone* ($A_{\alpha_0\sim\alpha_1}$): This subzone encompasses the areas that will become saturated only during the current rainfall event (or time step). The boundary α_1 for this subzone is where the $t_{s1} = \Delta t$ (for *Condition-I*) or $t_{s2} = \Delta t$ (for *Condition-II*).

For *Condition-I*, using Eq. (S20), the initial soil water deficit at α_1 is

$$\mathrm{D}_{0\alpha_1} = \left[\frac{\Delta t m_k(1-n)}{{S_b}^n}\right]^{\left(\frac{1}{1-n}\right)} \quad \text{(S28)}$$

For *Condition-II*, using Eq. (S25), the initial soil water deficit at α_1 is

$$\mathrm{D}_{0\alpha_1} = \left[\Delta t - \frac{{D_i}^{1-n}{S_b}^n}{m_k(1-n)}\right] i + D_i \quad \text{(S29)}$$

Substituting Eq. (S28) (for *Condition-I*) or Eq. (S29) (for *Condition-II*) into Eq. (S6) provides the upper boundary of this subzone:

$$\alpha_1 = 1 - \frac{1}{a} + \frac{\mathrm{D}_{0\alpha_1} + C_0 + (1-a)S_b}{a\sqrt{\left(\mathrm{D}_{0\alpha_1} + C_0 + S_b\right)^2 - 2aS_b\left(\mathrm{D}_{0\alpha_1} + C_0\right)}} \quad \text{(S30)}$$

c) *Partial event-induced infiltration excess subzone* ($A_{\alpha_1\sim\alpha_2}$): In this subzone, rainfall intensity exceeds the infiltration capacity for either the entire time step or a portion of it, yet the soil remains unsaturated throughout the event. Consequently, only infiltration excess runoff occurs. The upper boundary α_2 of this subzone is determined by setting t_p = Δt. Adjusting Eq. (S24) for D_0 and substituting Eq. (S22) for D_i yields initial soil water deficit at α_2.

$$D_{0\alpha_2} = \Delta \mathrm{t}\, i + \left(\frac{i}{m_k}\right)^{\frac{1}{n}} S_b \quad \text{(S31)}$$

Substituting Eq. (S31) into Eq. (S6), yields the upper boundary of this subzone:

$$\alpha_2 = 1 - \frac{1}{a} + \frac{\mathrm{D}_{0\alpha_2} + C_0 + (1-a)S_b}{a\sqrt{\left(\mathrm{D}_{0\alpha_2} + C_0 + S_b\right)^2 - 2aS_b\left(\mathrm{D}_{0\alpha_2} + C_0\right)}} \quad \text{(S32)}$$

d) *Infiltration dominated subzone* ($A_{\alpha_2 \sim 1}$): This subzone represents the areas where the soil's infiltration capacity consistently exceeds the rainfall intensity. Consequently, these areas neither experience saturation nor contribute to excess runoff infiltration. The dynamics in this subzone are primarily driven by the inherent soil properties and rainfall characteristics that allow the soil to absorb water efficiently throughout the event. This results in a steady infiltration process without the onset of ponding or surface runoff.

1.4 Spatial distribution of runoff generation

Depending on the subzone, runoff generation mechanisms and estimations differ. Below, we provide details on how runoff is computed.

a) $A_{0 \sim \alpha_0}$: This subzone is already saturated when rain starts; hence, all the rainfall becomes saturation excess runoff (m) and is computed as.

$$R_{s0} = P\alpha_0 \tag{S33}$$

Where P is the precipitation depth (m) in one timestep.

b) $A_{\alpha_0 \sim \alpha_1}$:

i) Condition -I

The saturation excess occurs after the saturation time t_{s1}, which can be computed for any point in this subzone as:

$$r_{s1} = i(\Delta t - t_{s1}) \tag{S34}$$

Substituting Eq. (S20) into Eq. (S34) and then substituting into Eq. (S6) yields a function for the spatial distribution of saturation excess in $A_{\alpha_0 \sim \alpha_1}$:

$$K_1(r_{s1}) = 1 - \frac{1}{a} + \frac{X_1}{a\sqrt{X_2 - X_3}} \tag{S35}$$

Where:

$$X_1 = \left[\frac{\left(\Delta t - \frac{r_{s1}}{i}\right) m_k (1-n)}{{S_b}^n}\right]^{\left(\frac{1}{1-n}\right)} + C_0 + (1-a)S_b \tag{S35-a}$$

$$X_2 = \left(\left[\frac{\left(\Delta t - \frac{r_{s1}}{i}\right) m_k (1-n)}{{S_b}^n}\right]^{\left(\frac{1}{1-n}\right)} + C_0 + S_b\right)^2 \tag{S35-b}$$

$$X_3 = 2aS_b\left(\left[\frac{\left(\Delta t - \frac{r_{s1}}{i}\right) m_k (1-n)}{{S_b}^n}\right]^{\left(\frac{1}{1-n}\right)} + C_0\right) \tag{S35-c}$$

Total saturation excess generated within this subzone is obtained by integrating Eq. (S35)

$$R_{sc1} = -P\alpha_0 + \int_0^P K_1(r_{s1})\, dr_s \tag{S36}$$

ii) Condition -II

For $\alpha_0 \sim \alpha^*$, the spatial distribution of saturation excess can be obtained by Eq. (S35). For $\alpha^* \sim \alpha_1$, the saturation excess after saturation time t_{s2} , at any spatial point within this area is computed as

$$r_{s2} = i(\Delta t - t_{s2}) \tag{S37}$$

Substituting Eq. (S37) into Eq. (S6) along with Eq. (S25), the spatial distribution of saturation excess is given by:

$$K_2(r_{s2}) = 1 - \frac{1}{a} + \frac{X_4}{a\sqrt{X_5 - X_6}} \tag{S38}$$

Where:

$$X_4 = \left(\Delta t - \frac{r_{s2}}{i} - \frac{{D_i}^{1-n}{S_b}^n}{m_k(1-n)}\right) i + D_i + C_0 + (1-a)S_b \tag{S38-a}$$

$$X_5 = \left(\left(\Delta t - \frac{r_{s2}}{i} - \frac{{D_i}^{1-n}{S_b}^n}{m_k(1-n)}\right) i + D_i + C_0 + S_b\right)^2 \tag{S38-b}$$

$$X_6 = 2aS_b\left(\left(\Delta t - \frac{r_{s2}}{i} - \frac{{D_i}^{1-n}{S_b}^n}{m_k(1-n)}\right) i + D_i + C_0\right) \tag{S38-c}$$

Total saturation excess generated within this subzone is obtained by integrating Eq. (S38)

$$R_{sc2} = -P\alpha_0 + \int_{i(\Delta t - t_{si})}^{P} K_1(r_{s1})\, dr_s + \int_{0}^{i(\Delta t - t_{si})} K_2(r_{s2})\, dr_s \tag{S39}$$

The infiltration excess runoff (m) for this subzone can be computed from the water balance as:

i) Condition I

$$R_{i1} = P(\alpha_1 - \alpha_0) - R_{sc1} - W_1 \tag{S40-a}$$

ii) Condition II

$$R_{i1} = P(\alpha_1 - \alpha_0) - R_{sc2} - W_1 \tag{S40-b}$$

Where W_1 is total soil wetting (m) within this subzone.

c) $A_{\alpha_1 \sim \alpha_2}$:

i) Condition-I:

Here D_i is located within the $A_{\alpha_1 \sim \alpha_2}$. It divides the subzone into two parts, $D_1 \sim D_i$ (i.e., $\alpha_1 \sim \alpha^*$) and $D_i \sim D_2$ (i.e., $\alpha^* \sim \alpha_1$). Infiltration excess computation differs between these two parts.

For $\alpha_1 \sim \alpha^*$, infiltration excess starts at the beginning of the current time step, and the infiltration excess runoff rate at a point is given by:

$$\alpha_1 < \alpha^* < \alpha_2$$

$$r_{i(\alpha_1 \sim \alpha^*)} = i\Delta t - \int_0^{\Delta t} f dt \tag{S41}$$

$$r_{i(\alpha_1 \sim \alpha^*)} = i\Delta t + \frac{S_b}{m_k^{\frac{1}{n}}}\left[m_k^{\frac{1-n}{n}}\left(\frac{D_0}{S_b}\right)^{1-n} - \frac{m_k^{\frac{1}{n}}(1-n)\Delta t}{S_b}\right]^{\frac{1}{1-n}} - D_0 \tag{S42}$$

For $\alpha^* \sim \alpha_2$,infiltration excess starts after ponding time, and infiltration excess runoff rate at a point is:

$$r_{i_{(\alpha^*\sim\alpha_2)}} = i(\Delta t - t_p) - \int_0^{\Delta t - t_p} f dt \tag{S43}$$

$$r_{i_{(\alpha^*\sim\alpha_2)}} = i\Delta t - D_0 + \frac{S_b}{m_k^{\frac{1}{n}}}\left[m_k^{\frac{1-n}{n}}\left(\frac{D_i}{S_b}\right)^{1-n} - \frac{m_k^{\frac{1}{n}}(1-n)\left(\Delta t - \frac{D_0 - D_i}{i}\right)}{S_b}\right]^{\frac{1}{1-n}} \tag{S44}$$

Total infiltration excess runoff in $A_{\alpha_1\sim\alpha_2}$ under this condition (R_{i2}) is the sum of infiltration excess from ($\alpha_1 \sim \alpha^*$) and ($\alpha^*\sim\alpha_2$) , and can be calculated by applying the finite difference method on Eq. (S42) and Eq. (S44).

ii) Condition-II:

D_i is located within the $A_{\alpha_0\sim\alpha_1}$. Infiltration excess at any spatial point in $A_{\alpha_1\sim\alpha_2}$ therefore starts after t_p and is calculated by Eq. (S44). Total infiltration excess runoff in $A_{\alpha_1\sim\alpha_2}$ area (R_{i2}) can be calculated by applying the finite difference method on Eq. (S44) only.

d) $A_{\alpha_2\sim 1}$:

No runoff is generated from this subzone.

$$R_s = R_i = 0 \tag{S45}$$

2 Additional details of demonstrating URSSIE over the U.S.

This section describes details in demonstrating the new runoff scheme, URSSIE, over the natural catchments in the contiguous U.S. through a simple hydrologic model. We provide the technical details of the simple hydrologic model, its application over the natural catchments (e.g., input data preparation, catchment selection, parameter determination), and additional modeling results (that are not included in the main text).

2.1 Technical details of the simple hydrologic model

Figure S1 is a conceptual diagram of the hydrologic model, which is designed by essentially expanding the PDM model by Yao and Wang (2022). The following equations describe how to compute the state and flux variables described in Fig. S1.

Total infiltration excess runoff R_i:

$$R_i = R_{i_1} + R_{i_2} \tag{S46}$$

Total saturation excess R_s:

$$R_s = R_{s_0} + R_{s1} \tag{S47}$$

Total soil wetting (W):

The soil wetting in $A_{\alpha_0\sim\alpha_1}$:

$$W_1 = \int_0^{D_{0\alpha_1}}(\alpha_1 - G(D))dD = \alpha_1 D_{0\alpha_1} - \int_0^{D_{0\alpha_1}} G(D)dD \tag{S48}$$

i) Soil wetting in $A_{\alpha_1\sim\alpha_2}$:

$$W_2 = P(\alpha_2 - \alpha_1) - R_{i2} \tag{S49}$$

ii) Soil wetting in $A_{\alpha_2 \sim 1}$:

$$W_3 = P(1 - \alpha_2) \quad \text{(S50)}$$

$$W = W_1 + W_2 + W_3 \quad \text{(S51)}$$

Evaporation (m):

$$E = \frac{W+S_0}{S_b}\frac{E_p+S_b-\sqrt{(E_p+S_b)^2-2aE_pS_b}}{a} \quad \text{(S52)}$$

Direct runoff (m):

$$R_d = R_i + \gamma R_s \quad \text{(S53)}$$

Groundwater recharge (m):

$$R_g = (1-\gamma)R_s \quad \text{(S54)}$$

Direct streamflow (m/s):

$$Q_d = k_d(R_d + S_{d0}) \quad \text{(S55)}$$

Baseflow (m/s):

$$Q_b = k_b(R_g + S_{g0}) \quad \text{(S56)}$$

Total streamflow (m/s):

$$Q = Q_d + Q_b \quad \text{(S57)}$$

Quick storage tank (m):

$$S_{d1} = (1-k_d)(R_d + S_{d0}) \quad \text{(S58)}$$

Slow storage tank (m):

$$S_{g1} = (1-k_b)\big(R_g + S_{g0}\big) \quad \text{(S59)}$$

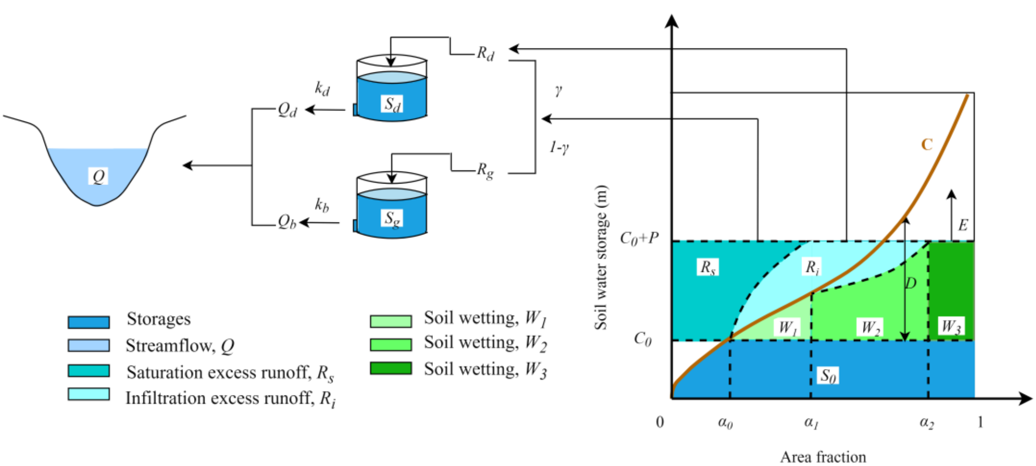


Figure S1. Conceptual diagram of the simple hydrologic model based on URSSIE. Q is total streamflow at the catchment outlet; Q_d is direct streamflow at the catchment outlet; Q_b is baseflow at the catchment outlet; k_d is the average characteristic time of the quick storage tank (s^{-1}); k_b is the average characteristic time of the slow storage tank (s^{-1}); S_d is the quick storage tank; S_g is the slow storage tank; R_d is direct runoff; R_g is groundwater recharge; γ is the partitioning parameter (-); C is soil water storage capacity at a point; Y(C) is the fraction of the catchment area for which the storage capacity is less than or equal to C; S_0 is initial soil water storage; P is precipitation depth; R_s is saturation excess runoff; R_i is infiltration excess runoff; E is actual evaporation; D is distance from water table to land surface(C-C_0); W_1 is the soil wetting in the area that will be saturated in the current timestep; W_2 is the soil wetting in the area that only infiltration excess will occur; W_3 is the soil wetting in the area where the infiltration capacity is greater than precipitation intensity; $A_{0\sim\alpha_0}$ is the subzone already saturated at the beginning of the current timestep; $A_{\alpha_0\sim\alpha_1}$ is the subzone that will be saturated in the current timestep; $A_{\alpha_1\sim\alpha_2}$ is the subzone where only infiltration excess occurs; $A_{\alpha_2\sim 1}$ is the subzone where no runoff is generated.

Eight parameters are needed to numerically implement the above equations in the hydrologic model, as summarized in Table S1.

Table S1. List of parameters in the simple hydrologic model

Parameter	Description	Range	Unit	References
S_b	Average soil water storage capacity over the catchment	0.05-1.5	m	
a	Shape parameter	0-2	–	
γ	Coefficient for partitioning runoff between direct runoff and groundwater recharge	0-1	–	Yao et al. (2020), Wang (2018)
k_b	Average characteristic times of the slow storage tank	$0\text{-}1.653 \times 10^{-6}$	s^{-1}	
k_d	Average characteristic times of the quick storage tank	$1.653 \times 10^{-6}\text{-}1.157 \times 10^{-5}$	s^{-1}	
m_s	Snowmelt coefficient	0-1	–	Wen et al. (2013), Martinez and Gupta (2010)
m_k	Potential maximum rainfall intensity	$0\text{-}2.315 \times 10^{-5}$	m/s	
n	Scaling coefficient	0.4-1	–	

2.2 Application of the hydrologic model

2.2.1 Input data preparation and catchment selection

The CAMELS dataset (Catchment Attributes and Meteorology for Large-sample Studies) is a comprehensive hydrometeorological dataset that includes daily streamflow data from 1985 to 2014 for 671 small, natural catchments across the United States (Newman et al., 2015). In addition to the time series data, the dataset also provides a wide range of catchment properties, such as topography, land cover, soil, and geology. This rich dataset has been extensively used in various hydrological studies, including rainfall-runoff modeling (Kratzert et al., 2018) and the analysis of conceptual model structure uncertainty (Knoben et al., 2020). For the purpose of this study, a subset of 181 catchments is selected from the CAMELS dataset based on two criteria: 1) the availability of complete daily hydroclimatic driving force data and 2) a snow fraction of less than 5% to minimize the impact of snow on runoff generation. The selected catchments exhibit a wide range of physical characteristics and climatic conditions. The drainage area of these catchments varies from 9 to 6,136 km^2, showcasing a diverse range of catchment sizes. Mean annual precipitation across the study catchments ranges from 240 mm/yr to 2,894 mm/yr, while mean annual potential evapotranspiration varies from 718 mm/yr to 1,732 mm/yr. To further characterize the climatic conditions of the study catchments, the climate aridity index (AI) is calculated (Fig. S2). The climate aridity index, defined as the ratio of mean annual potential evapotranspiration to mean annual precipitation, ranges from 0.248 to 4.763 across the 181 catchments. Among these, 14 catchments have an aridity index below 0.5, indicating relatively humid conditions, while 101 catchments have an aridity index between 0.5 and 1.0, representing moderately humid to semi-arid conditions. Furthermore, 49 catchments have an aridity index between 1.0 and 2.0, corresponding to semi-arid to arid conditions, and 17 catchments have an aridity index greater than 2.0, indicating highly arid conditions. The diverse range of climate zones represented by the selected catchments suggests that the contribution of saturation excess runoff and infiltration excess runoff to mean annual runoff is likely to exhibit significant spatial variability. This wide spectrum of climatic conditions allows for a comprehensive analysis of the dominant runoff generation mechanisms across different hydrologic regimes.

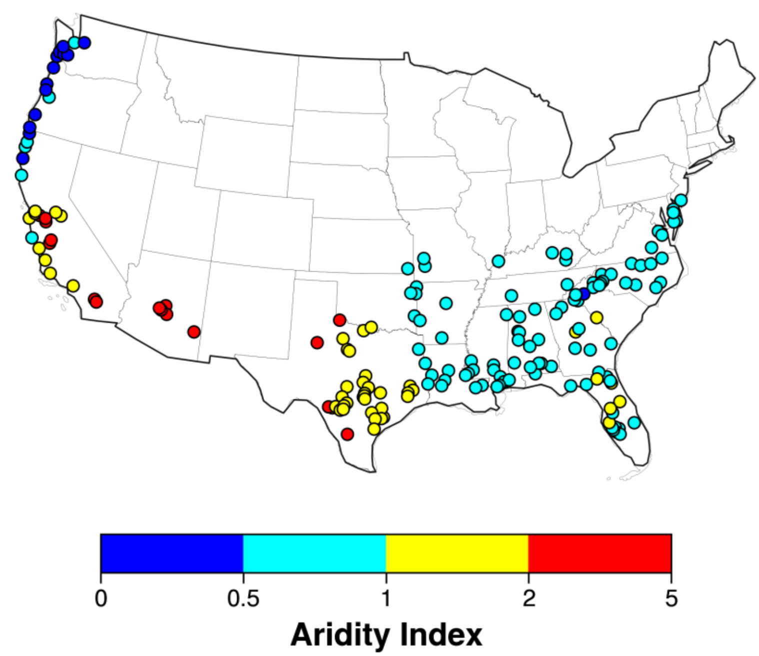


Figure S2. Spatial map of the 181 catchments and their Aridity index values.

The NLDAS-2 forcing dataset (Phase 2 of the North American Land Data Assimilation System) provides hourly precipitation, potential evapotranspiration, and air temperature in 1/8th-degree grid spacing over North America from 1979 to the present (Xia et al., 2012). For this study, the hourly forcing data for the 181 selected catchments from 1985 to 2014 are extracted from the NLDAS-2 dataset. The 1985-2014 period is further divided into three sub-periods: 1985–1986 for model spin-up, 1987–2005 for parameter determination, and 2006–2014 for model evaluation.

2.2.2 Input data preparation and catchment selection

The simple hydrologic model has eight parameters, including S_b, a, γ, k_b, k_d, m_s, m_k, and n. The ranges of these parameters are listed in Table S1. The optimal parameters are not calibrated, but "determined" through three major steps, following a strategy proposed by Abeshu et al. (2023), as shown in Fig. S3. The first step is parameter sampling. For each catchment, 100,000 parameter sets are sampled over the whole parameter space using a Latin hypercube sampling method (McKay et al., 1979) that samples the parameter values over a multidimensional space whilst preserving the statistical distribution of each parameter. Here, the number 100,000 is chosen by balancing the computational costs and representativeness of the parameter space. The second step is running URSSIE for each of 100,000 parameter sets at each catchment and archiving the simulated daily streamflow data. The third step is determining the optimal parameter set for each selected catchment in the following sub-steps:

(1) The initial 100,000 parameter sets are filtered to retain the top 10,000 that best preserve the annual water balance, i.e., based on the normalized root mean square error (NRMSE) between simulated and observed annual streamflow.

(2) These 10,000 sets are further filtered to the top 1,000 that best reproduce the seasonal water balance, i.e., based on the NRMSE between the simulated and observed regime curves.

(3) The 1,000 sets are then narrowed to the top 100 that best capture flood peaks, i.e., based on the NRMSE between the simulated and observed annual flood peaks.

(4) For the remaining 100 parameter sets, the Modified Kling-Gupta Efficiency (KGE) (Kling et al., 2012) values are computed based on the observed and corresponding simulated daily streamflow. The parameter set with the highest KGE value was selected as the optimal set.

After the above step, an optimal set of parameters for each catchment has been determined to facilitate further model-based analyses.

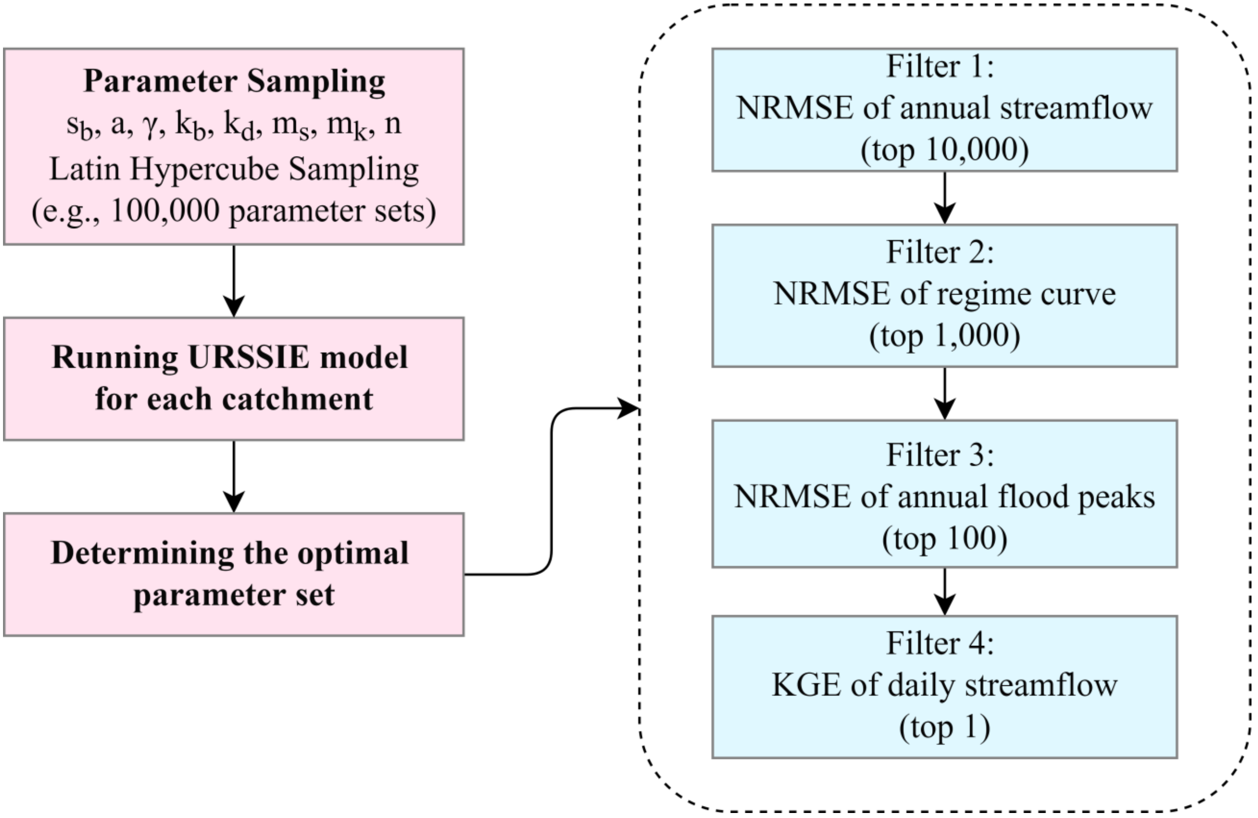


Figure S3. Flowchart for parameter determination

2.2.3 Additional modeling results

Figure S4 compares the KGE values during the parameter determination (1987-2005) and validation (2005-2014) stages. In the parameter determination stage (Fig. S4a), the KGE values are greater than 0.5 in 164 catchments, between 0 and 0.5 in 17 catchments. In the validation stage (Fig. S4b), the KGE values are greater than 0.5 in 126 catchments, between 0 and 0.5 in 48 catchments, and between -1 and 0 in 7 catchments. The worsened results in the validation period may be related to the exceptionally extreme rainfall events that did not occur in the determination period (Pendergrass et al., 2017), hence not captured by our parameter determination procedure.

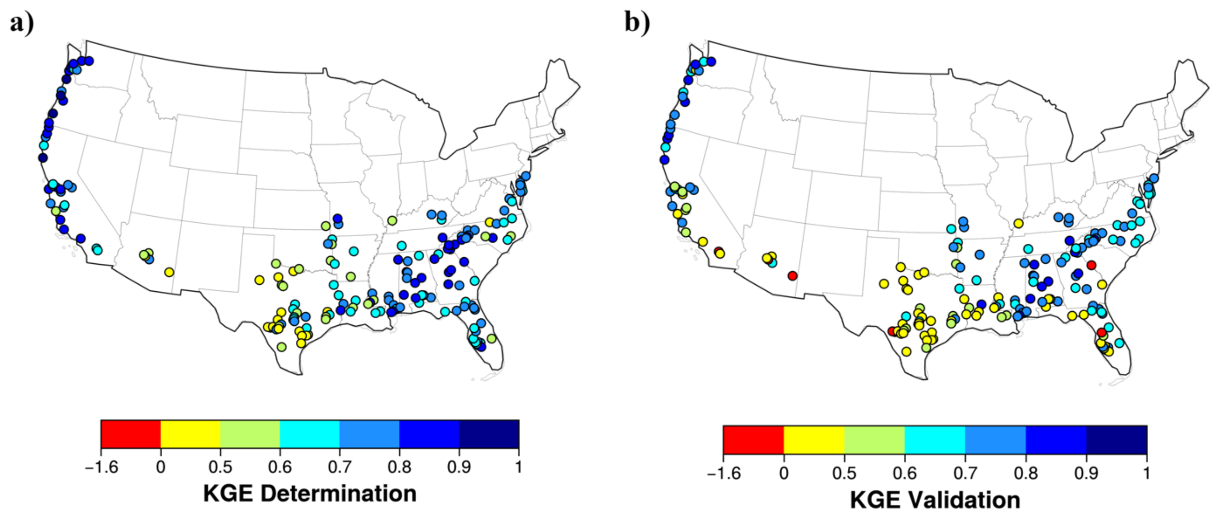
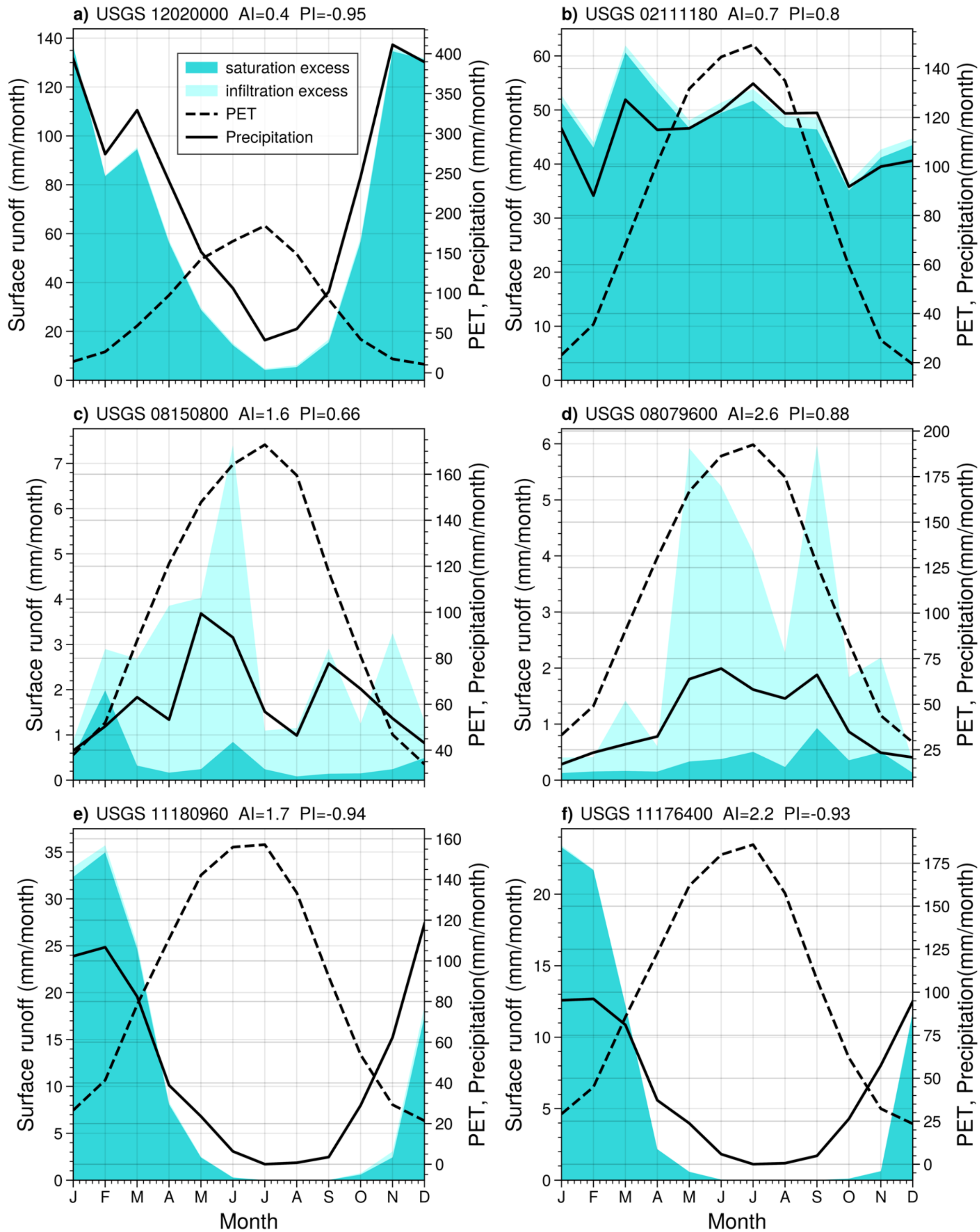


Figure S4. Spatial map of the KGE values at the 181 catchments.

Figure S5 shows the seasonality of infiltration excess and saturation excess in 6 representative catchments. In humid catchments (Figs. S5a, S5b), saturation excess dominates runoff generation regardless of whether seasonal evaporative energy and precipitation are in phase or not. In arid catchments with seasonal evaporative energy and precipitation in phase (Figs. S5c, S5d), infiltration excess dominates runoff generation. In arid catchments with seasonal evaporative energy and precipitation out of phase (Figs. S5e, S5f), saturation excess dominates runoff generation.

260



a) USGS 12020000 AI=0.4 PI=-0.95
b) USGS 02111180 AI=0.7 PI=0.8
c) USGS 08150800 AI=1.6 PI=0.66
d) USGS 08079600 AI=2.6 PI=0.88
e) USGS 11180960 AI=1.7 PI=-0.94
f) USGS 11176400 AI=2.2 PI=-0.93
saturation excess
infiltration excess
PET
Precipitation
Surface runoff (mm/month)
PET, Precipitation (mm/month)
J F M A M J J A S O N D
Month

Figure S5. Relative dominance of infiltration and saturation excess runoff at the seasonal scale at the representative catchments.

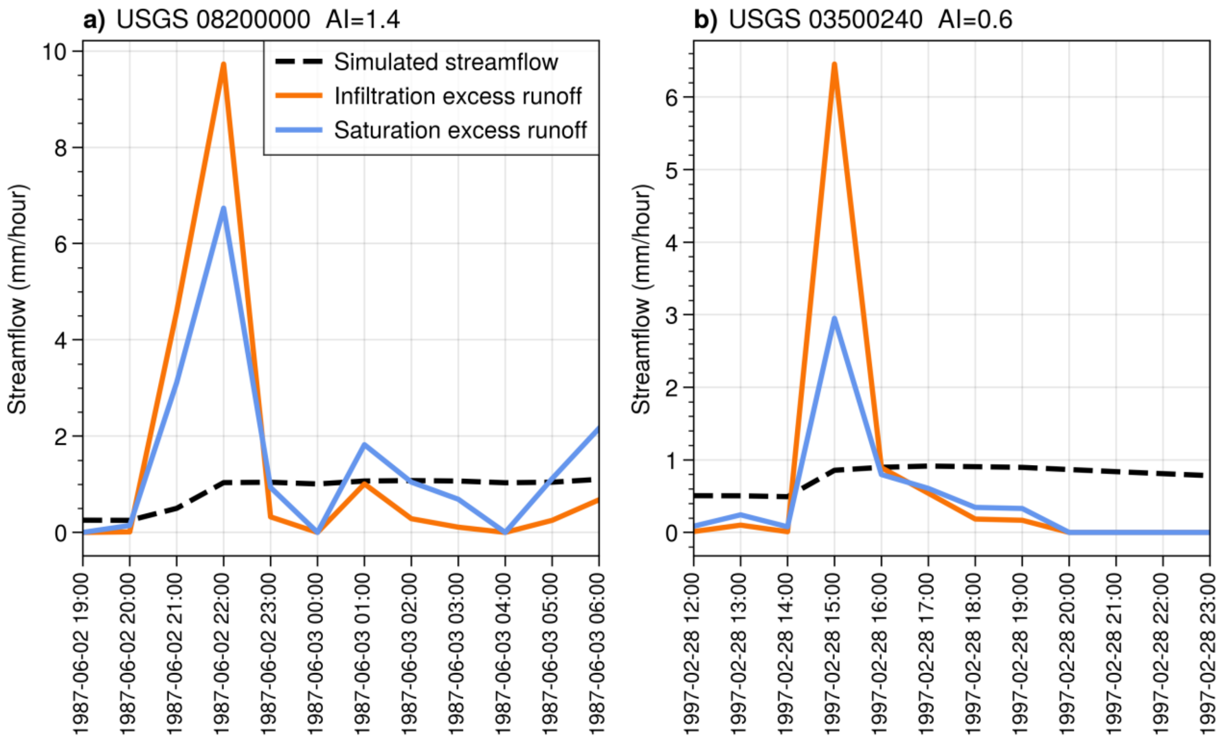


Figure S6. Relative dominance of infiltration and saturation excess runoff at the event scale at the representative catchments.

Figure S6 shows two example events where there are switches of relative dominance of infiltration and saturation excess runoff. Infiltration excess dominates at the beginning of the event. As more and more areas become saturated, saturation excess dominates runoff generation.

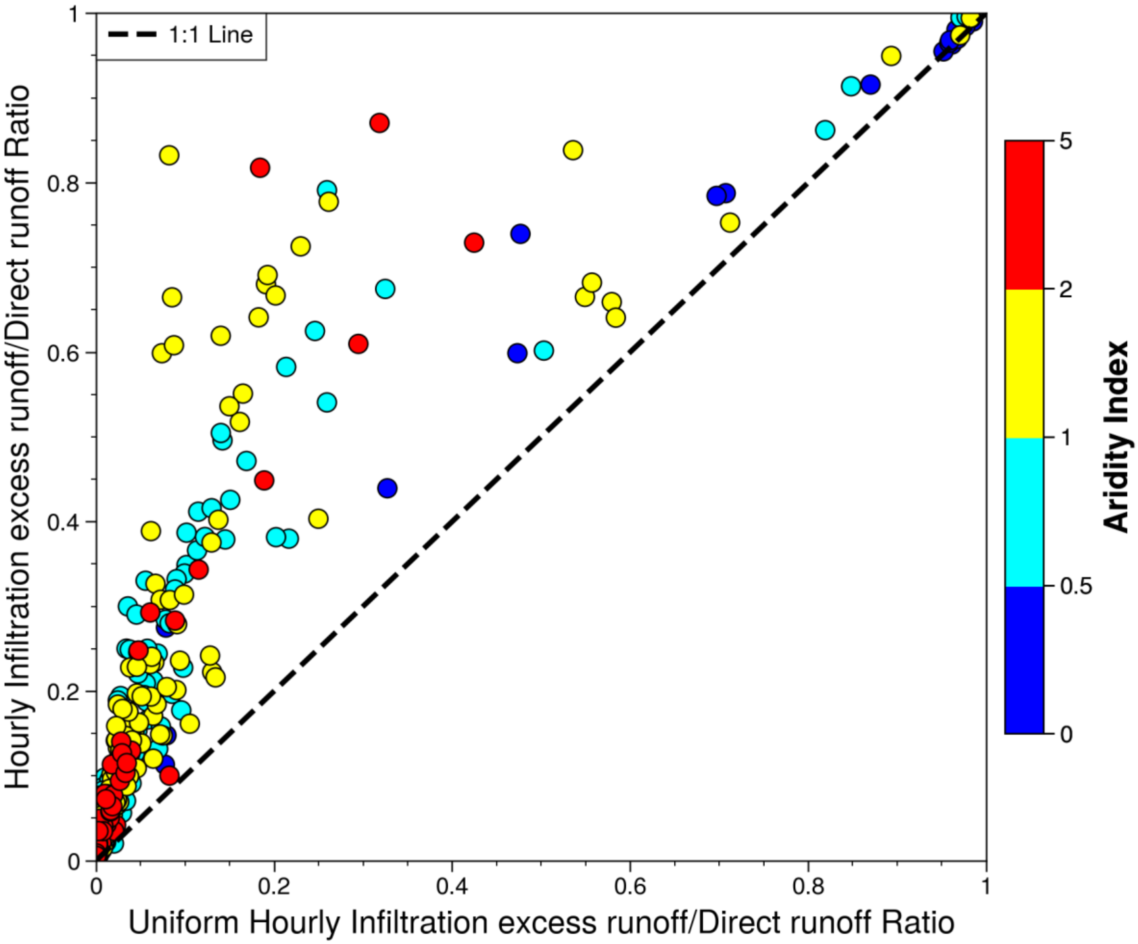


Figure S7. Comparison of long-term (1985-2014) infiltration excess runoff over total direct runoff ratio driven by hourly and 24-hour uniform precipitation data in 181 catchments.

270 To explore the influence of input precipitation data variability in each day to infiltration excess runoff ratio, we compared the long-term infiltration excess runoff ratio calculated using same calibrated best parameter sets but driven by hourly and 24-hour uniformed hourly precipitation data. Figure S7 depicts that in all the 181 catchments, the variability of precipitation within each day produces more infiltration excess runoff.

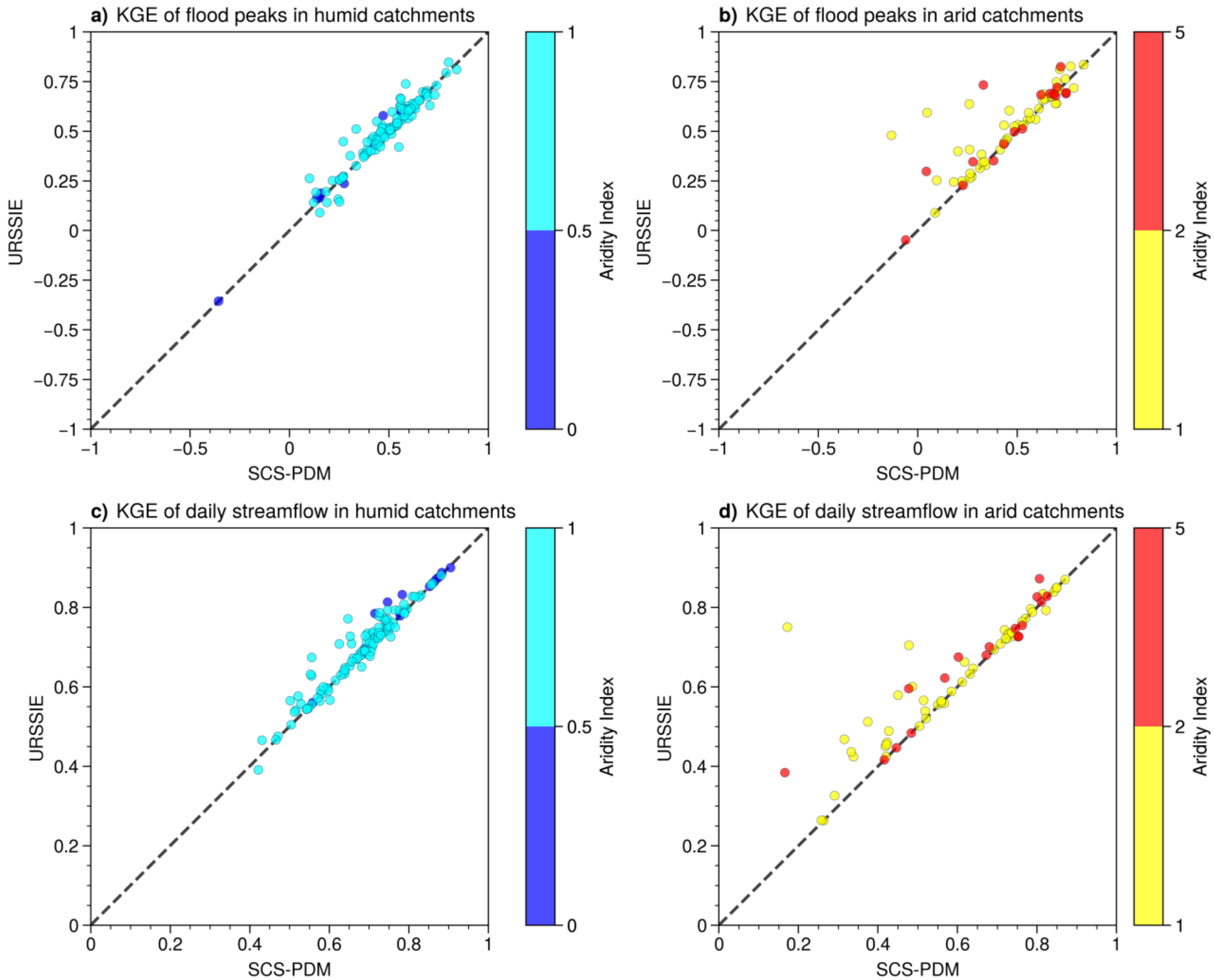


Figure S8. KGE of annual flood peaks (a, b) and daily streamflow (c, d) in humid (a, c) and arid (b, d) catchments, respectively, by URSSIE and SCS-PDM in 1987-2014.

To demonstrate the improvements of the URSSIE model, we compared URSSIE with the SCS-PDM model developed by Yao et al. (2020). Given that the SCS-PDM model operates on a daily timescale, we have likewise implemented the URSSIE model at a daily timescale for comparison. We apply the same climate forcing data, parameter optimization strategy, and study period as utilized in the URSSIE model, ensuring consistency across both models for a fair comparison. In humid catchments, the URSSIE model and SCS-PDM model has similar performance on capturing streamflow (Figs. S8a, S8c). In arid catchments, the KGE of annual flood peaks has increased for more than 0.05 in 20 of the 66 catchments (Fig. S8b) and the KGE of daily streamflow increased for more than 0.05 in 15 of the 66 catchments (Fig. S8d).

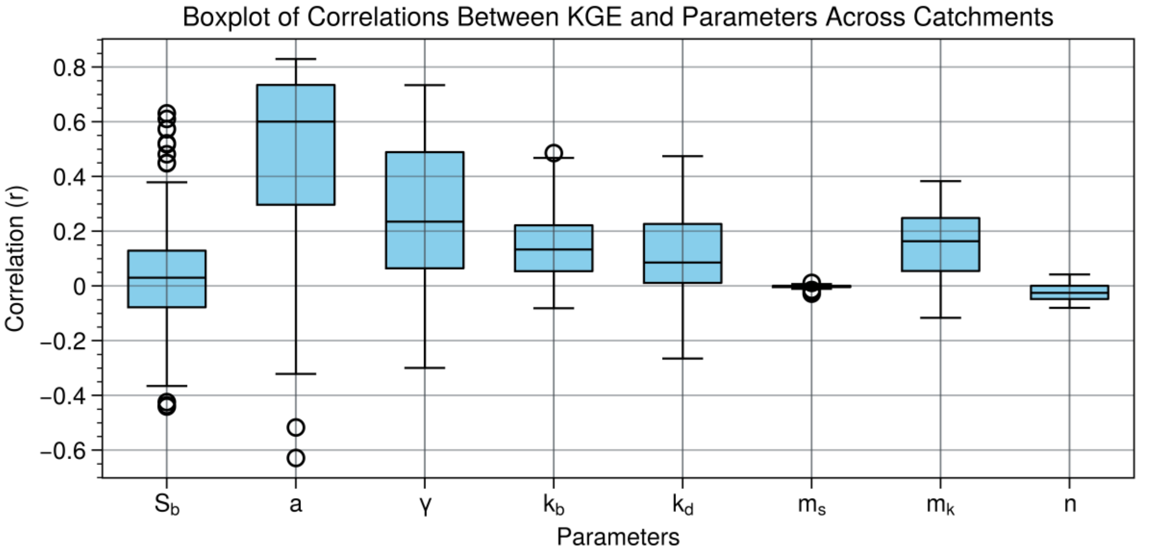


Figure S9. Boxplot of correlations between KGE and parameters across 181 catchments in 1987-2014.

To evaluate the sensitivity of the model performance to the eight parameters, we calculated the correlation coefficient (r) between the KGE and each parameter across multiple catchments. The parameter sets come from the 100,000 parameter sets generated by LHS and used for parameter determination. The boxplot presented in Fig. S9 shows that the model performance is most sensitive to shape parameter a, indicating that adjustments to this parameter have a stronger and more variable influence on streamflow simulation accuracy. The moderate correlations of γ, k_b, k_d, and m_k indicate they have a consistent but less pronounced effect. In contrast, S_b, m_s and n display weaker correlations, suggesting they have a minimal impact on model performance.

References

Abeshu, G. W., Tian, F., Wild, T., Zhao, M., Turner, S., Chowdhury, A. F. M., and Li, H.-Y.: Enhancing the representation of water management in global hydrological models, *Geosci. Model Dev.*, 16, 5449–5472, doi:10.5194/gmd-16-5449-2023, 2023.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, *J. Hydrol.*, 424, 264–277, doi:10.1016/j.jhydrol.2012.01.011, 2012.

Knoben, W. J., Freer, J. E., Peel, M. C., Fowler, K. J. A., and Woods, R. A.: A brief analysis of conceptual model structure uncertainty using 36 models and 559 catchments, *Water Resour. Res.*, 56, e2019WR025975, doi:10.1029/2019WR025975, 2020.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, *Hydrol. Earth Syst. Sci.*, 22, 6005–6022, doi:10.5194/hess-22-6005-2018, 2018.

Martinez, G. F. and Gupta, H. V.: Toward improved identification of hydrological models: A diagnostic evaluation of the "abcd" monthly water balance model for the conterminous United States, *Water Resour. Res.*, 46, W08507, doi:10.1029/2009WR008294, 2010.

McKay, M. D., Conover, W. J., and Beckman, R. J.: A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, 21, 239–245, 1979.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological dataset for the contiguous USA: Dataset characteristics and assessment of regional variability in hydrologic model performance, *Hydrol. Earth Syst. Sci.*, 19, 209–223, doi:10.5194/hess-19-209-2015, 2015.

Pendergrass, A. G., Knutti, R., Lehner, F., Deser, C., and Sanderson, B. M.: Precipitation variability in a warmer climate, *Sci. Rep.*, 7, 16966, doi:10.1038/s41598-017-17966-y, 2017.

Wang, D.: A new probability density function for spatial distribution of soil water storage capacity leads to the SCS curve number method, *Hydrol. Earth Syst. Sci.*, 22, 6567–6578, doi:10.5194/hess-22-6567-2018, 2018.

Wen, L., Nagabhatla, N., Lü, S., and Wang, S.-Y.: Impact of rain–snow threshold temperature on snow depth simulation in land surface and regional atmospheric models, *Adv. Atmos. Sci.*, 30, 1449–1460, doi:10.1007/s00376-012-2192-7, 2013.

Xia, Y., et al.: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *J. Geophys. Res.*, 117, D03109, doi:10.1029/2011JD016048, 2012.

Yao, L., Libera, D. A., Kheimi, M., Sankarasubramanian, A., and Wang, D.: The roles of climate forcing and its variability on streamflow at daily, monthly, annual, and long-term scales, *Water Resour. Res.*, 56, e2020WR027111, doi:10.1029/2020WR027111, 2020.

Yao, L. and Wang, D.: Climatic control on spatial distribution of water storage at the catchment scale: A framework for unifying saturation excess runoff models, *J. Geophys. Res. Atmos.*, 127, e2021JD036334, doi:10.1029/2021JD036334, 2022.