Case Study

Balancing Complexity, Parsimony, and Applicability in Hydrologic Modeling: A Comparative Evaluation of Four Infiltration Models across Parameterization Scenarios

Sara Rassa[1]; Gerhard Schoener, Ph.D., A.M.ASCE[2]; and Matthew Fleming, P.E.[3]

Abstract: Infiltration models are vital for predicting the onset and magnitude of flash floods. A multitude of infiltration models have been developed over the past century. In practice, model selection for a given study constitutes a compromise between complexity, parsimony, and feasibility. This study presents a comparative evaluation of four infiltration models of varying complexity available in the widely used HEC-HMS software: the curve number (CN), initial-constant (IC), Green-Ampt (GA), and the recently integrated linear-constant (LC) model. Using precipitation and runoff data from the Walnut Gulch Experimental watershed in Arizona, each model was analyzed across three different scenarios including, using published guidance, constrained calibration, and unconstrained calibration. Findings reveal that GA, the most complex model, excelled over the rest under published guidance and constrained calibration scenarios but was susceptible to equifinality issues under unconstrained calibration. Conversely, the simpler IC model performed well under unconstrained calibration. The results indicate the simple, two-parameter LC model offers a balance between complexity, parsimony, and feasibility in hydrologic modeling, particularly for arid and semiarid regions. This study provides valuable insight to hydrologists and researchers in selecting an appropriate infiltration model based on data availability and regional hydrologic conditions. **DOI: 10.1061/JHYEFF.HEENG-6258.**

Author keywords: Infiltration model optimization; Initial-constant (IC); Linear-constant (LC); Green-Ampt (GA); Curve number (CN).

Introduction

Flash flooding is among the most hazardous natural disasters globally and in the United States (Hapuarachchi et al. 2011). Flood prediction in arid and semiarid regions is challenging due to the spatial and temporal heterogeneities of hydrological characteristics within a catchment, including variations in precipitation patterns, runoff generation, and soil moisture conditions (Huang et al. 2016). Hydrologic models are essential in both managing and predicting flash floods. Event-based models are commonly used in arid and semiarid regions where flows are episodic and transient (Huang et al. 2016). Infiltration excess overland flow is the key hydrological process in many arid and semiarid watersheds (Huo et al. 2020; Kampf et al. 2018; Huang et al. 2016). Infiltration models are therefore vital for predicting the onset and magnitude of flash floods and play a key role in flood mitigation and emergency response planning.

A multitude of infiltration models of varying complexity have been developed over the past century (Mishra et al. 2003; Morbidelli et al. 2018; Chahinian et al. 2005). At one end of the spectrum, simple models have the advantage of requiring fewer parameters and are thus easier to implement. Ease of implementation is an important consideration, particularly in data-scarce areas (Pande et al. 2012). For this reason, the simple (one-parameter) NRCS curve number method (USDA 2004b) is still popular among practicing hydrologists (Moglen et al. 2022) despite its well-documented limitations (e.g., Ogden et al. 2017). Overly simplistic models, however, may lack the required complexity to adequately simulate physical processes; this can limit their predictive power (Hrachowitz et al. 2014).

Complex models, whereas perhaps more adept at simulating physical processes, are associated with their own set of challenges. In data-sparse regions, models with multiple parameters are more difficult to apply (Pande et al. 2012). Model calibration and validation are often impossible due to a lack of observations, especially in arid and semiarid regions where flows are episodic and gauge data are rare (Mengistu et al. 2019; Schoener 2018). Calibration and validation are important steps in hydrologic modeling (Mai 2023); during model calibration, parameters are adjusted, and the model is optimized by minimizing the difference between model predictions and observations. The optimized model is then validated using an independent data set. Even if rainfall–runoff data are available, overparameterization can lead to equifinality (Beven 2006), where different parameter combinations result in nearly identical model fits during calibration (Kirchner 2006). Equifinality (or nonuniqueness) can lead to poor model predictions despite satisfactory calibration performance (Hrachowitz et al. 2014).

Strategies for reducing the negative impact of equifinality include the implementation of a parsimonious model structure and

[1]Graduate Student, Dept. of Civil, Construction, and Environmental Engineering, Univ. of New Mexico, MSC01 1070, Albuquerque, NM 87131 (corresponding author). ORCID: https://orcid.org/0000-0002-0658-994X. Email: sarars892@gmail.com; srassa@unm.edu

[2]Senior Hydrologist, Southern Sandoval County Arroyo Flood Control Authority, 1041 Commercial Dr. SE, Rio Rancho, NM 87124. ORCID: https://orcid.org/0000-0002-1183-0419

[3]Chief, Water Management Systems Division, Hydrologic Engineering Center, 609 Second St., Davis, CA 95616.

Note. This manuscript was submitted on January 29, 2024; approved on June 25, 2024; published online on September 13, 2024. Discussion period open until February 13, 2025; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699.

constraining the range of model parameters during calibration (Hrachowitz et al. 2014; Guse et al. 2020). Pande et al. (2012) promoted a hydrologic modeling strategy with parsimonious functional relationships for regions with limited data. Tegegne et al. (2017) and Quichimbo et al. (2021) similarly argued for a parsimonious model structure in data-scarce dryland areas.

Multiple authors have argued that constraining the parameter space of hydrologic models can improve model performance and reduce equifinality problems (e.g., Guse et al. 2020; Wambura et al. 2018; Hrachowitz et al. 2014). Parameter constraints have been implemented using remotely sensed data (Nijzink et al. 2018), long-term average conditions (Wambura et al. 2018), or even anecdotal evidence and expert knowledge (Hrachowitz et al. 2014).

Selection of an appropriate infiltration model for a given study constitutes a compromise between three factors: (1) sufficient model complexity to adequately simulate underlying processes (Hrachowitz et al. 2014); (2) parsimony—to the greatest extent possible—i.e., representation of dominant processes in the simplest possible way (Pande et al. 2012); and (3) feasibility, given real-world limitations faced by the modeler (e.g., data availability, time, funding).

In this study, we investigate the performance of four infiltration models available within the widely used Hydrologic Modeling System (HEC-HMS) software. Developed and maintained by the Hydrologic Engineering Center (HEC) of the US Army Corps of Engineers, HEC-HMS is a free software package used for hydrologic modeling worldwide. HEC usage statistics show that version 4.10 of the software was downloaded more than 65,000 times in the 18-month period between its release in July 2022 and December 2023. Version 4.11, released in July 2023, had been downloaded more than 34,000 times by the end of 2023. Fig. 1 compares page visits associated with the HEC-HMS model by continent for the year 2023. The comparison illustrates that the model is used around the world; researchers and practitioners from more than 190 countries visited model-related pages in 2023.

The four infiltration models compared in this study are the curve number (CN) model (USDA 2004b), initial-constant (IC) model (USACE 2000), linear-constant (LC) model (Schoener et al. 2021), and the Green-Ampt (GA) model (Green and Ampt 1911). Models range in complexity from one parameter (CN) to four parameters (GA). Both the CN and GA models have been available for decades and are perhaps the most widely applied methods for runoff calculations in hydrologic practices (Grimaldi et al. 2013; Baiamonte 2019). With two parameters, the complexity of the IC model falls between CN and GA. However, Schoener et al. (2021) pointed out that the IC model can result in errors when precipitation intensity early during a storm event is high, a common occurrence during convective storms characteristic of many dryland areas. To overcome this problem, they proposed the linear-constant (LC) model, a modification of the IC method. The LC algorithm was subsequently added to the HEC-HMS software and became available with the release of version 4.12 Beta 1 in October 2023. Apart from initial development (Schoener et al. 2021) and benchmark testing by HEC during implementation, the LC model has not been applied widely in project applications.

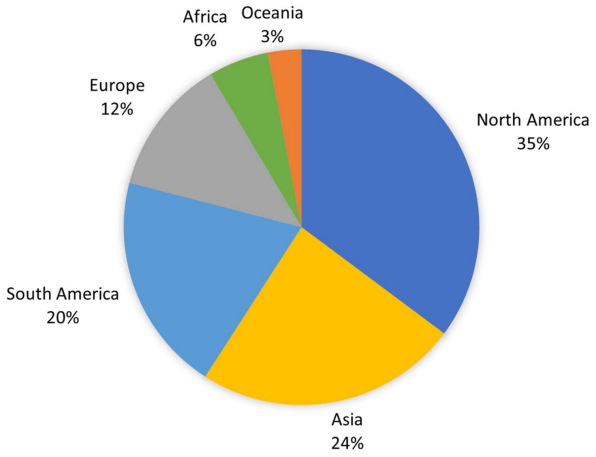

Fig. 1. Visits of webpages associated with the HEC-HMS model (download, user manual, technical reference manual, tutorials, training) by continent in 2023.

The objectives of this study were three-fold: (1) test the performance of the linear-constant infiltration model newly added to HEC-HMS in 2023; (2) compare performance with three other infiltration models (curve number, initial-constant, Green-Ampt) of varying complexity; and (3) assess the impact of different model parameterization strategies.

Testing was carried out using long-term precipitation and runoff data from the Walnut Gulch Experimental Watershed in Arizona, US. Models were parameterized based on published guidance (Schoener et al. 2023), and using constrained and unconstrained calibration approaches. Findings may have broad applicability to arid and semiarid areas by informing modeling strategies that aim to strike a balance between complexity, parsimony, and feasibility.

Methods

Study Area

In this study, seven basins were chosen within the Walnut Gulch Experimental Watershed (WGEW), managed by the US Department of Agriculture's Agricultural Research Service. The hydrology of this area is primarily influenced by the North American monsoon, which provides about two-thirds of the annual precipitation during the summer months (Goodrich et al. 2008). WGEW has extensive and well-monitored data collections including hydrological, meteorological, and soil data, enhancing the understanding of semiarid hydrology in this region and adjacent areas (Renard et al. 2008). Basins 4, 11, 102, 103, 104, 105, and 106 were selected within the WGEW because they are unaffected by urbanization and are representative of a range of spatial scales from 0.224 to 773 ha. Soils in the study area are comprised of gravelly sandy loams (Breckenfeld et al. 2008). Vegetation in this semiarid watershed consists largely of desert shrubs, interspersed with some grasses (Nearing et al. 2007).

Precipitation, Runoff, and Soil Moisture Data

Precipitation, runoff, and soil moisture data were obtained from the Southwest Watershed Research Center Data Access Project (Southwest Watershed Research Center 2024). Soil moisture records for the study area were only available for the period 2002–2019; because soil moisture is required for one of the infiltration models (GA); we therefore selected this time frame for our study.

The main focus of this study was flash flooding; we therefore decided to focus on the largest recorded events for each study basin. Runoff events for the 18-year period were ranked by peak discharge, and the 30 largest events (20 for basins 4 and 11) were selected for analysis. Runoff and precipitation data in breakpoint format were converted to uniform 1-min time steps. Locations of rain gauges, soil moisture probes, and flumes for discharge measurement are

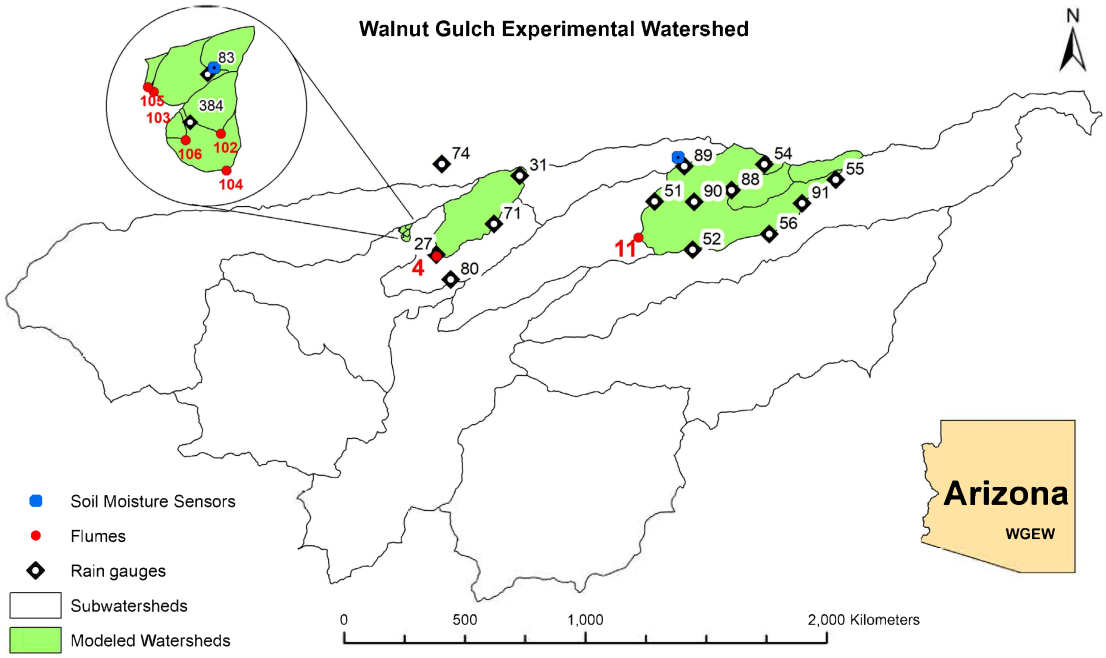

Fig. 2. Overview map of Walnut Gulch Experimental Watershed in southern Arizona, US.

indicated in Fig. 2. Detailed information about storm events including dates, peak rainfall and runoff, rainfall and runoff depths, and antecedent soil moisture conditions are included in Tables S1–S3. Despite the focus on large storms, selected events cover a range of magnitudes as indicated by event statistics.

Event data for each basin were ranked based on peak discharge and alternately split into a calibration and validation set. This was done to ensure storms of different magnitudes would be represented in each set. Average antecedent conditions across all basins and events were 0.10 m^3 m^{-3} (range 0.01–0.25 m^3 m^{-3}). The distribution of antecedent conditions for validation and calibration sets were similar in each test basin (see Tables S1–S3).

HEC-HMS Model

The US Army Corps of Engineers HEC-HMS software version 4.11 and 4.12 Beta 2 were used to build the rainfall-runoff models for individual basins.

Watershed and Subbasin Delineation

Basins were delineated utilizing a digital elevation model (DEM) with a 1 m vertical resolution obtained from the USGS (2024). Basins were first delineated using the geographical information systems tools available within the HEC-HMS model. All basin boundaries and flow paths were then verified manually and adjusted where necessary in ArcGIS (ESRI 2023) based on elevation contours generated from the DEM.

In a study based on Walnut Gulch data, Goodrich et al. (1997) found that basin runoff response became more nonlinear with increasing catchment size. They identified a critical threshold area of 40–60 ha and attributed nonlinear behavior for watersheds exceeding that threshold to spatial variation in rainfall and transmission losses in ephemeral channels. Based on these findings, in this study, small catchments (102–106) were modeled as a single basin. Larger watersheds were divided into subbasins not exceeding the 60-ha threshold (Goodrich et al. 1997). This allowed us to model transmission losses in larger watersheds (see section "Transform, Routing, and Transmission Losses" below). Following the temporal resolution of input data, models were run with a 1-min time step.

Precipitation

Rainfall was modeled in HEC-HMS using the interpolated precipitation method. The model develops a precipitation grid by interpolating between point precipitation gauges using the inverse distance method. The number of point gauges used in the interpolation depends on the size of each study basin: two rainfall gauges were used for basins 102–106 (see Fig. 2, gauges 83 and 384); five precipitation gauges were used for basin 4 (gauges 27, 31, 71, 74, and 80); and nine rainfall gauges were used to generate interpolated precipitation grids for basin 11 (gauges 51, 52, 54–56, and 88–91).

Infiltration Models

Four infiltration models were tested and compared in this study: curve number (CN), Green-Amp (GA), initial-constant (IC), and linear-constant (LC). Infiltration models are briefly described below with corresponding equations and references.

The curve number (CN) model is derived from the Natural Resources Conservation Service (NRCS) curve number equation (USDA 2004b). The curve number model is an empirical surface runoff model that estimates cumulative runoff based on the total precipitation. Green-Ampt (GA) is a physically based model that simulates ponded infiltration of water into a homogeneous soil with a uniform initial water content (Green and Ampt 1911):

$$f_t = K_{\text{eff}}\left(1 + \frac{\psi \Delta\theta}{F_t}\right) \tag{1}$$

where f_t (mm h^{-1}) = infiltration rate at time t; K_{eff} (mm h^{-1}) = effective hydraulic conductivity; ψ (mm) = wetting front suction; $\Delta\theta$(m^3 m^{-3}) = difference between initial moisture content and saturated content called soil moisture deficit; and F_t (mm) = cumulative infiltration at time t.

The initial-constant (IC) model (USACE 2000) assumes that the potential rate of infiltration remains constant throughout the storm

event after the initial abstraction is fulfilled. Initial abstraction refers to the precipitation lost prior to the onset of runoff due to interception, depression storage, and high initial infiltration rates while the soil has not become saturated

$$f_t = \left\{ \begin{array}{ll} p_t & \text{if } P_t \le I_{a(\text{IC})} \\ K_{\text{eff}} & \text{if } P_t > I_{a(\text{IC})} \end{array} \right\} \tag{2}$$

where p_t (mm h^{-1}) represents incremental precipitation rate; P_t (mm) = cumulative precipitation at time t; and $I_{a(\text{IC})}$ (mm) = initial abstraction. Because both the CN and IC model have an initial abstraction term, they are distinguished in this paper using notation $I_{a(\text{CN})}$ and $I_{a(\text{IC})}$, respectively. The IC model can result in modeling inaccuracies when precipitation intensity early during the storm event is high. No runoff is generated by the IC model until cumulative precipitation has exceeded the initial abstraction. To overcome this problem, Schoener et al. (2021) proposed the linear-constant (LC) model:

$$f_t = \left\{ \begin{array}{ll} mF_t + f_0 & \text{if } F_t < F_c \\ K_{\text{eff}} & \text{if } F_t \ge F_c \end{array} \right\} \quad (f_0 \ge K_{\text{eff}}) \tag{3}$$

Instead of applying an initial abstraction value, this method assumes that infiltration starts at an initial value of f_0 (mm h^{-1}) and decreases linearly with cumulative infiltration until it reaches to a constant rate of K_{eff} (mm h^{-1}) when cumulative infiltration, F_t (mm), equals or is greater than the initial deficit, F_c (mm). The parameter m (h^{-1}) represents the rate of infiltration potential decay with respect to cumulative infiltration (Schoener et al. 2021). Due to the linear relationship, only three parameters need to be specified: K_{eff}, m, and one other parameter (f_0 or F_c).

Comparability of Infiltration Models and Calibration Strategies

Infiltration models compared in this study have different theoretical backgrounds. The GA model is physically based; the CN model is empirical; and the IC and LC models are conceptual in nature. This raises the questions of comparability and consistency in calibration approaches between models. Three of the models (IC, LC, and GA) have structural similarities: they feature one parameter that describes the value toward which the infiltration curve converges, the effective hydraulic conductivity (K_{eff}). Additionally, all three models feature one or more parameters that describe the infiltration potential decay during the initial phase of the infiltration process. These parameters are the $I_{a(\text{IC})}$ for the IC model; the m and F_c for the LC model; and ψ and $\Delta\theta$ for the GA model.

For consistency, we chose to calibrate K_{eff}, the parameter common to all three models, along with one parameter describing infiltration potential decay. Sensitivity analysis performed by Schoener et al. (2023) showed that the LC model is more sensitive to changes in F_c compared with m, and that using a constant value of $m = -3$ yielded good results for an area similar to that of the present study. We therefore chose to calibrate F_c. Because soil moisture measurements were available for the study period, we chose to calibrate ψ for the GA model and calculated $\Delta\theta$ based on observations.

The empirical CN model differs in structure from the other three. It was not developed as an infiltration model but as a method for estimating total runoff from a storm event as a function of cumulative precipitation. In practice, the CN method has been incorporated into numerous hydrologic modeling programs and is therefore a widely used equivalent to an infiltration equation (Schoener et al. 2021). Applied in this way, the infiltration rate approaches zero for large values of cumulative precipitation. Despite its structural differences to the other three models, the CN method was included in this paper due to its widespread use in practice. Although it has two parameters CN and $I_{a(\text{CN})}$, the latter is typically calculated as a function of the curve number, and this approach was also adopted for the present study.

Parameterization Scenarios for Infiltration Models

Three different parameterization scenarios were tested in this study. The first scenario estimates infiltration model parameters based on published guidance. The other two scenarios optimize parameters based on calibration storm data using the optimization trial module available in HEC-HMS. Optimization trials evaluate simulated and observed information at the outlet of each basin where discharge measurements are available. Minimizing root mean square error (RMSE) was selected as the objective function for this study, in conjunction with the differential evolution algorithm and a population size of 30 parameter sets. For every iteration in the optimization search, the objective function RMSE is computed for all 30 parameter sets, and the variance is compared with a user-specified tolerance (in our case 0.0001). If the variance is larger than the specified tolerance, the parameter sets migrate in a random direction looking for the parameter space that improves the objective function and reduces the variance.

Parameterization Based on Published Guidance

In practice, models are often parameterized based on handbook values or other published guidance because data for model calibration and validation are unavailable or insufficient. This was also the approach applied in the first scenario. Parameter values are summarized in Table 1 (left column). For the curve number model, a CN value of 85 was used for all basins. This corresponds to a total maximum retention (S) of approximately 45 mm [see Table 1 and USDA (2004a) for the equation relating CN and S]. Because initial abstraction $I_{a(\text{CN})}$ was calculated as 20% of S as per standard practice (USDA 2004b), an initial abstraction value of 9 mm was used for this scenario. A CN value of 85 is consistent with plot-scale testing carried out by Schoener et al. (2023), and analysis reported by Stewart et al. (2012) who determined curve numbers for Walnut Gulch basins 102–106 from measured data.

Parameters for the initial-constant, linear-constant, and Green-Ampt models are all based on the study by Schoener et al. (2023), who developed parameter recommendations for all three infiltration models based on rainfall simulation for sandy soils and conditions similar to those in the study area. Parameter values for K_{eff}, $I_{a(\text{IC})}$, F_c, and ψ were selected based on average antecedent soil moisture conditions (0.09 m^3 m^{-3}) observed during calibration and validation events in the study area. Soil moisture deficit ($\Delta\theta$) at the onset of each event for the GA model was calculated as the difference between measured soil moisture and effective porosity. For this study, we assumed an effective porosity of 0.35 for all basins; this value is slightly lower than the recommended value in the HEC-HMS technical reference manual for sandy loam of 0.40 (USACE 2000). The downward adjustment is justified by the fact that sandy loams in the study area are generally stony (Breckenfeld et al. 2008).

Parameterization Based on Constrained Calibration

For the second scenario, infiltration model parameters CN, K_{eff}, $I_{a(IC)}$, F_c, and ψ were calibrated using optimization trials as described in section 2.3.5 above, whereby the search range for each parameter was constrained as shown in Table 1. Search ranges are based on the guidelines by Schoener et al. (2023) who suggested a range of reasonable values for IC, LC, and GA models parameters based on plot-scale testing. Although they did not include values for the CN method, a CN range was easily calculated from their data

Table 1. Infiltration model parameters (and parameter ranges) for the CN, IC, LC, and GA models under two different parameterization scenarios

Parameter	Parameterization based on published guidance	Parameter range for constrained calibration
CN	85	95–59
S (mm)	45	13–177
$I_{a(\mathrm{CN})}$ (mm)	9	3–35
K_{eff} (mm h^{-1})	15	11–21
$I_{a(\mathrm{IC})}$ (mm)	15	0–45
m (h^{-1})	−3	−3
F_c (mm)	26	0–72
ψ (mm)	83	0–183
$\Delta\theta$ (m^3 m^{-3})	0.35 − measured antecedent moisture	0.35 − measured antecedent moisture

Note: $S = 25{,}400/(S + 254)$; and $I_{a(\mathrm{CN})} = 0.2$ S.

using the method for estimating CN values from cumulative precipitation and runoff described by Hawkins et al. (2008). Consistent with the guidance scenario, initial abstraction for the CN method ($I_{a(\mathrm{CN})}$) was calculated as 20% of S, the parameter m was kept constant at −3, and $\Delta\theta$ was estimated as described above.

After completing optimization trials for all calibration storms, mean optimized parameter values were calculated for each basin. Mean values were then applied, and model performance was evaluated for validation events.

Parameterization Based on Unconstrained Calibration

In a third parameterization scenario, we used the same calibration approach as described above except using an unconstrained search range for infiltration model parameters CN, K_{eff}, $I_{a(\mathrm{IC})}$, F_c, and ψ. Because optimization trials in HEC-HMS require the selection of minimum and maximum values, we set the minimum to the smallest allowable value. Maximum values were based on the highest precipitation intensity and total rainfall depth for effective hydraulic conductivity ($K_{\mathrm{eff}} = 1$–200 mm h^{-1}), initial abstraction ($I_{a(\mathrm{IC})} = 0$–100 mm), and initial deficit ($F_c = 0$–100 mm). The range for suction was adopted from Rawls et al. (1983) for sandy loam (= 0–2,000 mm). Published guidance for curve numbers ranges from 30 to 100 (USDA 2004a), so a range of 20–100 was adopted for this study.

Transform, Routing, and Transmission Losses

Excess precipitation was converted into runoff hydrographs using the SCS unit hydrograph method (USDA 2007). Lag time was estimated based on the velocity method (USDA 2010) using flow path measurements performed in ArcGIS (ESRI 2023). Lag times for each basin were then calibrated by minimizing RMSE based on the runoff event with the highest measured peak discharge for each test basin. This resulted in lag values of 2–5 min for basins 102–106, and 4–25 min for individual subbasins in basins 4 and 11.

For routing in HEC-HMS, we used the Muskingum–Cunge method with the eight-point shape, which required selecting a simplified cross section with eight stations versus elevation values. The cross section represented the main channel and left and right overbank areas. Manning's n-values were estimated based on aerial orthoimages of the study area. Ephemeral channels in basins 4 and 11 are similar in geometry, sinuosity, and vegetation patterns. We therefore assigned the same Manning's n-value to all routing reaches. Phillips and Tadayon (2006) provide guidance for n-value selection in ephemeral streams of Arizona. Based on this information and experience of the authors in similar systems (Schoener 2017, 2022), an n-value of 0.025 was used.

Transmission losses were assumed to be negligible for smaller basins (102–106). This assumption is based on observations by Goodrich et al. (1997), who attribute nonlinear runoff response at Walnut Gulch for basins exceeding a threshold area of 40–60 ha to spatial variation in rainfall and transmission losses. For the two larger basins (4 and 11), transmission losses were accounted for in routing reaches following ephemeral channels. Whereas transmission loss is a transient process (Schreiner-McGraw and Vivoni 2018), HEC-HMS at present only allows specification of a constant percolation rate for each routing reach. The model then calculates the wetted perimeter for each time step and estimates reach-average transmission loss by multiplying wetted area and percolation rate. Given these limitations and based on other studies using HEC-HMS to simulate transmission losses in a similar environment (Schoener 2017, 2022), we set the percolation rate to 38 mm h^{-1} for all channels reaches.

Performance Metrics

To test model performance, we compared results from validation trials (i.e., 15 events for basins 102–106 and 10 events for basins 4 and 11) across all study basins and parameterization scenarios. Many metrics for assessing model performance have been developed and used in hydrology (Jackson et al. 2019). This study is focused on flash flood prediction. Two major concerns when modeling flash flooding are peak discharge (e.g., for assessment of infrastructure capacity such as channels or culverts, and inundation and flood plain mapping) and runoff volume (for design and evaluation of stormwater detention facilities). We therefore decided to first compare measured and simulated peak discharge and volume.

Because study basins range in size from 0.224 to 773 ha, peak and volume metrics were normalized by basin area and reported in units of mm h^{-1} and mm, respectively. The absolute error was used in this paper because relative error highlights discrepancies in low flows more than the differences in high flows (Jackson et al. 2019), and our main concern centered on larger events likely to induce the most destructive flash flood response. Whereas peak and volume comparisons are important, they may fail to reveal problems with simulated hydrographs related to timing and overall shape. We therefore computed the Nash-Sutcliffe efficiency (NSE) statistic (Nash and Sutcliffe 1970):

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2} \tag{4}$$

where O_i and S_i = values of observed and simulated discharge (mm h^{-1}) at time i of a particular storm event; and $\overline{O}$ = mean of observed discharge values. The NSE indicates how well observed and simulated hydrographs match, providing the degree of accuracy of the model's performance. NSE ranges from $-\infty$ to 1, with 1 indicating a perfect match between observed and computed

hydrographs. Typically, model simulations with an NSE of 0.5 or higher are considered satisfactory (Moriasi et al. 2007).

To further quantify infiltration model performance for each model parameterization scenario, a performance score was calculated from validation runs based on three metrics: root mean square error (RMSE) peak, RMSE volume, and average NSE. RMSE for each storm event and model scenario is calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (O_i - S_i)^2}{N}} \tag{5}$$

where O and I are observed and simulated values of peak discharge (for RMSE peak) or runoff volume (for RMSE volume) for all validation events in a study basin.

For each test basin, results were then ranked from 1 to 4 for each metric, with the highest rank assigned to minimum error and maximum NSE, respectively. Ranks for all seven test basins were added up to calculate a combined score for each infiltration model. This indicates each metric was assigned equal weight in calculating the overall score. The maximum possible score is 84 (3 performance metrics × 4 points per metric × 7 basins = 84).

Results

The results section is organized as follows: first, we present a comparison of model performance by study basin based on published guidance, constrained calibration, and unconstrained calibration. We then include statistical analysis of differences in parameterization scenarios, separate for each infiltration model. We conclude the results section with an overall comparison of infiltration model performance.

Comparison of Model Performance Using Published Guidance

Fig. 3 illustrates the error metrics of four infiltration models across all seven basins using published guidance for parameter estimation. Fig. 3(a) shows the absolute peak error (mm h^{-1}), Fig. 3(b) displays the absolute volume error (mm), and Fig. 3(c) indicates the percentage of Nash–Sutcliffe efficiency (NSE) equal to or greater than 0.5, a threshold commonly recognized to indicate acceptable model performance (Moriasi et al. 2007).

Results show that the GA model generally performed best across all test basins, followed by the LC, IC, and CN methods, in that order. The GA model achieved higher or equal NSE across all basins and the lowest average absolute peak error. The IC model displayed lower NSE compared with GA and LC but exceeded the CN model across four basins and matched it in two. Average absolute peak and volume errors for the IC model were lower than those of the CN model in all tested basins, except for basin 11.

For consistency with other parameterization scenarios, all results shown in Fig. 3 correspond to model runs using the validation storms only. Plots of measured and simulated runoff hydrographs for each validation event and all study basins are included in Figs. S1–S7.

Comparison of Model Performance Based on Constrained Calibration

Fig. 4 shows the same comparison for the model scenario *constrained calibration*, where infiltration model parameters were calibrated within a search range constrained by published guidance. Model performance follows patterns similar to those observed in Fig. 3, with the GA method generally performing best, followed by LC, IC, and CN.

Results displayed in Fig. 4 pertain to model runs using the validation storms only (see Figs. S1–S7 for plots of measured

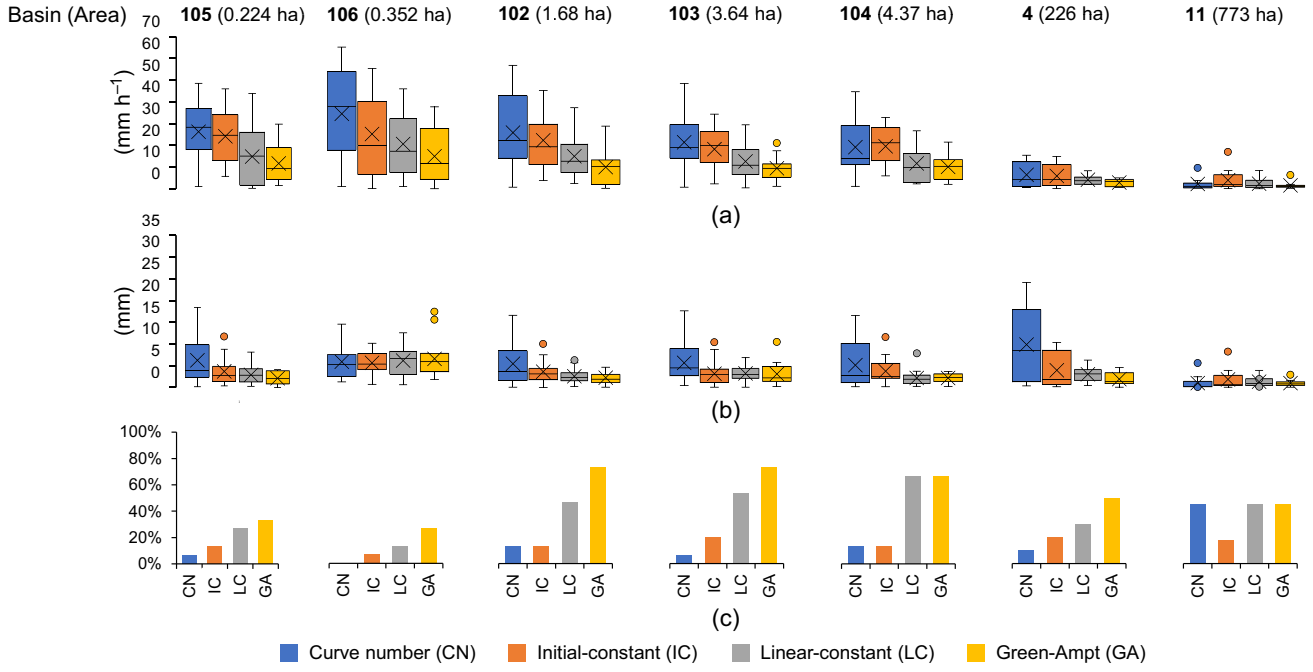

Fig. 3. Comparison of model performance for validation storms by infiltration model (CN, IC, LC, and GA), across seven test basins. Model parameters were estimated based on published guidance: (a) absolute peak error (mm h^{-1}); (b) absolute volume error (mm); and (c) Nash-Sutcliffe efficiency (NSE) ≥ 0.5.

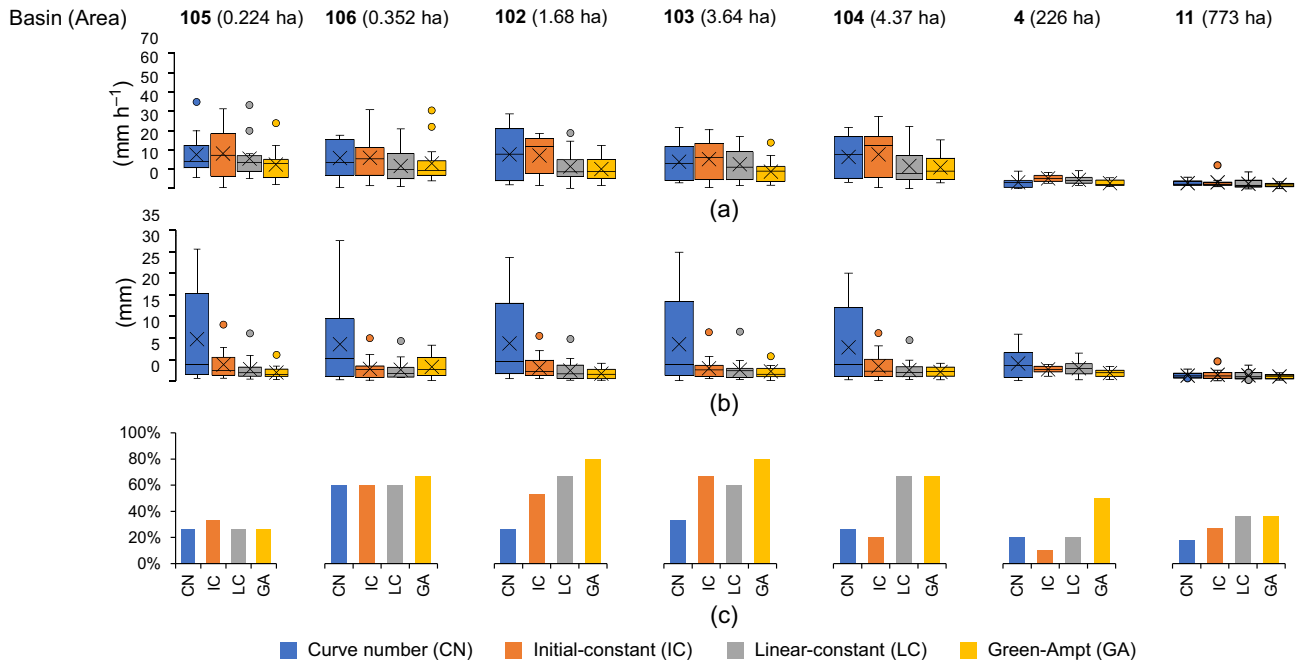

Fig. 4. Comparison of model performance for validation storms by infiltration model (CN, IC, LC, and GA), across seven test basins. Model parameters were estimated based on calibration with a constrained parameter range: (a) absolute peak error (mm h^{-1}); (b) absolute volume error (mm); and (c) Nash–Sutcliffe efficiency (NSE) ≥ 0.5.

and simulated hydrographs). Performance metrics for calibration runs are also included in the Supplemental Materials (see Tables S4–S17).

Comparison of Model Performance Based on Unconstrained Calibration

Results for unconstrained calibration are displayed in Fig. 5. Compared with Figs. 3 and 4, one crucial difference is apparent: the GA method clearly performed worst with respect to all three error metrics. Not a single model run resulted in NSE ≥ 0.5, and absolute peak errors were consistently greater than those of the other three models across all test basins. In line with the other two scenarios, results presented in Fig. 5 are from validation runs only (see Figs. S1–S7 for plots of measured and simulated hydrographs).

Comparison of Parameterization Scenarios by Infiltration Model

Fig. 6 shows a comparison of error metrics (validation events only) separately for each of the four infiltration models across all seven test basins. The left column contains the average absolute peak error (mm h^{-1}), and the column on the right average absolute volume error. Pair-wise assessment of mean error values was performed using the Wilcoxon signed ranks test; pairs with statistically significant differences ($\alpha = 0.05$) are indicated with an asterisk in Fig. 6. The nonparametric test was selected in this study because it does not assume normal distribution. Note that small differences in mean values can be significant if the values of one population are consistently higher. This is the case for the CN method, constrained and unconstrained calibration scenarios [see Fig. 6(b)].

For the CN method, results showed no clear pattern, with all three parameterization scenarios outperforming the others in some cases. Notably, the guidance method generally performed better with respect to runoff volume, whereas the calibration methods performed better with respect to peak. For the IC method, calibration with an unconstrained search range yielded the best results, whereby the difference was statistically significant in four and two cases with respect to peak and volume error, respectively. All three parameterization scenarios showed similar performance for the LC method. The largest and most consistent difference was observed for the GA model: unconstrained calibration consistently led to inferior model performance. The difference between unconstrained and constrained calibration was statistically significant in all cases. For all but the largest basin, using guidance values yielded better results than unconstrained calibration.

Comparison of Infiltration Model Performance

Fig. 7 shows the performance scores for all four infiltration models across three model parameterization scenarios based on validation events. Using published guidance or constrained calibration, the GA model performed the best, followed by the LC model. IC and CN models resulted in similar, markedly poorer performance. For unconstrained calibration, the IC model performed best, followed closely by the LC model, whereas the CN and GA models resulted in poor performance.

Discussion

Four infiltration models of varying complexity available in the widely used HEC-HMS software were evaluated in this study, including the new linear-constant model added in 2023. In the ensuing sections, we will discuss model performance for different parameterization scenarios, always with the compromise between complexity, parsimony, and feasibility in mind.

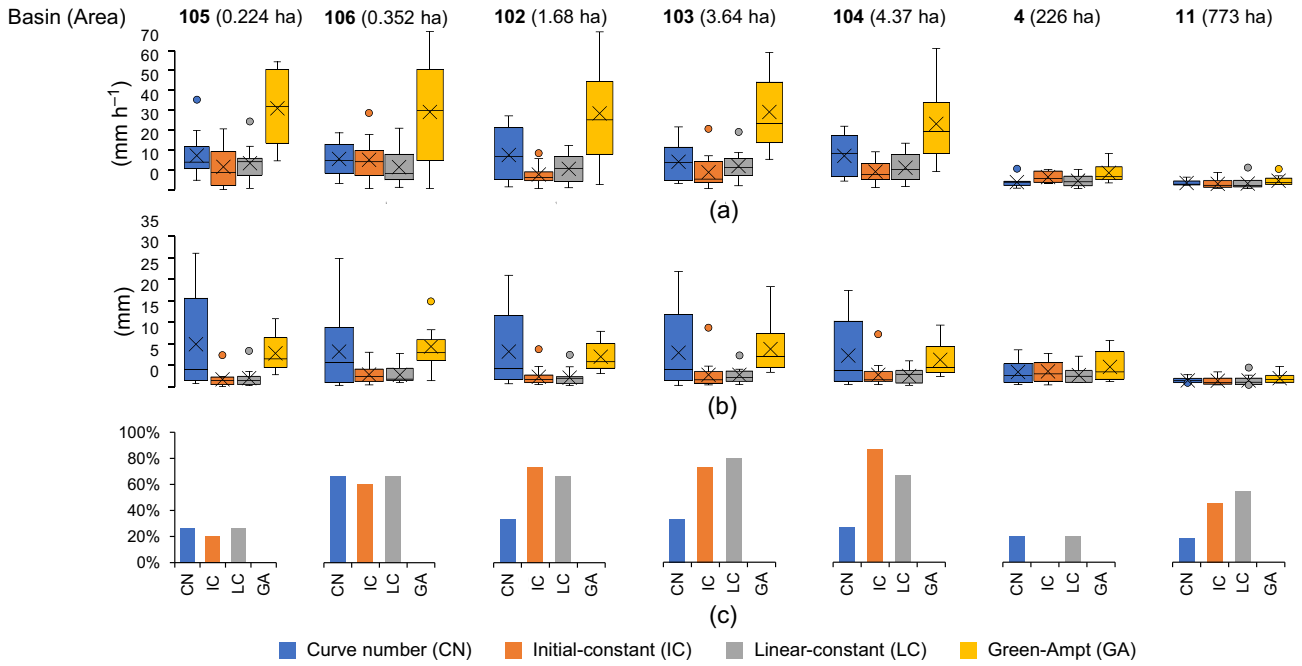

Fig. 5. Comparison of model performance for validation storms by infiltration model (CN, IC, LC, and GA), across seven test basins. Model parameters were estimated based on calibration with an unconstrained parameter range: (a) absolute peak error (mm h^{-1}); (b) absolute volume error (mm); and (c): Nash–Sutcliffe efficiency (NSE) ≥ 0.5.

Discussion of Model Performance Using Published Guidance

Using published guidance to parameterize infiltration models, performance across seven test basins generally followed the same pattern: the Green-Ampt (GA) model performed best, followed by the linear-constant (LC), initial-constant (IC), and curve number (CN) models, in that order. This pattern is evident in plots of error metrics (Fig. 3) and performance scores (Fig. 7). Differences in model performance followed a gradient of complexity: with four parameters, the GA method was the most complex model analyzed in this study and generally performed the best. On the other end of the spectrum, the CN model (one parameter) yielded the poorest performance. Performance of the two-parameter IC and LC models fell in between, whereby the LC model resulted in a higher performance score (Fig. 7). Findings suggest that over-simplification of underlying system dynamics reduces model performance (Hrachowitz et al. 2014), and a certain level of complexity is required for accurate predictions.

Discussion of Model Performance Based on Unconstrained Calibration

Model performance based on unconstrained calibration yielded strikingly different results (Fig. 5): under this scenario, the GA method performed worst as indicated by the lowest performance score (Fig. 7). This was due to high peak errors and poor agreement between measured and simulated hydrographs during model validation across all test basins (see Fig. 7; there was not a single model run with NSE ≥ 0.5). The cause for the poor performance lies in the equifinality conundrum. Two parameters for the GA model were calibrated: suction and hydraulic conductivity. Good model performance was achieved during calibration, with 85 out of 95 calibration runs across all seven test basins resulting in NSE values ≥ 0.5 (see Tables S5, S7, S9, S11, S13, S15, and S17). However, optimal parameter combinations for calibration runs occupy the extreme ends of the spectrum: high suction and low hydraulic conductivity, or low suction and high conductivity (see, for example, Table S17). This constitutes a clear example of equifinality: more than one parameter combination led to similar model performance. Without additional knowledge about underlying processes, it is not clear which combination gives the right answer for the right reason. During validation, where average values of optimized parameters were used, the GA model then yielded poor predictions.

Another interesting finding from the unconstrained calibration scenario is the superior performance of the IC model (see Figs. 5 and 6). Both parameters of this model (initial abstraction and hydraulic conductivity) were adjusted during calibration, however, results show no signs of equifinality. Optimal parameter values vary somewhat across basins (see Tables S4–S17) but are generally a combination of small initial abstraction values and large values of hydraulic conductivity. Although there appears to be a unique parameter combination, it is doubtful the IC model produces good results for the right reasons. Schoener et al. (2021) showed that the initial abstraction term can lead to large model errors, particularly for small basins. The IC model will not simulate any runoff until initial abstraction is fulfilled, regardless of the precipitation rate. During intense convective storms typical of the summer monsoon season in the study area, this can lead to an underestimation of the runoff response early in the storm (Schoener et al. 2021). As a result, parameter combinations with small values of initial abstraction and large values of hydraulic conductivity yield superior results during calibration and validation. Taking basin 102 as an example, the average optimized K_{eff} value is 39 mm h^{-1}, with a range spanning from 28.6 to 61.2 mm h^{-1} (see Table S5). Becker et al. (2018) estimated K_{eff} values in the study area at 20 mm h^{-1} based on field investigation. It therefore seems that the optimized K_{eff} parameter is much higher than the effective hydraulic conductivity expected for soil textures in the study area. If the goal is to generate a model that reproduces runoff irrespective of underlying processes, and if calibration data are available, the IC model may well produce

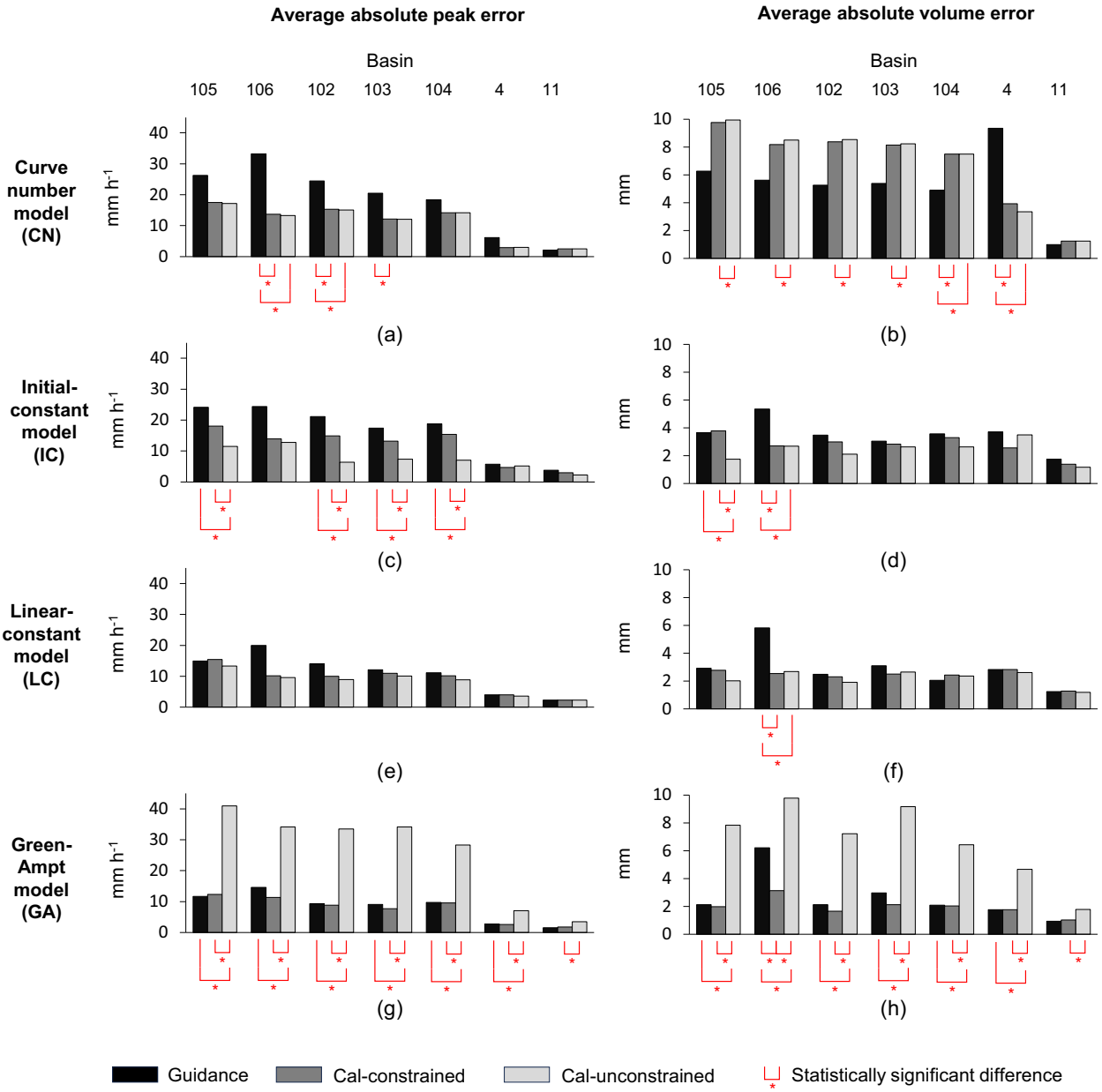

Fig. 6. Comparison of model performance for validation storms by infiltration model (CN, IC, LC, GA) across seven test basins for three different parameterization scenarios; model parameters were estimated based on published guidance (guidance), calibration with a constrained parameter range (cal-constrained), and calibration with an unconstrained parameter range (cal-unconstrained). Statistically significant differences between individual scenarios are indicated with an asterisk.

satisfactory results. It is, however, doubtful that parameters from a model calibrated for one area can be transferred to other sites.

Discussion of Model Performance Based on Constrained Calibration

Several authors have found that constraining the parameter range during calibration can effectively mitigate equifinality issues. Hrachowitz et al. (2014) found that complex models if calibrated with constraints, resulted in the best outcomes. Kelleher et al. (2017) used constraints based on observation and expert knowledge to reduce equifinality for a distributed catchment model. Athira (2021) concluded that constraining parameters during calibration resulted in improved representation of hydrologic processes. Results from this study—particularly for the GA model—highlight the importance of constraining the parameter search range during model calibration. Fig. 4 illustrates that equifinality issues for the GA model can be effectively mitigated by constraining the parameter range used for model calibration based on published guidance. Performance scores based on three error metrics for constrained calibration (Fig. 7) are nearly identical to those obtained when models are parameterized based on published guidance alone.

Discussion of Parameterization Scenarios

In addition to comparing performance among different infiltration models, we assessed the impact of different parameterization scenarios for each infiltration model (see Fig. 6). For the CN method, using published guidance generally resulted in larger peaks and smaller volume errors. This can be attributed to the fact that

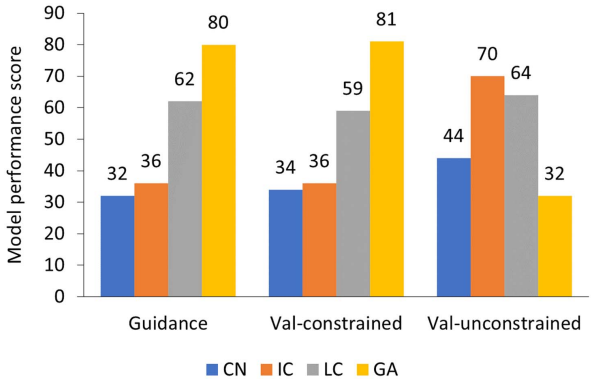

Fig. 7. Comparison of model performance scores for validation storms by parameterization scenario for four different infiltration models. Performance scores were calculated from three error metrics (RMSE peak, RMSE volume, and NSE). Higher scores indicate better model performance; the maximum possible score is 84.

guidance was developed by minimizing runoff volume error (Schoener et al. 2023), whereas calibration in this study minimized RMSE, a peak-weighted performance metric. Unconstrained calibration yielded the best results for the IC model for the reasons discussed above, whereby no significant differences between constrained calibration and guidance scenarios were observed for performance metrics [Figs. 6(c and d)]. For the LC and GA models, constrained calibration and guidance-based parameterization largely resulted in no significant differences. This comparison has implications for the practicing hydrologist: in the absence of calibration data, published guidance may serve as an acceptable strategy for parameterizing the LC and GA models. Granted, the present study examines specific environmental conditions featuring a particular soil texture, plant cover, and climate. Nevertheless, study results demonstrate that guidance for the LC and GA models developed in a different study area and at a different spatial scale (Schoener et al. 2023) can be applied successfully. Future work should extend the analysis to other areas and develop parameter guidance for varying soil and cover characteristics that may be used in ungauged basins. Guidance can then be used to constrain the parameter range during calibration or can be applied directly when no calibration data are available.

Model Performance and Storm Magnitude

The primary focus of this study as motivated in the introduction was the comparison of different infiltration models for flood simulation. Given this focus, we were mostly concerned about the largest storm events most likely to induce a flash flood response. We consequently chose the largest storms on record from each basin for our analysis. Nevertheless, the events we used for calibration and validation cover a range of magnitudes as indicated by event statistics included in the Supplemental Materials (see Tables S1–S3). To test whether model performance changed with storm magnitude, we plotted Nash–Sutcliffe efficiency against peak discharge for each basin, infiltration model, and model scenario (see Figs. S8–S11). Each plot includes a linear regression line along with the coefficient of determination (R^2). Most comparisons show an increase in NSE with increasing peak discharge. This indicates a tendency for models to perform better during larger events, although R^2 values suggest a weak to moderate association. This finding is consistent with other studies showing a relationship between runoff response and event magnitude in small catchments (e.g., Kokkonen et al. 2004). Kjeldsen et al. (2016) found that watershed response time decreases with increasing rainfall depth. This nonlinear flood response might not be adequately represented using a linear unit hydrograph approach such as the one employed in the present study. Findings from models calibrated based on large events may therefore not be applicable to small storms; this is identified as a limitation of the present study, and one that could be overcome using nonlinear routing options in HEC-HMS.

Assessment of the New Linear-Constant (LC) Infiltration Model

The linear-constant infiltration model is a new method incorporated into HEC-HMS in 2023. Apart from preliminary testing (Schoener et al. 2021), the model has not been thoroughly applied in real-world applications prior to this study. The present paper serves as the first test case to evaluate model performance and compare results against other commonly applied infiltration models. The LC model was developed as a modification of the IC model, with the goal to increase predictive power while maintaining a parsimonious model structure (Schoener et al. 2021). Results are encouraging. Compared with the CN and IC models, performance of the LC model is markedly improved with respect to peak discharge, runoff volume, and overall fit of simulated hydrographs as measured by NSE. Like the IC model, the LC method only has two parameters; it is therefore simple enough to be applied in practice, particularly in data-sparse areas.

The more complex (four-parameter) GA model still outperforms the LC model; however, the LC model appears less susceptible to equifinality, particularly under unconstrained calibration. This is somewhat surprising: equifinality has been shown to decrease with an increase in observations and a decrease in calibration parameters (Her and Chaubey 2015; Her et al. 2019). In this study, the same number of parameters were calibrated for both the LC and GA models, and the same number of observations were used to compare both models.

Given these findings, the LC model may be preferrable when: (1) a simple infiltration model is desired, as is often the case in practical applications with limited resources to conduct hydrologic analyses, and little or no observations; and (2) when information to constrain parameter ranges during model calibration is limited. Overall, the results from this study suggest that the LC model may serve as a suitable compromise between complexity, parsimony, and feasibility. More work is needed to vet the model in watersheds with differing soil, land use, and climate characteristics.

Conclusions

In this study, four infiltration models were evaluated within the HEC-HMS software: the curve number (CN), initial-constant (IC), linear-constant (LC) and Green-Ampt (GA) models. The GA model, being the most complex model, excelled over other infiltration models when parameterized using published guidance (Schoener et al. 2023) and under constrained calibration scenarios. However, the GA model faced the equifinality issue under unconstrained calibration. The IC model performed best under unconstrained calibration, but optimized parameters may not be representative of underlying physical processes.

The LC model, the new addition to HEC-HMS, showed promising results by outperforming CN and IC models, and being less susceptible to equifinality issue compared with the more complex GA model. The LC model may therefore offer a balance between complexity, parsimony, and feasibility; it may provide a suitable choice when: (1) a simple infiltration model is desired, as is often

the case in practical applications, and (2) when data to constrain parameters during calibration are limited.

Data Availability Statement

All data and models that support the findings of this study are available from the corresponding author upon request.

Acknowledgments

The authors express their sincere gratitude to David Gatterman, Executive Engineer at the Southern Sandoval County Arroyo Flood Control Authority (SSCAFCA), for his invaluable review and insightful comments on this work. We also extend our appreciation to the SSCAFCA board of directors; their unwavering support was pivotal in the successful completion of this study.

Supplemental Materials

Tables S1–S17 and Figs. S1–S11 are available online in the ASCE Library (www.ascelibrary.org).

References

Athira, P. 2021. "Calibration of hydrological models considering process interdependence: A case study of SWAT model." *Environ. Modell. Software* 144 (Oct): 105131. https://doi.org/10.1016/j.envsoft.2021.105131.

Baiamonte, G. 2019. "SCS curve number and green-Ampt infiltration models." *J. Hydrol. Eng.* 24 (10): 04019034. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001838.

Becker, R., M. Gebremichael, and M. Märker. 2018. "Impact of soil surface and subsurface properties on soil saturated hydraulic conductivity in the semi-arid Walnut Gulch Experimental Watershed, Arizona, USA." *Geoderma* 322 (Jul): 112–120. https://doi.org/10.1016/j.geoderma.2018.02.023.

Beven, K. 2006. "A manifesto for the equifinality thesis." *J. Hydrol.* 320 (1–2): 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007.

Breckenfeld, D. J., W. A. Svetlik, and C. E. McGuire. 2008. *Soil survey of Walnut Gulch Experimental Watershed, Arizona*. Tucson, AZ: USDA-SCS and USDA-ARS in Cooperation with Arizona Agricultural Experiment Station.

Chahinian, N., R. Moussa, P. Andrieux, and M. Voltz. 2005. "Comparison of infiltration models to simulate flood events at the field scale." *J. Hydrol.* 306 (1–4): 191–214. https://doi.org/10.1016/j.jhydrol.2004.09.009.

ESRI (Environmental Systems Research Institute). 2023. *ArcGIS desktop version 10.8.2 [computer software]*. Redlands, CA: ESRI.

Goodrich, D. C., T. O. Keefer, C. L. Unkrich, M. H. Nichols, H. B. Osborn, J. J. Stone, and J. R. Smith. 2008. "Long-term precipitation database, Walnut Gulch Experimental Watershed, Arizona, United States." *Water Resour. Res.* 44 (5): W05S04. https://doi.org/10.1029/2006WR005782.

Goodrich, D. C., L. J. Lane, R. M. Shillito, S. N. Miller, K. H. Syed, and D. A. Woolhiser. 1997. "Linearity of basin response as a function of scale in a semiarid watershed." *Water Resour. Res.* 33 (12): 2951–2965. https://doi.org/10.1029/97WR01422.

Green, W. H., and G. A. Ampt. 1911. "Studies on soil physics." *J. Agric. Sci.* 4 (1): 1–24. https://doi.org/10.1002/hyp.11453.

Grimaldi, S., A. Petroselli, and N. Romano. 2013. "Green-Ampt Curve-Number mixed procedure as an empirical tool for rainfall–runoff modelling in small and ungauged basins." *Hydrol. Processes* 27 (8): 1253–1264. https://doi.org/10.1002/hyp.9303.

Guse, B., J. Kiesel, M. Pfannerstill, and N. Fohrer. 2020. "Assessing parameter identifiability for multiple performance criteria to constrain model parameters." *Hydrol. Sci. J.* 65 (7): 1158–1172. https://doi.org/10.1080/02626667.2020.1734204.

Hapuarachchi, H. A. P., Q. J. Wang, and T. C. Pagano. 2011. "A review of advances in flash flood forecasting." *Hydrol. Processes* 25 (18): 2771–2784. https://doi.org/10.1002/hyp.8040.

Hawkins, R. H., T. J. Ward, D. E. Woodward, and J. A. Van Mullem. 2008. *Curve number hydrology: State of the practice*. Reston, VA: ASCE.

Her, Y., and I. Chaubey. 2015. "Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty." *Hydrol. Processes* 29 (19): 4220–4237. https://doi.org/10.1002/hyp.10487.

Her, Y., S. H. Yoo, J. Cho, S. Hwang, J. Jeong, and C. Seong. 2019. "Uncertainty in hydrological analysis of climate change: Multi-parameter versus multi-GCM ensemble predictions." *Sci. Rep.* 9 (1): 4974. https://doi.org/10.1038/s41598-019-41334-7.

Hrachowitz, M., O. Fovet, L. Ruiz, T. Euser, S. Gharari, R. Nijzink, J. Freer, H. H. G. Savenije, and C. Gascuel-Odoux. 2014. "Process consistency in models: The importance of system signatures, expert knowledge, and process complexity." *Water Resour. Res.* 50 (9): 7445–7469. https://doi.org/10.1002/2014WR015484.

Huang, P., Z. Li, J. Chen, Q. Li, and C. Yao. 2016. "Event-based hydrological modeling for detecting dominant hydrological process and suitable model strategy for semi-arid catchments." *J. Hydrol.* 542 (Nov): 292–303. https://doi.org/10.1016/j.jhydrol.2016.09.001.

Huo, W., Z. Li, K. Zhang, J. Wang, and C. Yao. 2020. "GA-PIC: An improved Green-Ampt rainfall-runoff model with a physically based infiltration distribution curve for semi-arid basins." *J. Hydrol.* 586 (Jul): 124900. https://doi.org/10.1016/j.jhydrol.2020.124900.

Jackson, E. K., W. Roberts, B. Nelsen, G. P. Williams, E. J. Nelson, and D. P. Ames. 2019. "Introductory overview: Error metrics for hydrologic modelling—A review of common practices and an open source library to facilitate use and adoption." *Environ. Modell. Software* 119 (Sep): 32–48. https://doi.org/10.1016/j.envsoft.2019.05.001.

Kampf, S. K., J. Faulconer, J. R. Shaw, M. Lefsky, J. W. Wagenbrenner, and D. J. Cooper. 2018. "Rainfall thresholds for flow generation in desert ephemeral streams." *Water Resour. Res.* 54 (12): 9935–9950. https://doi.org/10.1029/2018WR023714.

Kelleher, C., B. McGlynn, and T. Wagener. 2017. "Characterizing and reducing equifinality by constraining a distributed catchment model with regional signatures, local observations, and process understanding." *Hydrol. Earth Syst. Sci.* 21 (7): 3325–3352. https://doi.org/10.5194/hess-21-3325-2017.

Kirchner, J. W. 2006. "Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology." *Water Resour. Res.* 42 (3): W03S04. https://doi.org/10.1029/2005WR004362.

Kjeldsen, T. R., H. Kim, C. H. Jang, and H. Lee. 2016. "Evidence and implications of nonlinear flood response in a small mountainous watershed." *J. Hydrol. Eng.* 21 (8): 04016024. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001343.

Kokkonen, T., H. Koivusalo, T. Karvonen, B. Croke, and A. Jakeman. 2004. "Exploring streamflow response to effective rainfall across event magnitude scale." *Hydrol. Processes* 18 (8): 1467–1486. https://doi.org/10.1002/hyp.1423.

Mai, J. 2023. "Ten strategies towards successful calibration of environmental models." *J. Hydrol.* 620 (May): 129414. https://doi.org/10.1016/j.jhydrol.2023.129414.

Mengistu, A. G., L. D. van Rensburg, and Y. E. Woyessa. 2019. "Techniques for calibration and validation of SWAT model in data scarce arid and semi-arid catchments in South Africa." *J. Hydrol.: Reg. Stud.* 25 (Oct): 100621. https://doi.org/10.1016/j.ejrh.2019.100621.

Mishra, S. K., J. V. Tyagi, and V. P. Singh. 2003. "Comparison of infiltration models." *Hydrol. Processes* 17 (13): 2629–2652. https://doi.org/10.1002/hyp.1257.

Moglen, G. E., H. Sadeq, L. H. Hughes, M. E. Meadows, J. J. Miller, J. J. Ramirez-Avila, and E. W. Tollner. 2022. "NRCS curve number method: Comparison of methods for estimating the curve number from rainfall-runoff data." *J. Hydrol. Eng.* 27 (10): 04022023. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002210.

Morbidelli, R., C. Corradini, C. Saltalippi, A. Flammini, J. Dari, and R. S. Govindaraju. 2018. "Rainfall infiltration modeling: A review." *Water* 10 (12): 1873. https://doi.org/10.3390/w10121873.

Moriasi, D. N., J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. 2007. "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations." *Trans. ASABE* 50 (3): 885–900. https://doi.org/10.13031/2013.23153.

Nash, J. E., and J. V. Sutcliffe. 1970. "River flow forecasting through conceptual models part I—A discussion of principles." *J. Hydrol.* 10 (3): 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.

Nearing, M. A., M. H. Nichols, J. J. Stone, K. G. Renard, and J. R. Simanton. 2007. "Sediment yields from unit-source semiarid watersheds at Walnut Gulch." *Water Resour. Res.* 43 (6): 1–10. https://doi.org/10.1029/2006WR005692.

Nijzink, R. C., et al. 2018. "Constraining conceptual hydrological models with multiple information sources." *Water Resour. Res.* 54 (10): 8332–8362. https://doi.org/10.1029/2017WR021895.

Ogden, F. L., R. P. Hawkins, M. T. Walter, and D. C. Goodrich. 2017. "Beyond the SCS-CN method: A theoretical framework for spatially lumped rainfall-runoff response." *Water Resour. Res.* 53 (7): 6345–6350. https://doi.org/10.1002/2016WR020176.

Pande, S., H. H. Savenije, L. A. Bastidas, and A. K. Gosain. 2012. "A parsimonious hydrological model for a data scarce dryland region." *Water Resour. Manage.* 26 (4): 909–926. https://doi.org/10.1007/s11269-011-9816-z.

Phillips, J. V., and S. Tadayon. 2006. *Selection of Manning's roughness coefficient for natural and constructed vegetated and nonvegetated channels, and vegetation maintenance plan guidelines for vegetated channels in central Arizona*. Reston, VA: USGS.

Quichimbo, E. A., M. B. Singer, K. Michaelides, D. E. Hobley, R. Rosolem, and M. O. Cuthbert. 2021. "DRYP 1.0: A parsimonious hydrological model of DRYland partitioning of the water balance." *Geosci. Model Dev.* 14 (11): 6893–6917. https://doi.org/10.5194/gmd-14-6893-2021.

Rawls, W. J., D. L. Brakensiek, and N. Miller. 1983. "Green-ampt infiltration parameters from soils data." *J. Hydraul. Eng.* 109 (1): 62–70. https://doi.org/10.1061/(ASCE)0733-9429(1983)109:1(62).

Renard, K. G., M. H. Nichols, D. A. Woolhiser, and H. B. Osborn. 2008. "A brief background on the U.S. Department of Agriculture Agricultural Research Service Walnut Gulch Experimental Watershed." *Water Resour. Res.* 44 (5): W05S02. https://doi.org/10.1029/2006WR005691.

Schoener, G. 2017. "Quantifying transmission losses in a New Mexico ephemeral stream: A losing proposition." *J. Hydrol. Eng.* 22 (3): 05016038. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001473.

Schoener, G. 2018. "Time-lapse photography: Low-cost, low-tech alternative for monitoring flow depth." *J. Hydrol. Eng.* 23 (2): 06017007. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001616.

Schoener, G. 2022. "Impact of urbanization and stormwater infrastructure on ephemeral channel transmission loss in a semiarid watershed." *J. Hydrol.: Reg. Stud.* 41 (Jun): 101089. https://doi.org/10.1016/j.ejrh.2022.101089.

Schoener, G., S. Rassa, M. Fleming, D. Gatterman, and J. Montoya. 2023. "Infiltration model parameters from rainfall simulation for sandy soils." *J. Hydrol. Eng.* 29 (1): 06023001. https://doi.org/10.1061/JHYEFF.HEENG-6025.

Schoener, G., M. C. Stone, and C. Thomas. 2021. "Comparison of seven simple loss models for runoff prediction at the plot, hillslope and catchment scale in the semiarid southwestern U.S." *J. Hydrol.* 598 (Jul): 126490. https://doi.org/10.1016/j.jhydrol.2021.126490.

Schreiner-McGraw, A. P., and E. R. Vivoni. 2018. "On the sensitivity of hillslope runoff and channel transmission losses in arid piedmont slopes." *Water Resour. Res.* 54 (7): 4498–4518. https://doi.org/10.1029/2018WR022842.

Southwest Watershed Research Center. 2024. "Data access project." US Department of Agriculture, Agricultural Research Service. Accessed March 7, 2024. https://www.tucson.ars.ag.gov/dap/.

Stewart, D., E. Canfield, and R. Hawkins. 2012. "Curve number determination methods and uncertainty in hydrologic soil groups from semiarid watershed data." *J. Hydrol. Eng.* 17 (11): 1180–1187. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000452.

Tegegne, G., D. K. Park, and Y. O. Kim. 2017. "Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin." *J. Hydrol.: Reg. Stud.* 14 (Dec): 49–66. https://doi.org/10.1016/j.ejrh.2017.10.002.

USACE. 2000. *Hydrologic modeling system HEC-HMS technical reference manual*. Davis, CA: USACE.

USDA. 2004a. *National engineering handbook, Chapter 9, Hydrologic soil-cover complexes*. Washington, DC: USDA, Natural Resources Conservation Service.

USDA. 2004b. *National engineering handbook, Chapter 10, Estimation of direct runoff from storm rainfall*. Washington, DC: USDA, Natural Resources Conservation Service.

USDA. 2007. *National engineering handbook, Chapter 16, Hydrographs*. Washington, DC: USDA, Natural Resources Conservation Service.

USDA. 2010. *National engineering handbook, Chapter 15, Time of concentration*. Washington, DC: USDA, Natural Resources Conservation Service.

USGS. 2024. "National geospatial program: The national map." US Department of the Interior. Accessed April 19, 2023. https://www.usgs.gov/programs/national-geospatial-program/national-map.

Wambura, F. J., O. Dietrich, and G. Lischeid. 2018. "Improving a distributed hydrological model using evapotranspiration-related boundary conditions as additional constraints in a data-scarce river basin." *Hydrol. Processes* 32 (6): 759–775. https://doi.org/10.1002/hyp.11453.