

## Analysis of scaling relationships for flood parameters and peak discharge estimation in a tropical region

Charles Mazivanhanga<sup>a</sup>, Robert C. Grabowski<sup>id</sup><sup>a,\*</sup>, Eunice Pérez-Sánchez<sup>b</sup> and Víctor R. Carballo-Cruz<sup>b</sup>

<sup>a</sup> School of Water Energy and Environment, Cranfield University, Bedford, United Kingdom

<sup>b</sup> División Académica de Ciencias Biológicas, Universidad Juárez Autónoma de Tabasco, Tabasco, Mexico

\*Corresponding author. E-mail: r.c.grabowski@cranfield.ac.uk

RCG, 0000-0002-0926-1202

### ABSTRACT

Relationships between peak discharges and catchment size (e.g., flood scaling) in a catchment have the potential to support new river flood forecasting approaches but have not been tested in tropical regions. This study determined flood scaling relationships between peak discharge and nested drainage areas in the La Sierra catchment (Mexico). A statistical power law equation was applied to selected rainfall–runoff events that occurred between 2012 and 2015. Variations in flood scaling parameters were determined in relation to catchment descriptors and processes for peak downstream discharge estimation. Similar to studies in humid temperate regions, the results reveal the existence of log-linear relationships between the intercept ( $\alpha$ ) and exponent ( $\theta$ ) parameter values and the log–log power–law relationships between ( $\alpha$ ) and the peak discharge observed from the smallest headwater catchments. The flood parameter values obtained were then factored into the scaling equation ( $Q_p = \alpha A^\theta$ ) and successfully predicted downstream flood peaks, especially highly recurrent flood events. The findings contribute to a better understanding of the nature of flood wave generation and support the development of new flood forecasting approaches in unregulated catchments suitable for non-stationarity in hydrological processes with climate change.

**Key words:** catchment, flood forecasting, flood peak, hydrological scaling, rainfall event, streamflow response

### HIGHLIGHTS

- Peak discharge in the La Sierra catchment follows a power–law relationship, similar to humid temperate regions.
- Log-linear and log–log relationships can be used to estimate flood parameters and peak downstream discharge especially for frequent events.

## 1. INTRODUCTION

Flood prediction in ungauged regions is a global challenge. Most tropical regions with increasing flood risk do not have enough hydrometeorological records required for setting up flood models, and, as a result, people are exposed to elevated and increasing flood risk. To address this challenge, the United States Geological Survey (USGS) pioneered a flood estimation method that employs regional quartile equations to build regression relationships between catchment descriptors and variables where no streamflow data are available (Gupta *et al.* 2007). The equations are statistical relationships that use observations of annual maximum discharge over large, homogeneous regions to estimate flood magnitudes (Mandapaka *et al.* 2009b). The use of homogeneous regions enables the pooling of streamflow data from various basins within the same region with streamflow records to address the lack of streamflow data in other ungauged basins within the same region (Gupta *et al.* 2007). Regional flood frequency (RFF) analysis is still used currently to estimate annual peak flow quantiles in ungauged basins.

Many regions of the world, however, remain poorly gauged, resulting in limited data to apply the RFF method. Furthermore, the significant alterations to catchment hydrology caused by climate change and land use change make past observations less useful for flood estimation in many regions. Also, the current flood quantile equations do not account for the physical processes that cause flooding and make providing flood estimates at smaller basins challenging (Gupta *et al.* 2007). A physical solution has long been required to improve the accuracy of flood quantile estimates and compensate

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

for the limited data records (Gupta 2017; Gupta *et al.* 2010; Furey *et al.* 2016). There has been a need for methodologies that estimate flood frequencies based on physical principles of water movement and general knowledge of the geographic and geomorphologic characteristics of upstream catchments for the location of interest (Perez Mesa 2019).

A solution is to estimate flood parameters and peak discharge at a rainfall–runoff event scale rather than current large-scale resolutions based on annual temporal scales and large homogeneous spatial regional scales. To switch from RFF approaches to physics-based flood frequency estimates, it is important to find the patterns between rainfall and runoff at increasing scales of sub-catchments within a river network, i.e., flood scaling (Perez Mesa 2019). In this regard, several studies have provided the theoretical and empirical basis for formulating a geophysical or a scaling theory of floods focused on scalable drainage areas and rainfall event scales (Gupta *et al.* 1996, 2010, 2015; Robinson & Sivapalan 1997; Menabde & Sivapalan 2001; Ogden & Dawdy 2003; Gupta 2004, 2017; Dawdy *et al.* 2012; Farmer *et al.* 2015; Medhi & Tripathi 2015; Furey *et al.* 2016; Lee & Huang 2016; Ayalew *et al.* 2018; Yang *et al.* 2020; Mantilla *et al.* 2011).

The scaling theory of floods is based on the idea that, at the individual rainfall–runoff event scale, there is a scale-invariant spatial structure of peak discharge with drainage areas defined by power–law relationships (Ayalew *et al.* 2018). In other words, the peak flood discharge will increase in a predictable fashion as it flows downstream (and catchment area increases). The theory focuses on river drainage areas, abandoning the concept of large, homogeneous regions in favour of incorporating flood-producing physical processes at a small rainfall–runoff scale (Gupta *et al.* 2010). Studies have found that peak discharge, from a single rainfall–runoff event scale, has a power–law relationship with a drainage area given as

$$Q_p = \alpha A^\theta \quad (1)$$

where  $Q_p$  is the peak discharge,  $A$  is the drainage area,  $\alpha$  is the scaling intercept, and  $\theta$  is the scaling exponent values (Gupta *et al.* 2010; Medhi & Tripathi 2015).

The intercept ( $\alpha$ ) and exponent ( $\theta$ ) of the peak discharge power–law relations are referred to as flood scaling parameters. An important discovery is that the values of these two scaling parameters change from one rainfall event to another, and this variation determines the magnitude of peak discharge (Ogden & Dawdy 2003; Ayalew *et al.* 2018). Thus, a better understanding of catchment factors and the processes governing the variation of these parameters is critical in developing scaling equations capable of estimating peak discharge, contributing towards solving the problem of flood prediction in ungauged basins (Sivapalan 2003; Hrachowitz *et al.* 2013; Yang *et al.* 2020).

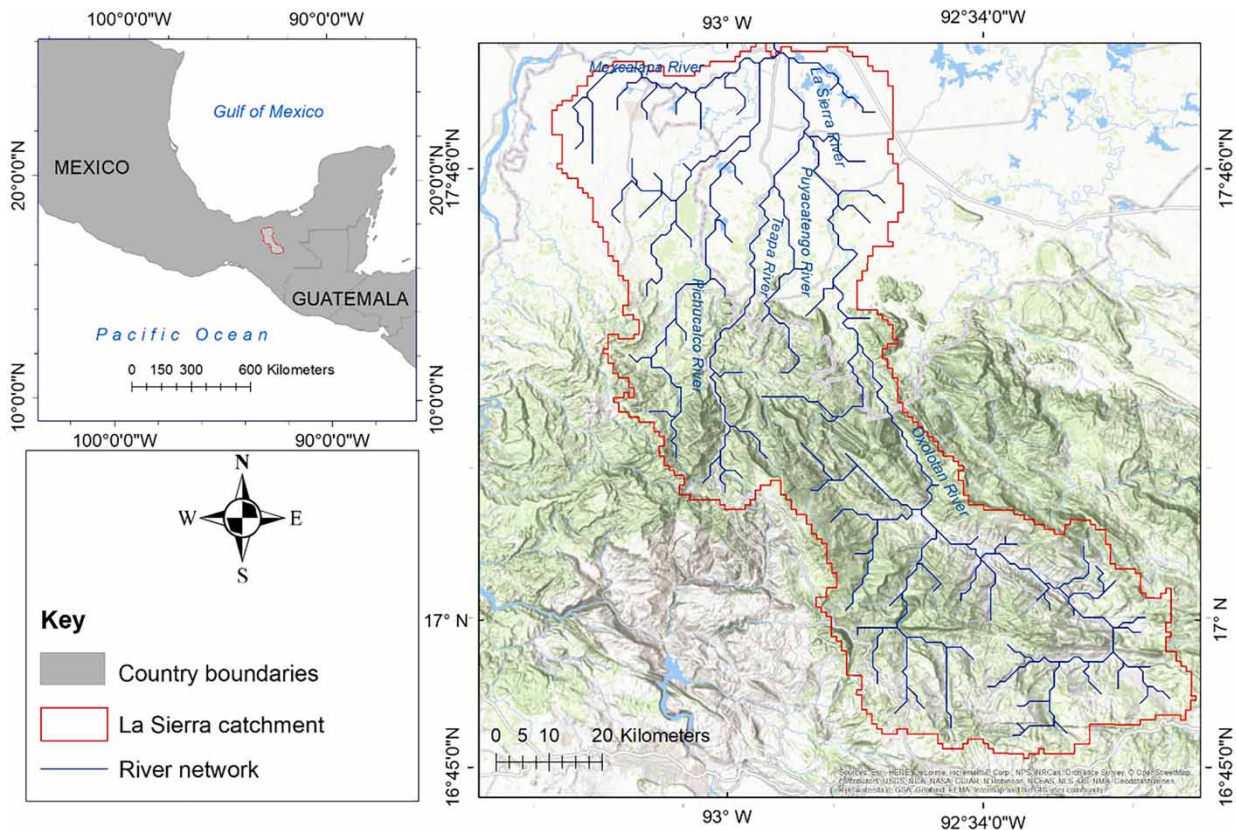
The focus of most research studies has been on establishing and validating the scaling relationships between rainfall, peak discharge, and drainage areas, as well as other factors that control flood parameters (Ayalew *et al.* 2014b). However, most of these studies have been done in humid northern latitude regions, with several studies conducted in the United States and the United Kingdom, and it is not known whether the scaling flood theory is applicable in other climatic regions with different rainfall types and coverages (Wilkinson & Bathurst 2018). There is a need to generalise the scaling theory of floods to medium and large river basins spanning different climatic regions with a greater variety of rainfall event types (Gupta 2017).

This study aims to determine if flood scaling relationships are valid in large tropical regions to support flood forecasting. The three main objectives are (i) to investigate flood scaling relationships between peak discharge and nested drainage areas in an unregulated tropical catchment, (ii) to identify and explain factors influencing variations in the scaling parameters and the confidence levels, and (iii) to provide a unified framework for estimating flood parameter values and peak flood magnitudes across data-scarce tropical catchments.

## 2. METHODS

### 2.1. Study area and data

The La Sierra catchment area, located in southeast Mexico, was used to investigate the scaling theory of floods. The catchment covers an area of 6,743 km<sup>2</sup>, drains a mountainous area with a maximum elevation of 2,216 m, and is located in the east of the Grijalva basin (Figure 1). The Sierra River is a tributary of the Grijalva River, with its confluence located downstream of the city of Villahermosa. It originates in the central highlands of Chiapas, with headwater catchments covering primarily the Sierra Mountains. The topography of the headwater catchments comprises highly dissected mountainous areas with very steep slopes (>40%) and narrow V-shaped river valleys. However, catchment areas in the downstream sections consist of broad and gentle (2%) low-lying plains with a minimum elevation of 6 m above sea level. The La Sierra catchment area is



**Figure 1** | The Sierra catchment area in southeast Mexico (Rabus *et al.* 2003; Esri 2012; Lehner & Grill 2013).

Mexico's wettest catchment, with an annual rainfall of approximately 4,000 mm and monthly averages of 200 mm throughout the year, with extremely high totals occurring from June to October and lower totals occurring from November to May (Mercado *et al.* 2016; INEGI 2018). The Sierra River has a significant impact on flood discharges and is a major source of flooding in Villahermosa City (López-Caloca *et al.* 2017).

The La Sierra catchment area experiences two types of rainfall: large-scale rainfall events that affect the entire catchment and intense convective rainfall events that affect localised areas of the catchment. The three major sources of rainfall in the catchment area are tropical cyclones from the Caribbean Sea and the Atlantic Ocean; the Inter-Tropical Convergence Zone (ITCZ), which extends to higher latitudes during the summer and affects the upper Grijalva basin; and late summer tropical waves, which cause significant rainfall in the northern parts of the catchment (Arreguín-Cortés *et al.* 2014).

Seven nested catchment areas in the La Sierra catchment area were used to investigate the scaling relationships and peak flow generation at various spatial scales, based on the location of existing flow gauging stations (GSs) (Table 1 and Figure 2). Discharge data of the El Puente GS were not used in the analysis because of insufficient hydrometeorological records for the years under consideration. Using the Spatial Analyst extension tool in ArcGIS on a 30-arc-second-resolution digital elevation model (DEM), each nested catchment was delineated along the watershed, beginning with where flow GSs were located in the catchment (Rabus *et al.* 2003). The resulting multi-scale nested catchment structure consisted of units of drainage areas nested in the main catchment, the La Sierra catchment area in this case (Paschalis *et al.* 2014). In this way, the nested structure allowed the quantification of the spatial process by which floods are generated in small headwater catchments, building up into floods at larger-scale catchments (Wilkinson & Bathurst 2018).

## 2.2. Hydrometeorological data

Rainfall data for the La Sierra catchment area were obtained from a total of 12 rain gauge stations. However, these rainfall data only encompassed the time period spanning from 2012 to 2015, as depicted in Table 2. The data were obtained from the Mexican rainfall database (CLIMCOM 2013) maintained by the Mexican Meteorological Service (Servicio Meteorológico

**Table 1** | Characteristics of nested catchment areas in the Sierra main catchment

Catchment name (names derived from outlet flow GSs)	Catchment area ×1,000,000 (m <sup>2</sup> )	Length of the main channel ×1,000 m	Minimum elevation of the main channel (m)	Maximum elevation of the main channel (m)	Average slope of the main channel (%)	Standard average annual rainfall (mm)	Standard runoff (%)
Gaviotas (for La Sierra)	6,743	173,615	6	2,216	1.27	1,730	20–30
Pueblo nuevo	4,748	157,442	9	2,177	1.38	1,730	20–30
Tapijulapa	3,698	131,192	19	2,216	10.67	1,602	>30
Oxolotan	2,416	115,064	109	2,216	20.08	1,602	>30
El Puente	1,787	97,219	7	902	1.00	2,101	20–30
Teapa	438	42,027	35	1,106	20.55	2,513	20–30
Pichulcalco	431	3,992	10	2,100	30.00	2,600	20–30
Puyacatengo	229	15,370	59	644	30.81	2,561	>30

Note: the catchment area and channel length are in km and can be expressed in square metres and metres (metre unit) when multiplied by 100,000 and 1,000, respectively.

Nacional, SMN) and the Mexican Water Commission (CONAGUA). The nested hydrometric network's discharge data were collected from seven GSs located at the outlets of each nested catchment (Figure 2), spanning from 2012 to 2015. The flow data were obtained from CONAGUA, the Mexico Surface Water Management and Rivers Engineering (GASIR) database, and the National Surface Water Data Bank (BANDAS) database. However, discharge data for El Puente and Tapijulapa nested catchment areas were limited, with missing values or having some gaps in their time series. As a result, the El Puente sub-catchment area was not included in the analysis (Table 2).

### 2.3. Investigation of scaling relationships

The analysis sought to investigate the scaling relationships between peak discharge and nested catchment drainage areas, as well as to identify underlying factors driving variations in the relationships, to enable flood prediction in the La Sierra catchment. The scaling relationships between selected rainfall–runoff events between 2012 and 2015 were investigated using a statistical power law equation (Equation (1)). The study investigated scaling relationships not only with drainage area data but also with other catchment descriptors and processes that contribute significantly to the mechanism of flood peak generation in the study area (Formetta *et al.* 2021).

#### 2.3.1. Selection of observed peak discharge events

A total of 59 rainfall events that generated peak discharges greater than a 2-year return period between 2012 and 2015 were identified for analysis. The events were selected following a preliminary analysis of rainfall–discharge records to identify the time of concentration of flows in the catchment area. Figure 3 illustrates an excerpt of the 2015 peak discharge events that were analysed. For the La Sierra catchment, the estimated time of concentration was 3 h, which is the smallest time window size that separates individual rainfall–runoff peak events in the area (Ayalew *et al.* 2015). The time of concentration was calculated using the velocity method (Equations (2) and (3)) that sum flow journey times for segments along the hydraulically most remote open channel flow (NRCS 1997).

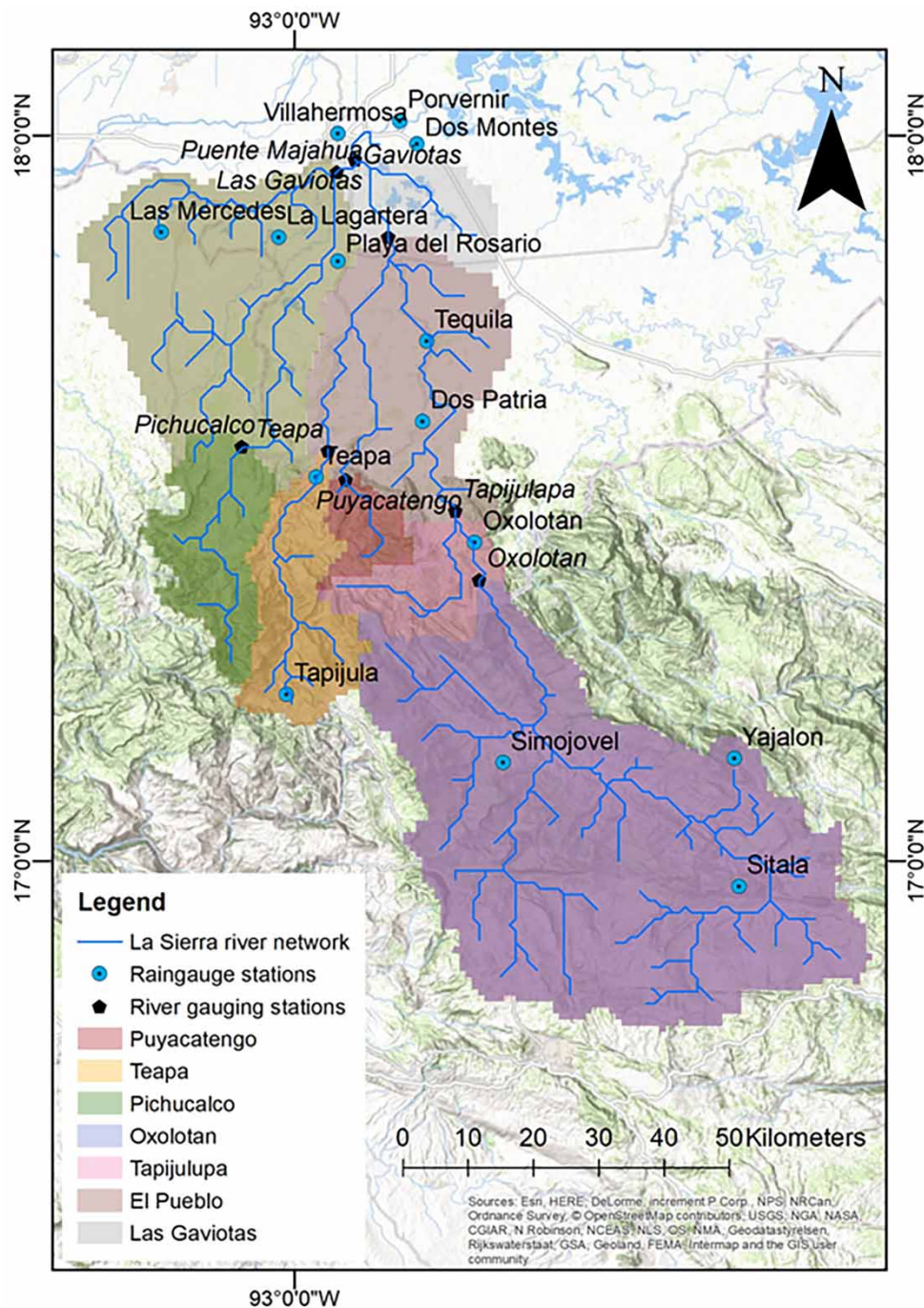
$$T_c = T_{i1} + T_{i2} + T_{i3} + \dots + T_{in} \quad (2)$$

where  $T_c$  is the time of concentration in hours,  $T_i$  is the travel time of a segment, and  $n$  is the number of segments comprising the total hydraulic length.

The travel time for any given segment is equal to

$$T_t = l / 3,600 V \quad (3)$$

$T_t$  is the travel time, in hours,  $l$  is the flow length, in metres,  $V$  is the average velocity (m/s), and 3,600 is the conversion factor (seconds to hours).



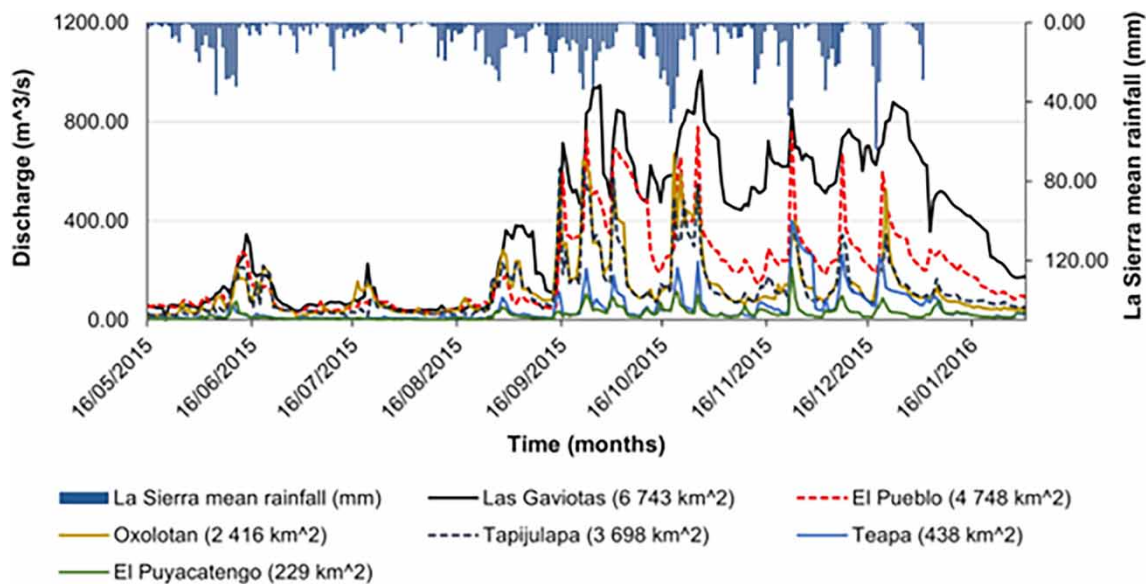
**Figure 2** | Nested catchments and gauge network in the La Sierra catchment (Rabus *et al.* 2003; Esri 2012; CLIMCOM 2013; Lehner & Grill 2013).

### 2.3.2. Peak scaling analysis

The following procedure was followed to establish significant scaling relationships for estimating flood parameter values for the La Sierra catchment area (Ayalew *et al.* 2015). The  $\alpha$ ,  $\theta$ , and coefficient of determination ( $R^2$ ) values were determined using the power-law equation (Equation (1)) by establishing scaling relationships between the observed peak discharge for each rainfall event and its corresponding nested catchment drainage area plotted on a logarithmic scale scatter plot with a straight line fitted between data points. The  $R^2$  was used to evaluate the reliability of the scaling parameters ( $\alpha$  and  $\theta$ ) and thus the accuracy of prediction of the scaling equation (Chen *et al.* 2020). Individual rainfall events were selected

**Table 2** | Discharge data availability across GSs for the nested catchment areas

Gauging Stations	Years with data								
	1940	1949	1959	1968	1978	1987	1996	2006	2015
Puyacatengo									
Teapa									
Pichucalco									
Tapijulapa									
Oxolotan									
Puente de la Majahua									
Pueblo Nuevo									
Gaviotas									

**Figure 3** | Extract of some peak discharge events analysed for the La Sierra nested catchments, 2012–2015. (The 2015–2016 time series for El Puente and Pichucalco nested catchment regions are not shown owing to limited data.)

based on their  $R^2$  values; those with values greater than 0.5 were considered, while those with values less than 0.5, indicating that the scaling relationship or equation contained discrepancies, were discarded (Ayalew *et al.* 2015, 2018; Farmer *et al.* 2015).

### 2.3.3. Catchment descriptors and processes affecting parameters

A linear regression model was used to determine the relationships between estimated flood parameter values and other measurable catchment descriptors and processes that include rainfall location, soil moisture, rainfall accumulation, river levels, and peak discharge (Gupta *et al.* 1996; Mandapaka *et al.* 2009a; Ayalew *et al.* 2015, 2018).

Pearson's correlation analysis was conducted using IBM SPSS Statistics for Windows, Version 25, to assess both the strength and the direction of established relationships (IBM Corp. 2017). This study employed the Pearson correlation coefficient (CC) ( $r$ ) to measure the relationship between scaling parameters and physical catchment characteristics. The resulting values were evaluated on a scale ranging from  $-1$  to  $1$ . When  $r$  exceeds  $0.5$  or falls below  $-0.5$ , it is inferred that the data points are in close proximity to the line of best fit.

Furthermore, linear and multiple regression analyses were conducted to investigate the dependency of the flood scaling parameters on multiple variables from the catchment's properties and processes. A forward stepwise multiple regression was employed to identify the strongest predictors of the flood scaling parameter values from the measurable catchment

descriptors and processes identified (Formetta *et al.* 2021). The regression equations were further assessed based on the goodness-of-fit measure shown by the  $R^2$ , the normal distribution of errors, the  $t$ -statistic, collinearity statistics, and the Shapiro–Wilk residual tests (Ayalew *et al.* 2015, 2018).

The Shapiro–Wilk test was used to check the normality of the data and produces a  $W$  value; low values indicate that the sample was not normally distributed. The null hypothesis of normally distributed data was rejected if the alpha ( $\alpha$ ) value was 0.05 and the  $p$ -value was less than 0.05. The  $p$ -values indicate the probability that any observed difference between datasets was attributable to chance. Checks for multicollinearity were done to check if it occurs when independent variables in a model are correlated. Two variables were regarded as perfectly collinear if their CC was within the tolerance of  $>0.1$  and the variance inflation factor (VIF) is  $<10$ . The VIF detected and quantified collinearity in the regression models developed (Bruin 2006).

The best optimal equation, which should not overfit data and perform well in new contexts, was selected using the XLSTAT software, version 2019.3.2, from a pool of several equations developed. The software employs the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), to select optimal equations for predicting flood parameter values for the catchment (Ding *et al.* 2018). The two model selection methods were selected due to their ability to effectively determine the optimal statistical model for fitting hydrological extremes. Di Baldassarre *et al.* (2009) did a study to see how well the AIC, the BIC, and the Anderson–Darling criterion work in statistical flood model selection. The results of their study revealed that the model selection methods were valuable tools for model selection and for reducing the uncertainty of design flood estimation.

The best regression equations selected were then used in a flood prediction framework to estimate the scaling parameter values and peak discharge across nested catchments in the study area. The model results were validated by comparing them to observed values from different time periods and sub-catchments that were not included in the development of equations.

#### 2.4. Framework for predicting flood parameters and peak discharge magnitudes

A flood prediction framework, consisting of the following steps (Figure 4), was employed to estimate flood parameter values and peak discharge magnitudes across the catchment area (Ayalew *et al.* 2014a). First, a log–log regression equation was used to calculate the scaling  $\alpha$  value from the observed peak discharges from the nested catchments. Second, a log–linear regression equation was applied to estimate the corresponding scaling  $\theta$  value from the natural logarithm of the scaling  $\alpha$  estimated in Step 1. Third, the estimated flood parameters ( $\theta$  and  $\alpha$ ) were then used in the flood scaling (Equation (1)) to determine the expected magnitude value of peak discharge across the La Sierra catchment.

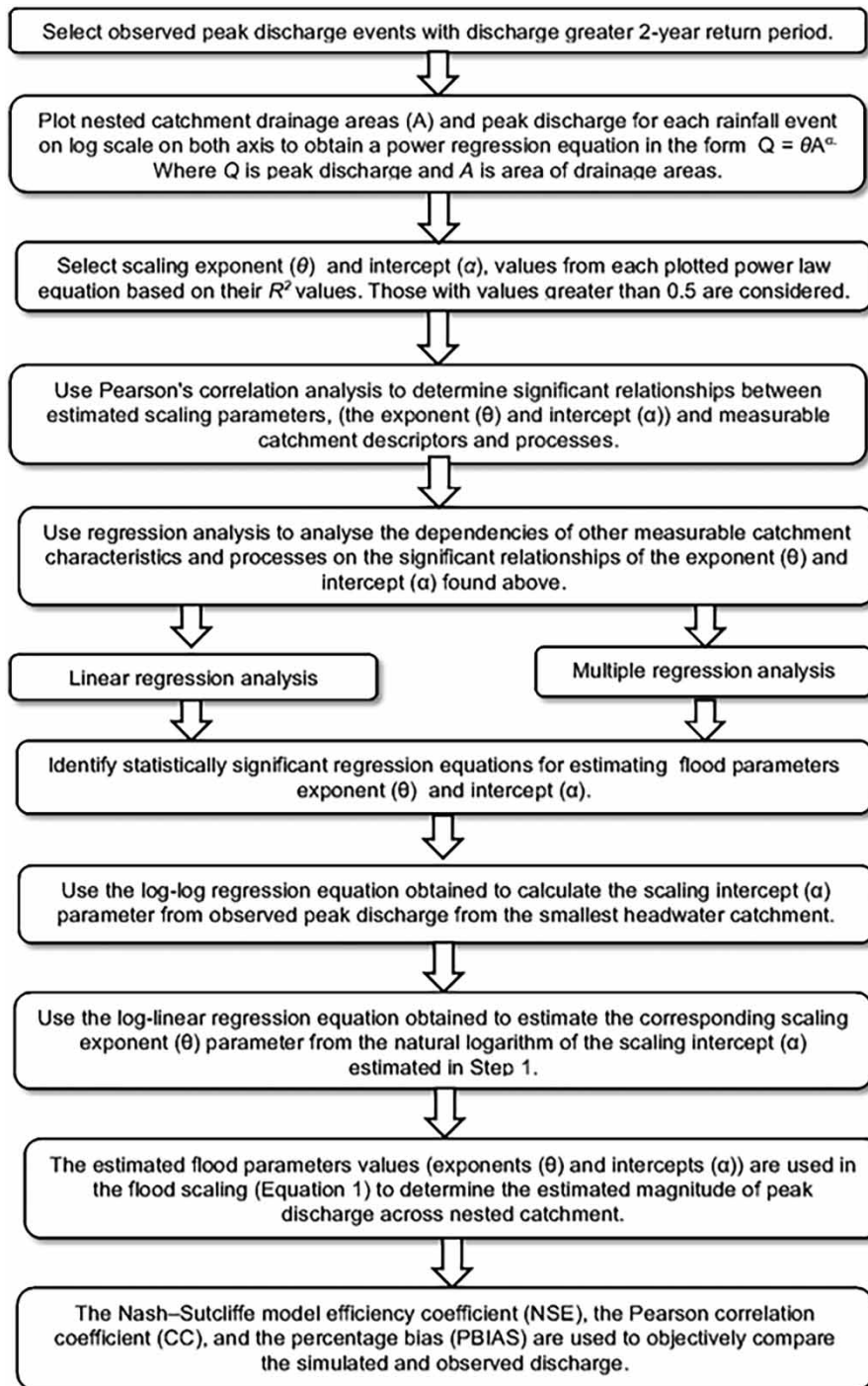
To validate the model results and confirm that the equations developed were sufficiently predictive, the simulated  $\alpha$  and  $\theta$  values were compared to observed values derived from scaling relationships from a different sub-catchment and time (2016–2018) than used in model development. The model was run with the Teapa nested catchment added to the flood scaling equation (Equation (1)), along with the estimated parameters. This was done to test the model's ability to predict in different spatial conditions and see if it can make simulations that are accurate enough. The Nash–Sutcliffe efficiency (NSE), the Pearson CC, and the percentage bias (PBIAS) were calculated to objectively compare the simulated and observed discharge obtained (Moriasi *et al.* 2007). The root mean square error (RMSE) was also used to estimate the errors between observed and simulated peak discharge values found. A flowchart to visually depict the statistical scaling modelling stages is shown in Figure 4.

Uncertainties in the estimated flood scaling parameters and peak discharge values were shown using confidence and prediction intervals. The confidence interval bounds of 5 and 95% were used to represent the range of frequently varying discharge and parameter values, illustrating the bounds of their natural variability, and thus expressing the uncertainty in the flood parameters and peak discharge estimates found (Wadsworth 1990; NIST 2012). The prediction interval was taken as the range likely to include the mean value of the estimated flood parameter and peak discharge values given the specific values of independent variables (Helsel & Hirsch 2002).

### 3. RESULTS AND DISCUSSION

#### 3.1. Investigation of scaling relationships in the La Sierra catchment

The results of flood scaling analysis of peak discharge data based on hydrometric records of peak flow from 59 rainfall events in the La Sierra catchment from 2012 to 2015 demonstrate that peak discharge exhibits a power–law relationship with drainage areas, which is consistent with findings from humid temperate regions. The average  $R^2$  is 0.86, with  $R^2$  values greater than 0.7 for 60% of all 59 rainfall–runoff events under consideration (Table 3).



**Figure 4** | Flowchart to visually depict the modelling process.

## 3.2. Other catchment variables affecting flood scaling parameter estimation

### 3.2.1. ( $\alpha$ ) parameter estimation

Correlation analysis between flood parameter values and other catchment variables revealed that peak discharge observed at the Puyacatengo GS has the strongest positive relationship with scaling  $\alpha$  values with a CC of  $r = 0.86$  and  $p = 0.01$ . The scaling intercept values when regressed on the natural logarithm of peak discharge results in a model with a good overall fit with  $R^2 = 0.74$  (Figure 5).



**Table 3** | Peak scaling relationships from selected individual rainfall events in the La Sierra catchment ( $\alpha$  and  $\theta$ )

Rainfall event No.	Peak discharge date	Scaling parameters			Rainfall event No.	Peak discharge date	Scaling parameters		
		$\alpha$	$\theta$	$R^2$			$\alpha$	$\theta$	$R^2$
1	1 January 2012	6.72	0.46	0.76	31	24 May 2014	29.07	0.31	0.66
2	08 January 2012	3.98	0.51	0.97	32	28 May 2014	5.89	0.81	0.90
3	30 January 2012	12.32	0.34	0.88	33	03 June 2014	5.11	0.61	0.82
4	17 April 2012	37.38	0.14	0.51	34	05 July 2014	3.39	0.49	0.78
5	14 May 2012	30.40	0.22	0.52	35	14 July 2014	5.84	0.50	0.73
6	21 May 2012	80.85	0.20	0.53	36	03 August 2014	5.67	0.70	0.75
7	26 June 2012	10.76	0.36	0.59	37	15 August 2014	5.04	0.45	0.52
8	12 August 2012	50.40	0.22	0.57	38	26 August 2014	5.03	0.56	0.53
9	27 September 2015	17.46	0.35	0.59	39	31 August 2014	4.75	0.58	0.85
10	20 December 2012	10.76	0.36	0.59	40	12 September 2014	5.09	0.54	0.78
11	20 January 2013	33.01	0.24	0.56	41	19 September 2014	4.39	0.66	0.83
12	08 June 2013	5.41	0.84	0.58	42	23 September 2014	5.78	0.73	0.79
13	06 July 2013	5.25	0.59	0.81	43	14 October 2014	16.22	0.35	0.58
14	18 August 2013	28.08	0.24	0.60	44	26 October 2014	10.87	0.39	0.50
15	27 August 2013	11.57	0.38	0.59	45	29 October 2014	51.13	0.47	0.55
16	07 September 2013	10.75	0.44	0.53	46	08 November 2014	28.60	0.24	0.57
17	14 September 2013	48.45	0.17	0.59	47	13 November 2014	55.09	0.20	0.64
18	26 September 2013	46.15	0.18	0.51	48	27 November 2014	15.09	0.37	0.60
19	14 October 2013	12.14	0.41	0.71	49	13 June 2015	10.80	0.40	0.85
20	28 October 2013	1.72	0.73	0.79	50	29 August 2015	7.73	0.45	0.92
21	16 November 2013	12.32	0.44	0.87	51	15 September 2015	5.39	0.81	0.82
22	27 November 2013	1.15	0.77	0.86	52	23 September 2015	9.33	0.60	0.82
23	13 December 2013	0.61	0.95	0.80	53	01 October 2015	6.62	0.63	0.94
24	16 December 2013	19.91	0.39	0.89	54	20 October 2015	9.72	0.61	0.66
25	13 March 2014	2.30	0.61	0.97	55	26 October 2015	17.52	0.45	0.91
26	08 April 2014	10.12	0.88	0.57	56	14 November 2015	5.41	0.56	0.70
27	16 April 2014	27.30	0.25	0.50	57	23 November 2015	82.32	0.26	0.70
28	20 April 2014	5.60	0.69	0.88	58	08 December 2015	19.70	0.42	0.75
29	05 May 2014	5.37	0.70	0.78	59	20 December 2015	12.24	0.50	0.92
30	14 May 2014	4.64	0.60	0.76					

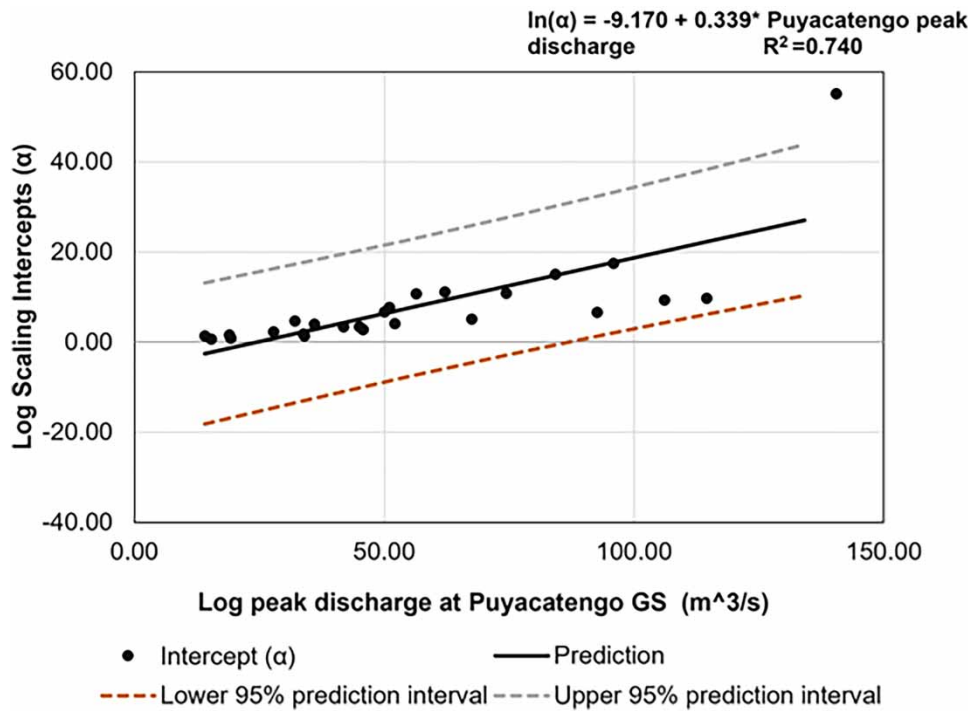
$R^2$  is the coefficient of determination.

The results indicate that the peak discharge observed at the Puyacatengo GS can be used to predict the  $\alpha$  parameter values across the catchment area using a log–log regression equation with an  $\alpha$  value of  $-9.17$ , varying from  $-14.00$  to  $-4.34$ , and an estimated mean slope value of  $0.34$ , varying from  $0.25$  to  $0.42$  at a 95% confidence level (Table 4).

Equation (4) summarises the resulting overall log–log regression equation for estimating  $\alpha$  values.

$$\theta = 5.96 \times 10^{-1} - 7.43 \times 10^{-3} \ln \alpha \quad (4)$$

Student's  $t$ -test on the coefficients of Equation (4) indicates that they are statistically significant at a 95% confidence level. Checks on the Q–Q plot for residuals and the Shapiro–Wilk test for normality yield  $W = 0.96$  and  $p = 0.007$ , indicating that the residuals are normally distributed.



**Figure 5** | Natural logarithm of the scaling  $\alpha$  regressed against the natural logarithm of peak discharge observed at Puyacatengo GS from 59 rainfall–runoff events that occurred between 2012 and 2015.

**Table 4** | The log–log relationship between  $\alpha$  values and Puyacatengo peak discharge

Source	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
$\ln(\alpha)$	-9.170	2.399	-3.822	<0.005	-14.002	-4.338
$\ln(\text{Puyacatengo})$	0.339	0.042	7.995	<0.001	0.254	0.424

Notes: t is the Student’s t-test statistic. The Pr(>t) relates to the probability of observing any value equal to or larger than t.

### 3.2.2. $\theta$ parameter estimation

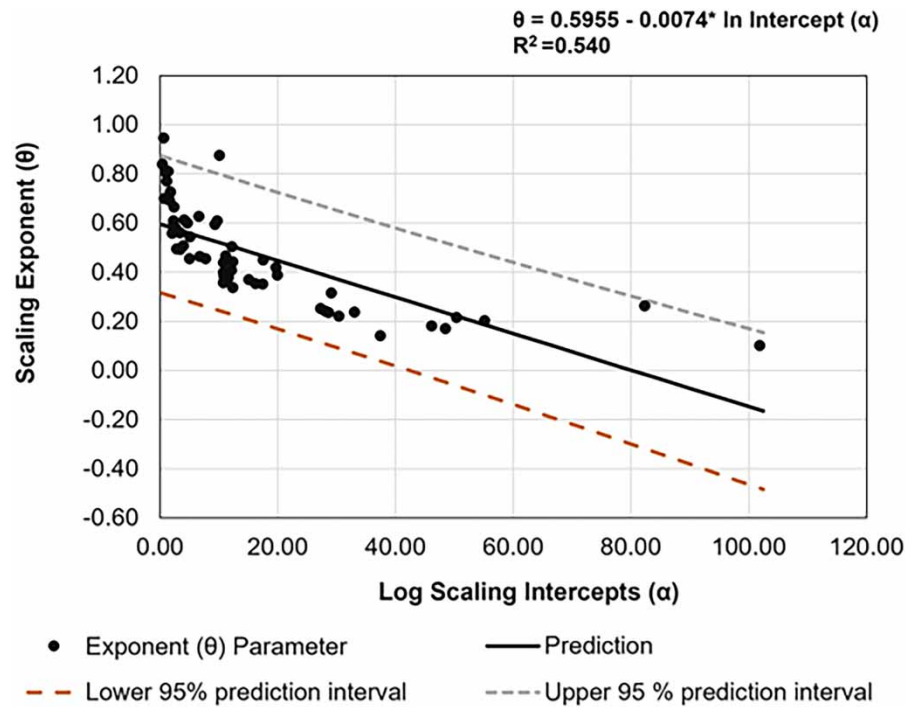
Correlation analysis shows the strongest linear negative relationship between scaling  $\alpha$  and  $\theta$  values, with a CC of  $r = -0.74$  and a p-value of 0.01. The scaling  $\theta$  values when regressed on the natural logarithm of the scaling  $\alpha$  values result in a model with an overall good fit with  $R^2 = 0.54$  (Figure 6). This result shows that the explanatory  $\alpha$  variables in this log-linear relationship can predict 54% of the variability of the dependent scaling  $\theta$  values.

Analysis results (Table 5) show the robustness of the log-linear relationship between the scaling  $\theta$  and  $\alpha$  parameters, with significant coefficients ( $p = 0.05$ ). This shows that  $\alpha$  values can significantly predict  $\theta$  values using a log-linear equation with an  $\alpha$  value of 0.60 ranging from 0.55 to 0.64 and an estimated mean slope value of  $-0.01$  varying from  $-0.01$  to  $-0.006$  at the 95% confidence level.

Equation (5) summarises the overall log-linear regression model.

$$\theta = 0.596 - 7.426E^{-03} \times \ln \alpha. \tag{5}$$

Student’s t-test on the equation coefficients (Equation (5)) reveals that they are statistically significant at a 95% confidence level. Checks on the Q–Q plot for residuals between  $\alpha$  and  $\theta$  values as well as the Shapiro–Wilk test for normality ( $W = 0.89$  and  $p = 0.01$ ) show that the residuals are normally distributed.



**Figure 6** | Scaling  $\theta$  regressed against the natural logarithm of the scaling  $\alpha$  values (adapted from Ayalew *et al.* 2018).

**Table 5** | The log-linear regression analysis between  $\theta$  and  $\alpha$  parameters

Source	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
$\theta$	0.596	0.023	25.947	<0.0012	0.550	0.641
$\alpha$	-0.007	0.001	-8.184	<0.0014	-0.009	-0.006

Notes:  $t$  is the Student's  $t$ -test statistic. The  $\text{Pr}(>t)$  relates to the probability of observing any value equal to or larger than  $t$ .

### 3.2.3. $\alpha$ values multi-regressed with other catchment variables

A multiple regression analysis was also used to estimate  $\alpha$  and  $\theta$  parameter values, this time using more than one exploratory variable. Analysis results show that  $\alpha$  parameter values have the strongest relationship with the observed discharge at the Puyacatengo GS ( $r = -0.77, p = 0.01$ ), followed by discharge at the Teapa GS ( $r = 0.38, p = 0.01$ ) (Table 6). The natural logarithm of  $\alpha$  values was multi-regressed on the natural logarithm of peak discharges observed at Puyacatengo and Teapa GSs. A simple equation with all significant coefficients ( $p = 0.05$ ) was then obtained using a stepwise forward elimination procedure.

**Table 6** | Log-log multiple regression analysis of the relationships between  $\alpha$  parameters, for Puyacatengo GS and Teapa GS peak discharge

Model	Unstandardised coefficients $B^*$	Standardised coefficients		t	Sig.	Correlations			Collinearity statistics	
		Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
ln(constant)	0.667	0.04		17.73	0.004					
ln(Puyacatengo)	-0.003	0.00	-0.81	-8.03	0.003	-0.73	-0.77	-0.77	0.91	1.11
ln(Teapa)	0.0001	0.00	0.27	2.65	0.011	0.02	0.38	0.26	0.91	1.11

Notes:  $B^*$  is the unstandardised beta value representing the slope of the line between the predictor variable and the dependent variable.  $t$  is the Student's  $t$ -statistic. VIF is the factor for quantifying collinearity (see Section 2.3 for explanation).

Equation (6) summarises the resulting log–log multiple regression equation.

$$\ln \text{Intercept} = 0.667 - 0.003 \ln \text{Puyacatengo peak flows} + 0.0001 \ln \text{Teapa flows.} \tag{6}$$

Student’s *t*-test on the equation coefficients reveals that Puyacatengo and Teapa peak discharge values can significantly ( $p = 0.05$ ) predict  $\alpha$  parameter values for the catchment area. Furthermore, checks for multicollinearity in the model show that it is at 1.1, which is within the tolerance of  $>0.1$  and  $\text{VIF} < 10$ .

### 3.3. Developing flood scaling equations

Using the methods described above, six scaling equations were developed for estimating flood parameter values for the entire La Sierra the catchment. Four of the six equations were for the  $\theta$  parameter estimation, with an average good model fit ( $R^2 = 0.57$ ). The remaining two were found to be valid for estimating  $\alpha$  parameter values, both with an overall good model fit and an  $R^2 = 0.62$ .

### 3.4. Selection of the best scaling equations

Equation (1) was found to have the lowest AIC (−232.30) and BIC (−228.14) values and was thus chosen as the best log-linear regression model for estimating  $\theta$  parameter values for the La Sierra catchment (Table 7). Equation (5) was found to be the best for estimating  $\alpha$  parameter values, with the lowest AIC (−191.15) and BIC (−185.66) values.

### 3.5. Flood prediction framework

The estimated values for the intercept and exponent parameters were incorporated in the scaling equation to estimate discharge values in the study area. By comparing the values that were predicted and those that were observed, it was possible to see how well the equations were able to predict flood parameters and peak discharge across the La Sierra catchment.

### 3.6. $\alpha$ parameter comparison

The simulated intercept values mirrored changes in the observed values, with an overall equation fit with  $R^2 = 0.51$  (Figure 7). However, errors were generated in estimating the values, shown by a high RMSE of  $11.36 \text{ m}^3/\text{s}$  and a PBIAS of  $-10.39\%$ , which indicates an underestimation in the predictions.

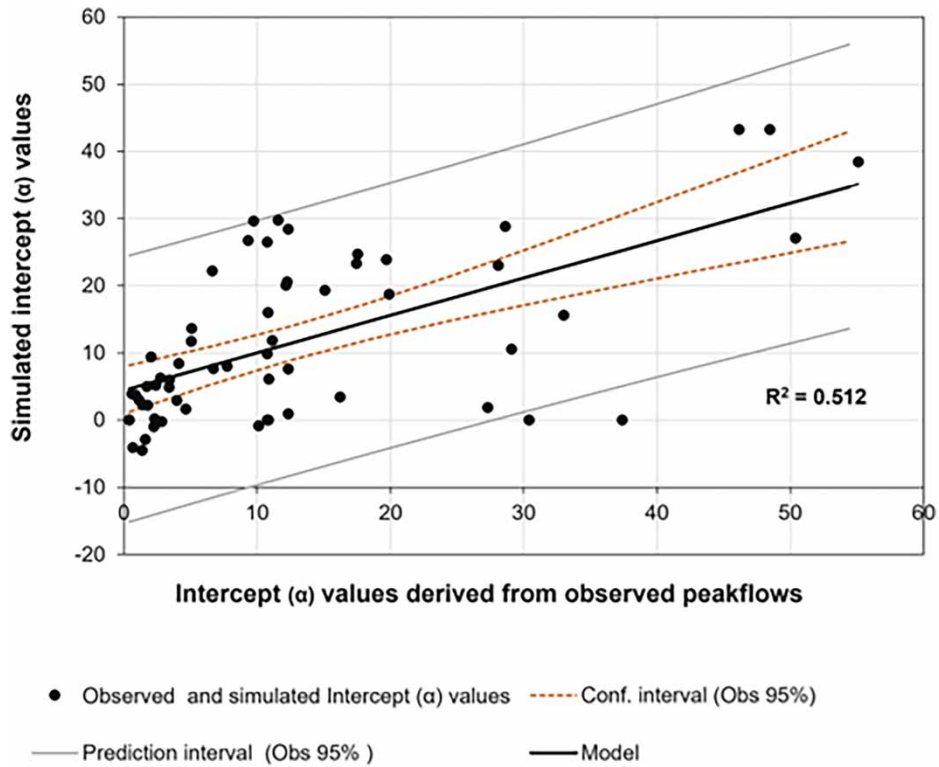
#### 3.6.1. $\theta$ parameter comparison results

The log-linear equation can estimate the value of the  $\theta$  with an overall equation fit with  $R^2 = 0.56$  (Figure 8). However, there were some errors in the estimates, as evidenced by the high RMSE of 12.57 and a PBIAS of 2.08%, indicating overestimation in predictions.

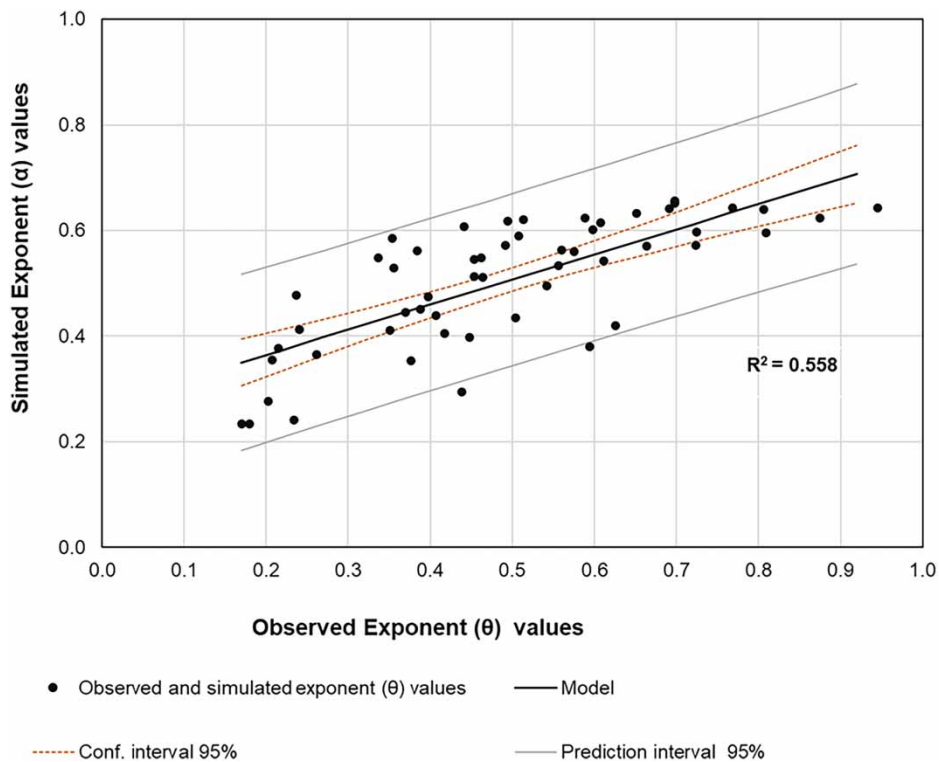
**Table 7** | Scaling equation ranking according to the AIC and BIC estimators

Rank No	Overall equations	Equation selection metrics			
		MSE	Adj. $R^2$	AIC	BIC
1	$\theta = 5.96 \times 10^{-1} - 7.43 \times 10^{-3} \ln \alpha$	0.02	0.53	−232.30	−228.14
2	$\theta = 6.99 \times 10^{-1} - 3.01 \times 10^{-3} \ln \text{Puyacatengo peak discharge}$	0.02	0.49	−192.54	−188.80
3	$\theta = 3.17 \times 10^{-1} - 3.21 \times 10^{-4} \ln 2 \text{ h Oxolotan peak discharge}$	0.01	0.56	−34.16	−34.00
4	$\ln \alpha = 9.17 \times 10^0 + 3.40 \times 10^{-1} \ln \text{Puyacatengo peak discharge}$	78.19	0.74	211.19	214.94
5	$\ln \alpha = 6.67 \times 10^{-1} - 5.92 \times 10^{-4} \ln (\text{Puyacatengo peak discharge} + 3.46\text{E-}03 \times \ln \text{Teapa peak discharge})$	0.015	0.58	−191.15	−185.66
6	$\ln \theta = 4.93 \times 10^{-1} - 3.33 \times 10^{-3} \ln (\text{Puyacatengo peak discharge} + 4.99\text{E-}04 \times \text{Gaviota peak discharge})$	0.006	0.83	−232.20	−226.71

Note: Adj.  $R^2$  is a corrected goodness-of-fit measure for assessing models.



**Figure 7** | Comparison between observed and simulated scaling  $\alpha$  values from rainfall-runoff events (2012–2015) across the La Sierra catchment.



**Figure 8** | Comparison of estimated  $\theta$  and observed values from rainfall-runoff events in the La Sierra catchment from 2012 to 2015.

### 3.6.2. Peak discharge comparison results

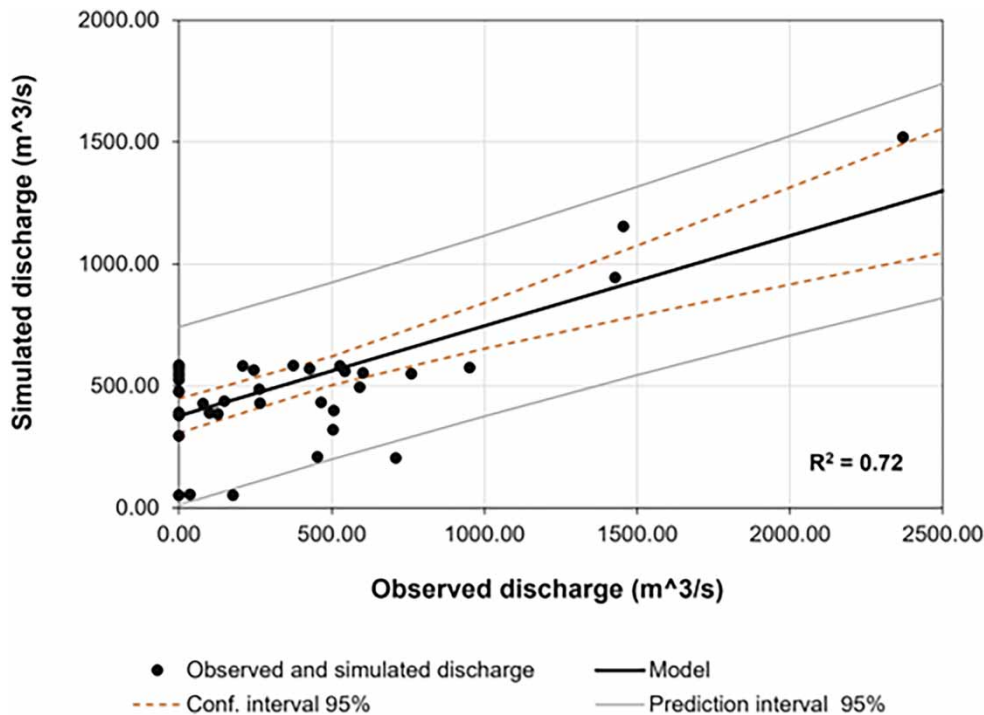
A comparison of the simulated and observed peak discharge from 2016 to 2018 at the Teapa GS shows the strongest good model fit, shown by  $R^2 = 0.72$  (Figure 9), a good CC ( $r = 0.79$ ), and a good NSE (0.60), but again with a large RMSE of  $9.04 \text{ m}^3/\text{s}$ .

### 3.7. Summary of statistical significance

The  $\alpha$  and  $\theta$  parameter equations developed were found to be statistically significant in their relationship with catchment descriptors and processes in the La Sierra catchment and validated by comparing them to historical data. A series of statistical tests were conducted to validate the robustness of the regression equations obtained. The test results show that the peak flood observed at the smallest sub-catchments in the basin significantly predicted the intercept ( $p < 0.05$ ). The residuals were also confirmed to be normally distributed through the inspection of their Q-Q plot and the Shapiro-Wilk test ( $W = 0.96$  and  $P = 0.12$ ) for the  $\alpha$  and between the  $\alpha$  and  $\theta$  values ( $W = 0.89$  and  $p = 0.01$ ), showing that the residuals were normally distributed. Similarly, the study checked for multicollinearity in the model data, whether it was within the tolerance of  $>0.1$  and  $\text{VIF} < 10$  and whether the data were normally distributed. Based on these findings, it is possible to conclude that the regression models generated were reliable though with uncertainties that were quantified and reflected in the results.

## 4. DISCUSSION

The results reveal that peak discharge in the La Sierra catchment exhibits power-law relationships. The study also demonstrates how the estimated scaling parameters are linked to catchment descriptors and processes, allowing statistical equations to be developed and utilised to estimate flood parameter values and the magnitude of peak discharge across the studied catchment area. Similarly, recent studies on scaling laws have demonstrated the potential of estimating flood parameters by analysing the physical characteristics of drainage areas (Menabde & Sivapalan 2001; Furey & Gupta 2005, Gupta *et al.* 2007; Mantilla *et al.* 2006; Mandapaka *et al.* 2009a; Ayalew *et al.* 2014a, 2014b, 2015; Farmer *et al.* 2015; Medhi & Tripathi 2015; Furey *et al.*, 2016; Lee & Huang 2016; Wilkinson & Bathurst 2018; Perez Mesa 2019; Yang *et al.* 2020).



**Figure 9** | Comparison results of simulated and observed discharge from Teapa GS from 2016 to 2018.

A noteworthy result from present study is that peak discharge exhibits a power-law relationship with drainage areas, similar to studies from humid temperate regions (Table 2). The results of the scaling relationship analysis in the La Sierra sub-catchment confirmed that the flood parameters,  $\theta$  and  $\alpha$  in the scaling equations, vary with rainfall events, which is consistent with findings in the United States and the United Kingdom (Gupta *et al.* 1996, 2007; Ogden & Dawdy 2003; Gupta 2004; Furey & Gupta 2005; Mantilla 2007; Mantilla *et al.* 2011; Ayalew *et al.* 2014a, 2014b, 2015).

The study reveals the existence of a log-linear relationship between  $\alpha$  and  $\theta$  values (Equation (5)) and a log-log power-law relationship between  $\alpha$  and the peak discharge observed from the smallest headwater catchments (Equation (4)). The results show that the log-linear and log-log relationships can be used to estimate the values of flood scaling parameters across the nested catchments, even though there are some quantifiable errors in the predictions. This finding is consistent with the research conducted by Ayalew *et al.* (2015, 2018) in Iowa River basin, which revealed a log-linear relationship between the intercept and exponent as well as a log-log relationship between the intercept and the peak flood observed in the smallest gauged sub-catchments within the nested catchment. These relationships, when combined with the scaling law equation for peak floods, can be used to reasonably predict peak floods at any location within the nested catchments after a rainfall-runoff event (Ayalew *et al.* 2018).

Chen *et al.* (2020) recently revealed that using log-transformation methods in peak discharge power-law analysis produces errors and uncertainties when compared to nonlinear regression methods, which is noted in this study. The study found that the log-log-linear regressions produced higher peak discharge prediction errors than nonlinear regressions, and the logarithmic transformation results in smaller peak and drainage area data points that are more heavily weighted. As a result, Chen *et al.* (2020) recommend using nonlinear regression on the arithmetic scale to estimate the scaling parameters when performing peak discharge scaling analysis, particularly for prediction purposes. This study was conducted before the study by Chen *et al.* (2020), so it did not incorporate their findings into the design of this study. However, there is some evidence in this work that the log transformations show a satisfactory fitting on the log-log scale, reliably control the size-dependent variability present, and ensure a normal distribution of the residuals. The regression equations obtained from using multi-scale nested catchments, each less than 10,000 km<sup>2</sup> in size, were found to be adequately fitted, not overfitting the data, and performing well in new contexts. The errors and uncertainties generated by the log-transformation approach used in this study were quantified to demonstrate the variability in the estimated flood scaling parameters and peak discharge values. Thus, the knowledge of quantified uncertainties improves the flood modelling approach employed and allows for a clear uncertainty analysis aimed at taking into account errors and uncertainties in the modelling process.

The ability of the scaling equations to predict small-scale flood events has been demonstrated with a degree of confidence (Figure 9); however, the equation's ability to predict large-scale flood events is questioned. The results show that the simulated peak discharges have greater variability around their expected values, which is particularly noticeable in the estimation of high-magnitude flood events (Figure 9). These results could support the findings of Chen *et al.* (2020); however, it is believed that the uncertainties in the scaling relationships can be explained in part by the lack of long records of hydrometeorological datasets. This study was limited to a short window, between 2012 and 2015, as it was the only period that had data across all gauging sites. Thus, in their current form, the equations are suitable for estimating low- to medium-peak discharges with high recurrence intervals. Therefore, caution should be exercised when extrapolating the findings of this study over time. Future research is recommended to repeat this work in locations with long hydrometeorological records extending across several years to fully capture the low frequent severe flood events.

The findings of this study are significant because they quantify relationships between drainage size, peak discharge, and other catchment processes that cause flooding from small to larger downstream catchments in a tropical region. The results provide statistical relationships (Equations (4) and (5)) that shed light on the broader catchment flood response and flood generation across catchments. In other words, the findings contribute to establishing how flood scaling relationships are quantifiably linked to larger catchment-scale descriptors and processes to enable the estimation of flooding. The most important finding in this regard is how point-scale processes at the smallest nested catchment, such as observed discharge, are related to catchment-wide flooding, which enables new strategies in flood monitoring and prediction. The result contributes to the formulation of a flood risk strategy that promotes a network of flow GSs in smaller headwater streams for peak discharge monitoring and prediction for larger downstream catchment areas (Gupta 2017). This application is especially important for the La Sierra catchment area, which drains floodwaters from smaller mountainous catchments to frequently flood larger downstream catchments in the Gaviotas district areas of Villahermosa City.

Although this research has established valid flood scaling relationships between small headwater catchments and larger downstream catchment areas, not all flooding factors were included in formulating the equations developed. Several other physical catchment properties and processes that directly or indirectly control flooding were not considered due to the lack of relevant data. For example, this study did not consider channel geometry and roughness, key factors that strongly influence river flooding. Although some additional flood-generating variables were not considered, the results obtained were statistically significant and were validated through comparison with historical data.

Thus, this study supports the potential use of scaling equations to simulate discharge without using hydrological models, a method which is essential and relevant under current conditions of climate change. Model calibration has always been a major challenge in the hydrological community (Zheng *et al.* 2021). In changing climatic conditions, the accuracy of historical data is unreliable, and parameter re-calibration using historical data does not reflect future conditions. The scaling equations provide a flood estimation technique that does not require model calibration against historical observations to forecast peak discharge in the future (Gupta 2017). As a result, this study has important implications for the long-term use of statistical equations for flood monitoring and prediction in the face of climate change. It is also suggested that more research should be done to find out what roles climate variability and change play in scaling peak floods.

## 5. CONCLUSIONS

The significance of this study lies in extending the scaling theory of floods to tropical regions and investigating catchment descriptors and processes that influence the flood scaling parameters and the estimation of the same parameter values and peak discharge magnitudes. This study has shown that peak discharge in the La Sierra catchment area exhibits power-law relationships with drainage areas similar to findings from humid temperate regions. The findings substantiate the log-linear relationship identified by Ayalew *et al.* (2015) between the  $\theta$  and the  $\alpha$  of the power-law relationships, not often used in other studies, as well as the significance of the log-log relationship between the  $\alpha$  and peak discharge observed from the smallest nested catchment. In this regard, this study contributes to the ongoing application of the scaling theory of floods and to efforts in predicting floods in data-scarce regions. However, the question of whether the relationship between the intercept and exponent is universal remains unanswered. It is hoped that this question may spark further research.

Nonetheless, the current study has provided a framework for estimating flood parameter values and peak discharge magnitudes that enables flood prediction in data-scarce tropical catchments. The results of the study have led to the identification of some influential factors and provide the foundation for operational flood prediction tools. The study has demonstrated the potential of scaling relationships to estimate the magnitude of peak floods by leveraging the intrinsic relationship between peak discharge and geophysical characteristics and processes within nested catchment areas. The relationships obtained offer valuable insights into the wider catchment flood response and the generation of floods linked to headwater flood peaks (Gupta 2017). One potential implication of this finding is the potential for developing a flood risk strategy that advocates for the setting up of a network of flow GSs in smaller headwater streams that serve the purpose of monitoring peak discharge and estimating flood levels in larger downstream catchment areas.

The self-similarity of river networks and the flood scaling parameters are not expected to change as the climate changes. The scaling laws and statistical models developed will remain unchanged for some time (Gupta 2017). Therefore, there is potential for the long-term use of statistical models for flood monitoring and prediction. This study provides a new research direction to make flood predictions at multiple spatial and temporal scales under a changing global hydroclimate. However, there is a need to investigate the role of climate variability and change using the scaling framework for floods discussed.

Thus, the persistent challenge that remains is to 'bridge the scaling gap' by extending the scaling theory of floods (scaling relationships) from the smallest rivers on the mainland to large regional and continental rivers in different climate regions. There is still a shortage of scaling relationship studies in multi-scale catchments across climatic regions. In several catchments, the process by which floods are generated from small headwater catchments to larger downstream catchments is still not quantified. Therefore, there is a need for further studies that apply the scaling theory of floods and establish patterns of scaling relationships across regional and continental river catchments.



## ACKNOWLEDGEMENTS

This work was supported by funding from the UK National Environment Research Council (NE/M009009/1). The authors would like to acknowledge the guidance and technical advice of Tim Brewer and thank CONAGUA for providing the data for the analyses. The authors also thank two anonymous reviewers who provided constructive feedback on an earlier draft of the manuscript.

## DATA AVAILABILITY STATEMENT

The data used in this study have been uploaded to Cranfield University's open access data repository (CORD - DOI: 10.17862/cranfield.rd.23530956).

## CONFLICT OF INTEREST

The authors declare there is no conflict.

## REFERENCES

- Arreguín-Cortés, F. I., Rubio-Gutiérrez, H., Dominguez-Mora, R. & Luna-Cruz, F. D. 2014 Análisis de las inundaciones en la planicie tabasqueña en el periodo 1995-2010. *Tecnología y Ciencias del Agua* 5 (3), 5–32. Available from: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-24222014000300001](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-24222014000300001).
- Ayalew, T. B., Krajewski, W. F. & Mantilla, R. 2014a Connecting the power-law scaling structure of peak-discharge to spatially variable rainfall and catchment physical properties. *Advances in Water Resources* 71, 32–43. <https://doi.org/10.1016/j.advwatres.2014.05.009>.
- Ayalew, T. B., Krajewski, W. F., Mantilla, R. & Small, S. J. 2014b Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak discharge. *Advances in Water Resources* 64, 9–20. <https://doi.org/10.1016/j.advwatres.2013.11.010>.
- Ayalew, T. B., Krajewski, W. F. & Mantilla, R. 2015 Analysing the effects of excess rainfall properties on the scaling structure of peak discharge: Insights from a mesoscale river basin. *Water Resources Research* 51 (6), 3900–3921. <https://doi.org/10.1002/2014WR016258>.
- Ayalew, T. B., Krajewski, W. F., Mantilla, R. & Zimmerman, D. L. 2018 Can floods in large river basins be predicted from floods observed in small subbasins? *Journal of Flood Risk Management* 11 (3), 331–338. <https://doi.org/10.1111/jfr3.12327>.
- Bruin, J. 2006 newtest: Command to compute new test. UCLA: Statistical Consulting Group. Available from: <https://stats.oarc.ucla.edu/stata/ado/analysis/>.
- Chen, B., Ma, C., Krajewski, W. F., Wang, P. & Ren, F. 2020 Logarithmic transformation, and peak-discharge power-law analysis. *Hydrology Research* 51 (1), 65–76. <https://doi.org/10.2166/nh.2019.108>.
- CLICOM of the SMN. 2013 CICESE web platform. Available from: <http://clicom-mex.cicese.mx>.
- Dawdy, D. R., Griffis, V. W. & Gupta, V. K. 2012 Regional flood frequency analysis: how we got here and where we are going. *Journal of Hydrologic Engineering* 17 (9), 953–959. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000584](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000584).
- Di Baldassarre, G., Laio, F. & Montanari, A. 2009 Design flood estimation using model selection criteria. *Physics and Chemistry of the Earth, Parts A/B/C* 34 (10–12), 606–611. ISSN 1474-7065. <https://doi.org/10.1016/j.pce.2008.10.066>.
- Ding, J., Tarokh, V. & Yang, Y. 2018 Model selection techniques: An overview. *IEEE Signal Processing Magazine* 35 (6), 16–34. Available from: <https://arxiv.org/pdf/1810.09583.pdf>.
- Environmental Systems Research Institute (ESRI) 2012 World Topo Base [basemap]. global coverage from ~1:591M scale down to ~1:577k. World Topographic Map. <https://wtb.maptiles.arcgis.com/arcgis/rest/services>. (Accessed 29 November 2018).
- Farmer, W. H., Over, T. M. & Vogel, R. M. 2015 Multiple regression and inverse moments improve the characterization of the spatial scaling behavior of daily streamflows in the Southeast United States. *Water Resources Research* 51 (3), 1775–1796. <https://doi.org/10.1002/2014WR015924>.
- Formetta, G., Over, T. & Stewart, E. 2021 Assessment of peak flow scaling and its effect on flood quantile estimation in the United Kingdom. *Water Resources Research* 57, e2020WR028076. <https://doi.org/10.1029/2020WR028076>.
- Furey, P. R. & Gupta, V. K. 2005 Effects of excess rainfall on the temporal variability of observed peak discharge power laws. *Advances in Water Resources* 28 (11), 1240–1253. <https://doi.org/10.1016/j.advwatres.2005.03.014>.
- Furey, P. R., Troutman, B. M., Gupta, V. K. & Krajewski, W. F. 2016 Connecting event-based scaling of flood peaks to regional flood frequency relationships. *Journal of Hydrologic Engineering* 21 (10), 04016037.
- Gupta, V. K. 2004 Emergence of statistical scaling in floods on channel networks from complex runoff dynamics. *Chaos, Solitons & Fractals* 19 (2), 357–365. [https://doi.org/10.1016/S0960-0779\(03\)00048-1](https://doi.org/10.1016/S0960-0779(03)00048-1).
- Gupta, V. K. 2017 Scaling theory of floods for developing a physical basis of statistical flood frequency relations. In: *Oxford Research Encyclopaedia of Natural Hazard Science*. Oxford University Press, Oxford, UK. <https://doi.org/10.1093/acrefore/9780199389407.013.301>.
- Gupta, V. K., Castro, S. L. & Over, T. M. 1996 On scaling exponents of spatial peak flows from rainfall and river network geometry. *Journal of Hydrology* 187 (1–2), 81–104. [https://doi.org/10.1016/S0022-1694\(96\)03088-0](https://doi.org/10.1016/S0022-1694(96)03088-0).

- Gupta, V. K., Troutman, B. M. & Dawdy, D. R. 2007 Towards a non-linear geophysical theory of floods in river networks: An overview of 20 years of progress. In: *Non-linear Dynamics in Geosciences*. Springer, New York, USA, pp. 121–151. [https://doi.org/10.1007/978-0-387-34918-3\\_8](https://doi.org/10.1007/978-0-387-34918-3_8).
- Gupta, V. K., Mantilla, R., Troutman, B. M., Dawdy, D. & Krajewski, W. F. 2010 **Generalising a non-linear geophysical flood theory to medium-sized river networks**. *Geophysical Research Letters* **37**, L11402. <https://doi.org/10.1029/2009GL041540>.
- Gupta, V. K., Ayalew, T. B., Mantilla, R. & Krajewski, W. F. 2015 **Classical and generalised Horton laws for peak flows in rainfall-runoff events**. *Chaos* **25**, 075408. <https://doi.org/10.1063/1.4922177>.
- Helsel, D. R. & Hirsch, R. M. 2002 Statistical methods in water resources. In: *Hydrologic Analysis and Interpretation*. US Geological Survey, Reston, United States. <https://doi.org/10.3133/twri04A3>.
- Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. & Cudennec, C. 2013 **A decade of predictions in ungauged basins (PUB): A review**. *Hydrological Sciences Journal* **58** (6), 1198–1255. <https://doi.org/10.1080/02626667.2013.803183>.
- IBM Corp. Released. 2017 *IBM SPSS Statistics for Windows, Version 25.0*. IBM Corp, Armonk, NY. Available from: <https://www.ibm.com/products/spss-statistics>.
- Instituto Nacional de Estadística y Geografía (INEGI) (Mexican Institute of Statistics and Geography), ‘Topography’ (‘Hydrography,’) 2018 Available from: <http://en.www.inegi.org.mx/temas/imagenes/espaciomapa/> (accessed January 2018).
- Lee, K. T. & Huang, J. K. **Influence of storm magnitude and watershed size on runoff nonlinearity**. *Journal of Earth System Science* **125**, 777–794. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.00014](https://doi.org/10.1061/(ASCE)HE.1943-5584.00014).
- Lehner, B. & Grill, G. 2013 **Global river hydrography and network routing: Baseline data and new approaches to study the world’s large river systems**. *Hydrological Processes* **27** (15), 2171–2186. <https://doi.org/10.1002/hyp.9740>.
- López-Caloca, A. A., Tapia-Silva, F. O. & Rivera, G. 2017 Sentinel-1 satellite data as a tool for monitoring inundation areas near urban areas in the Mexican tropical wet. In: *Water Challenges of an Urbanising World*. IntechOpen. <https://www.intechopen.com/books/water-challenges-of-an-urbanizing-world/sentinel-1-satellite-data-as-a-tool-for-monitoring-inundation-areas-near-urban-areas-in-the-mexican> (accessed 6 September 2019).
- Mandapaka, P. V., Krajewski, W. F., Ciach, G. J., Villarini, G. & Smith, J. A. 2009a **Estimation of radar-rainfall error spatial correlation**. *Advances in Water Resources* **32** (7), 1020–1030. <https://doi.org/10.1016/j.advwatres.2008.08.014>.
- Mandapaka, P. V., Krajewski, W. F., Mantilla, R. & Gupta, V. K. 2009b **Dissecting the effect of rainfall variability on the statistical structure of peak flows**. *Advances in Water Resources* **32** (10), 1508–1525. <https://doi.org/10.1016/j.advwatres.2009.07.005>.
- Mantilla, R., Gupta, V. K. & Mesa, O. J. 2006 Role of coupled flow dynamics and real network structures on Hortonian scaling of peak flows. *Journal of Hydrology* **322** (1–4), 155–167.
- Mantilla, R. 2007 *Physical Basis of Statistical Scaling in Peak Flows and Streamflow Hydrographs for Topologic and Spatially Embedded Random Self-Similar Channel Networks*. Doctoral dissertation, the University of Colorado at Boulder. Available from: [https://www.iuhr.uiowa.edu/rmantilla/files/2015/03/Thesis Mantilla.pdf](https://www.iuhr.uiowa.edu/rmantilla/files/2015/03/Thesis%20Mantilla.pdf) (accessed 6 May 2019).
- Mantilla, R., Gupta, V. K. & Troutman, B. M. 2011 **Scaling of peak flows with constant flow velocity in random self-similar networks**. *Non-linear Processes in Geophysics* **18** (4). <https://doi.org/10.5194/npg-18-489-2011>.
- Medhi, H. & Tripathi, S. 2015 **On identifying relationships between the flood scaling exponent and basin attributes**. *Chaos* **25** (7), 075405. Erratum in *Chaos*. 2015; **25** (7), 079901. <https://doi.org/10.1063/1.4916378>.
- Menabde, M. & Sivapalan, M. 2001 **Linking space-time variability of river runoff and rainfall fields: A dynamic approach**. *Advances in Water Resources* **24** (9–10), 1001–1014. [https://doi.org/10.1016/S0309-1708\(01\)00038-0](https://doi.org/10.1016/S0309-1708(01)00038-0).
- Mercado, V. D., Perez, G. C., Solomatine, D. & Van Lanen, H. A. J. 2016 **Spatiotemporal analysis of hydrological drought at catchment scale using a spatially distributed hydrological model**. *Procedia Engineering* **154**, 738–744. <https://doi.org/10.1016/j.proeng.2016.07.577>.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 **Model evaluation guidelines for systematic quantification of accuracy in watershed simulations**. *Transactions of the ASABE* **50** (3), 885–900. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- NRCS (National Research Conservation Service). 1997 Ponds—Planning, design, construction. Agriculture Handbook no. 590. US Natural Resources Conservation Service, Washington, DC. Available from: <https://nrcspad.sc.egov.usda.gov/distributioncenter/product.aspx?ProductID=115> (accessed 15 September 2023).
- NIST Handbook 44; Specifications, Tolerances, and Other Technical Requirements for Weighing and Measuring Devices. 2012 *National Institute of Standards and Technology OR National Bureau of Standards*. U.S. Department of Commerce, Gaithersburg. Available from: <https://www.nist.gov/> (accessed 26 October 2018).
- Ogden, F. L. & Dawdy, D. R. 2003 **Peak discharge scaling in small Hortonian watershed**. *Journal of Hydrologic Engineering* **8** (2), 64–73. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2003\)8:2\(64\)](https://doi.org/10.1061/(ASCE)1084-0699(2003)8:2(64)).
- Paschalis, A., Faticchi, S., Molnar, P., Rimkus, S. & Burlando, P. 2014 **On the effects of small-scale space-time variability of rainfall on basin flood response**. *Journal of Hydrology* **514**, 313–327. <https://doi.org/10.1016/j.jhydrol.2014.04.014>.
- Perez Mesa, G. J. 2019 *Explaining the Physics Behind Regional Peak Flow Equations Using the Scaling Theory of Floods and River Network Descriptors*. PhD (Doctor of Philosophy) thesis, University of Iowa. <https://doi.org/10.17077/etd.dabq-5b5d>.

- Rabus, B., Eineder, M., Roth, A. & Bamler, R. 2003 The shuttle radar topography mission – a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing* **57** (4), 241–262. [https://doi.org/10.1016/S0924-2716\(02\)00124-7](https://doi.org/10.1016/S0924-2716(02)00124-7).
- Robinson, J. S. & Sivapalan, M. 1997 An investigation into the physical causes of scaling and heterogeneity of regional flood frequency. *Water Resources Res* **33** (5), 1045–1059. <https://doi.org/10.1029/97WR00044>.
- Sivapalan, M. 2003 Prediction in ungauged basins: A grand challenge for theoretical hydrology. *Hydrological Processes* **17** (15), 3163–3170. <https://doi.org/10.1002/hyp.5155>.
- Wadsworth, H. M. 1990 *Handbook of Statistical Methods for Engineers, and Scientists*. McGraw-Hill, Inc., New York, USA.
- Wilkinson, M. E. & Bathurst, J. C. 2018 A multi-scale nested experiment for understanding flood wave generation across four orders of magnitude of the catchment area. *Hydrology Research* **49** (3), 597–615. <https://doi.org/10.2166/nh.2017.070>.
- Yang, W., Yang, H. & Yang, D. 2020 Classifying floods by quantifying driver contributions in the Eastern Monsoon Region of China. *Journal of Hydrology* **585**, 124767. <https://doi.org/10.2166/nh.2017.070>.
- Zheng, Y., Li, J., Zhang, T., Rong, Y. & Feng, P. 2021 Exploring the application of flood scaling property in hydrological model calibration. *Journal of Hydrometeorology*. <https://doi.org/10.1175/JHM-D-21-0123.1>.

First received 23 June 2023; accepted in revised form 5 January 2024. Available online 23 January 2024

# Analysis of scaling relationships for flood parameters and peak discharge estimation in a tropical region

Mazivanhanga, Charles

2024-01-23

Attribution 4.0 International

---

Mazivanhanga C, Grabowski R, Perez Sanchez E, Carballo Cruz VR. (2024) Analysis of scaling relationships for flood parameters and peak discharge estimation in a tropical region. *Hydrology Research*, Volume 55, Issue 2, February 2024, pp. 161-179

<https://doi.org/10.2166/nh.2024.111>

*Downloaded from CERES Research Repository, Cranfield University*