



# Statistical analysis of NRCS curve number (NRCS-CN) in arid basins based on historical data

Mohammed M. Farran<sup>1</sup> · Amro M. Elfeki<sup>1</sup>

Received: 6 June 2019 / Accepted: 18 November 2019 / Published online: 3 January 2020  
© Saudi Society for Geosciences 2020

## Abstract

CN method is a commonly used technique for estimation of direct runoff for flood mitigation studies. In Saudi Arabian (SA) arid basins, there is no analysis of NRCS-CN obtained from measured rainfall-runoff events. The research objectives are to find out the actual range of CN values in arid basins, the statistical distribution of CN, the confidence intervals of CNs, and the relation between the CN and initial abstraction factor,  $\lambda$ . Five basins with 19 sub-basins located in the southwest of Saudi Arabia were considered, and 161 rainfall-runoff events were analyzed during the period 1984–1987. The least squares method was used to obtain the optimum range of values for  $\lambda$  and CN based on three estimation methods. The rainfall and runoff exhibited log-normal distribution. The analysis showed that CN varied between 45 and 85 at  $\lambda = 0.2$ . The low CN values account for the transmission losses which is a typical phenomenon in arid regions. The initial abstraction ratio  $\lambda = 0.01$  is found to be more representative to arid basins rather than  $\lambda = 0.2$ . The Beta distribution is the best to fit CN at both  $\lambda = 0.2$  and  $0.01$ . The confidence intervals are estimated and tabulated at different significant levels for flood risk assessment studies in arid basins.

**Keywords** NRCS-CN method · Curve number · Least square method · Arid regions · Initial abstraction · Flood studies

## Introduction

The method of curve number (CN) for estimation of direct runoff in a catchment from storm rainfall is becoming a fundamental technique in flood impact analyses, design of hydraulic structures, and hydrological engineering studies (Ponce 1996). The technique was developed by the Soil Conservation Service (SCS) in 1954 (Rallison 1980), and was described in NRCS National Engineering Handbook Section 4 (NEH-4) (USDA-SCS 1985). The Soil Conservation Service (SCS) name has been changed in 1994 to Natural Resources Conservation Service (NRCS). The NRCS-CN method is based on the principle of the mass conservation law and two essential hypotheses given

by Mishra and Singh (2002). The first hypothesis says that the ratio of direct runoff,  $Q$ , to potential maximum runoff ( $P-I_a$ ), where  $I_a$  is the initial abstraction, is identical to the ratio of infiltration,  $F$ , to potential maximum retention,  $S$ . The second hypothesis says that  $I_a$  is proportional to the potential maximum retention,  $S$  (i.e.,  $I_a = \lambda S$ ). The NRCS-CN method equations which are introduced by the USDA-SCS (1985) and USDA-NRCS (2004) are given by:

$$Q = \begin{cases} \frac{(P-\lambda S)^2}{P + (1-\lambda)S}, & P > \lambda S \\ 0, & P \leq \lambda S \end{cases} \quad (1)$$

Where  $P$  is the rainfall depth,  $Q$  is the direct runoff depth (effective rainfall),  $\lambda$  is the initial abstraction ratio, and  $S$  is the potential maximum retention of the catchment.

There is no specific approach to estimate NRCS-CN from rainfall-runoff events. Through the review of literature, it has been shown that many formulas are presented, e.g., USDA-SCS (1972); Rallison and Cronshey (1979); Hjelmfelt (1980); Rallison and Miller (1982); Hawkins et al. (1985); Zevenbergen (1985); Hawkins (1993). These formalisms show different ways in handling measured precipitation ( $P$ ) and runoff ( $Q$ ) data for 24-h rainfall duration. However, in arid

Responsible Editor: Broder J. Merkel

✉ Mohammed M. Farran  
mohamadalfaran@yahoo.com

<sup>1</sup> Department of Hydrology and Water Resources Management, Faculty of Meteorology, Environment and Arid Land Agriculture, King Abdulaziz University, Jeddah, Saudi Arabia

regions, most storms have short durations which are less than 24 h with intense rainfall. These storms are considered in the analysis.

Several studies have been conducted to describe the variability of the curve number (CN). Most of these studies, if not all, are performed in climate conditions that are not arid or semi-arid like Saudi Arabia. Moreover, the catchments that have been considered in most of these studies are agriculture catchments; therefore, the tables of the NRCS-CN method are mainly estimated based on these types of catchments, e.g., Rallison (1980); Rallison and Miller (1982); McCuen (1982); Hjelmfelt (1991); Rawls and Maidment (1993); Pilgrim and Cordery (1993); Hawkins (1993); Banasik et al. (1994); Ponce (1996); Schneider and McCuen (2005); Soulis and Dercas (2007); Soulis and Valiantzas (2012, 2013); and Banasik et al. (2014).

Many studies have dealt with NRCS-CN method to find out a theory behind the method and applications in different regions in the world under various climate conditions. Rallison and Cronshey (1979) presented a document where data from daily precipitation and runoff for yearly flood events were used in the original Soil Conservation Service handbook. A validation technique of the CN procedure was presented in the Soil Conservation Service handbook USDA-SCS (1972) which compared the frequency distributions of the synthesized and measured runoff volumes. Hjelmfelt (1980) estimated CNs from measured data of  $P$  and  $Q$  pairs by the use of frequency distributions, a CN transformation model, and by fitting techniques. Bonta (1997) used derived frequency distributions as a technique to determine curve numbers from field data, treating rainfall and runoff measurements as discrete distributions. He concluded that the derived distribution method had a potential for determining curve numbers when rainfall and runoff data are limited. According to McCuen (2002) the CN was treated based on the theory of random variables. A general technique is presented to estimate CN for making confidence intervals. Since CN is a value between 0 and 100, it was found that its distribution is negatively skewed. However, common probability distribution are positively skewed; therefore, the use of 100-CN shows positively skewed distributions and consequently is considered in the study. The amount "100-CN" was fitted by the gamma distribution, and therefore, CN confidence intervals were estimated for the range of CN between 65 to 95 of his data.

NRCS-CN adopted the standard value of the initial abstraction ratio,  $\lambda$  (the ratio between initial abstraction,  $I_a$ , and the maximum retention potential,  $S$ ), equals 0.2. The validity and applicability of  $\lambda = 0.2$  has been questioned by many researchers like Ponce (1996); Hawkins et al. (2001); Jiang (2001); and Baltas et al. (2007). Woodward et al. (2003) stated that the effect of using  $\lambda$  equal to 0.05 in place of 0.20 leads either to lower CNs or lower rainfall depths. However, the

values documented in the literature were varying in the range of (0.0 up to 0.3) in a number of studies that include different geographic locations in the US and many countries e.g. Cazier and Hawkins (1984); Ramasastri and Seth (1985); Bosznay (1989); Woodward et al. (2003); Mishra and Singh (2004); and Mishra and Singh (2013). Yuan et al. (2014) stated that the estimation of direct runoff from rainfall using curve number (CN) for semiarid catchments can be inaccurate. The inaccuracy was due to the customary NRCS value of  $\lambda$  equals 0.2. They had shown that runoff estimation is sensitive to the initial abstraction ratio, in case of relatively low rainfall depth and for basins covered by coarse soil conditions that are dominant in many semiarid basins in the world. For Walnut Gulch catchments, optimal values of  $\lambda$  ranging from 0.01 to 0.53 improved runoff estimation. It was concluded that the effect of  $\lambda$  on runoff estimation increases with decreasing the value of CN.

From the aforementioned review of literature on the topic, and to the best of the authors' knowledge, there is no extensive research work on CN from measurements of rainfall and runoff on arid basins especially in Saudi Arabia except the work by Alagha et al. (2016). They studied the relationship between CN and morphometric parameters of some sub-basins in Saudi Arabia and deduced some empirical equations to estimate CN from the wadi morphology. However, in their study, there was no detailed statistical analysis on the CN values. They considered only a single catchment which made the study not sufficient for a reliable statistical analysis.

The objective of this research is to perform a detailed study on NRCS-CN in arid and semi-arid basins that cover the following items: (1) determining the range of values of the curve number CN and  $\lambda$  under arid conditions in the study region from field measurements of rainfall and runoff events, (2) derivation of the best probability distribution of precipitation,  $P$ , runoff,  $Q$ , and curve number CN values, and (3) estimation of the confidence intervals of CN values for arid and semi-arid basins. The expected results would improve the flood predictions in arid and semi-arid regions especially in Saudi Arabia and consequently would come up with reliable parameters for the design of the flood protection schemes in the region.

## Material and methods

### Study area

The study basins are located in the southwestern part of Saudi Arabia. The region is located in the eastern Asir escarpment with elevations of up to 3000 m which runs towards the west to the Red Sea coast. Saudi Arabian Dames and Moore (1988) developed a detailed study of five selected basins with their

sub-basins (19 sub-basins) which include: Al-Lith, Yiba, Habawnah, Liyyah, and Tabalah basins. Al-Lith, Yiba, and Liyyah basins drain towards the Red Sea. However, Habawnah and Tabalah basins drain from the mountains to the interior, towards the Rub al Khali. Figure 1a shows the geographic location of the study area and the locations of the five basins. Each of these basins is divided into sub-basins with runoff stations at the outlet of the each sub-basin. Figure 1b shows the locations of the runoff basin in each of the five basins. Table 1 summarizes the information for the basins and sub-basins.

**Data collection**

The data has been measured by Saudi Arabian Dames and Moore (1988). They installed an extensive measuring network. In arid region in general and in Saudi Arabia in particular, runoff measurements are often rare. The only limited detailed runoff measurements are found in the period 1984–1987. The company that made such measurements had a contract with the ministry for 4 years. During these 4 years, there were 161 storms recorded in the study area which are considered statistically acceptable. In the current study, the climate change is out of the scope of this study and could be considered in the future if there would be extra measurements. Figure 2 shows a sample of rainfall and runoff events (27 April 1985) on station N-404 in Habawnah basin, in the south-west of Saudi Arabia.

**Estimation of the CN and  $\lambda$**

Since  $P$  and  $Q$  data are available for the study catchments,  $P$  and  $Q$  pairs are used directly to estimate the maximum potential retention,  $S$ , characterizing the catchment as (Chen 1982),

$$S = \frac{P}{\lambda} + \frac{(1-\lambda)Q - \sqrt{(1-\lambda)^2 Q^2 + 4\lambda PQ}}{2\lambda^2} \tag{2}$$

Assuming  $\lambda = 0.2$  (the common value), Eq. (2) becomes the common quadratic formula for  $S$  (Hawkins 1973) as,

$$S = 5 \left[ (P + 2Q) - \sqrt{Q(4Q + 5P)} \right] \tag{3}$$

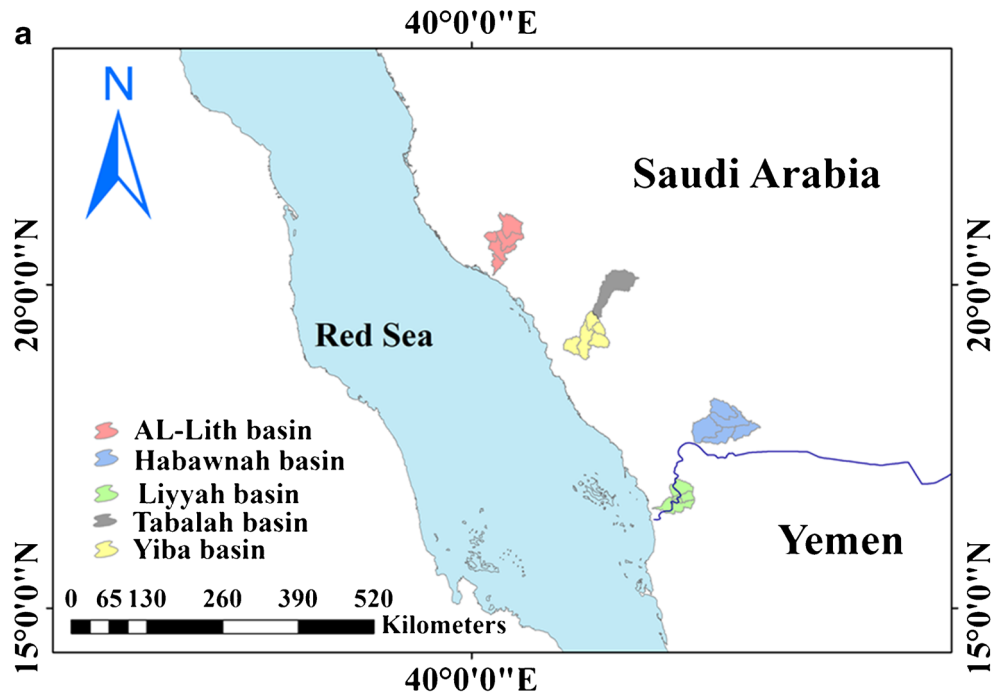
and,

$$CN = \frac{25400}{S + 254} \tag{4}$$

In the current analysis, three different ways of the estimation of CN and  $\lambda$  are suggested as follows:

- Method 1: estimation of CN and  $\lambda$  from all events on the sub-basins of each basin to obtain the best values. The obtained values are called  $\lambda_{best}$  and  $CN_{best}$ .
- Method 2: estimation of CN at  $\lambda = 0.2$  (the common value) from all events on the sub-basins for each basin to obtain the value for CN at  $\lambda = 0.2$ . This is called  $CN_{(\lambda = 0.2)}$ , and.

**Fig. 1** (a) Location of the studied basins in the southwestern part of Saudi Arabia. (b) The studied basins and locations of the runoff stations



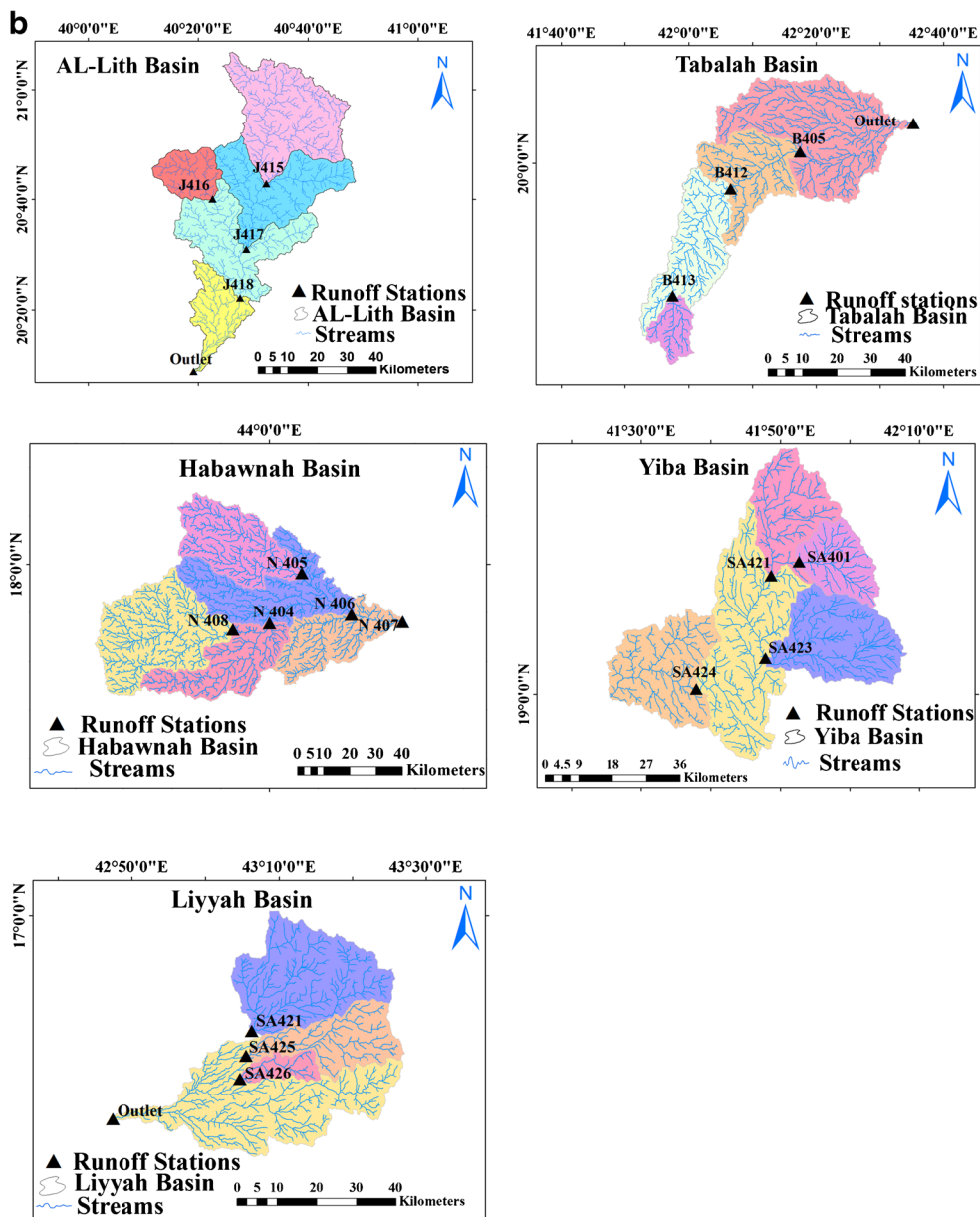


Fig. 1 (continued)

Method 3: estimation of CN and  $\lambda$  from each event on the sub-basins of the basin to obtain the best values per event and averaging over the number of events. The obtained values are called  $\bar{\lambda}_{best}$ , and  $\overline{CN}_{best}$ .

Figure 3 shows a flowchart of the procedure used to estimate CN for every storm at each sub-basin based on inverse modeling technique. The details of the procedure is presented in the following section.

### The Least Squares Method

The LSM for CN estimation was introduced by Hawkins et al. (2002) and Hawkins et al. (2008). The method determinates the parameters  $\lambda$  and CN of the well-known runoff equation. The LSM is a minimization of the sum of squared differences between the observed direct runoff ( $Q$ ) and the estimated runoff ( $\hat{Q}$ ). It gives the best estimation, of the parameters. In the current study, the application of the least square method is

**Table 1** Summary of the basins and the sub-basins data

Label	Basin	Runoff stations	Coordinates		Drainage Area (km <sup>2</sup> )
			E	N	
a	Yiba	SA401	41° 48' 42.35"	19° 17' 2.15"	785
		SA422	41° 52' 41.98"	19° 19' 4.26"	322
		SA423	41° 47' 53.73"	19° 5' 10.44"	597
		SA424	41° 37' 58.89"	19° 0' 45.53"	2305
b	AL-Lith	J415	40° 32' 21.24"	20° 42' 49.59"	949
		J416	40° 22' 37.47"	20° 40' 3.59"	274
		J417	40° 28' 44.14"	20° 30' 56.65"	1657
		J418	40° 27' 37.63"	20° 22' 9.68"	2672
c	Tabalah	B405	42° 17' 28.97"	20° 1' 47.92"	1270
		B412	42° 6' 36.17"	19° 55' 57.45"	730
		B413	41° 57' 28.93"	19° 39' 15.23"	170
d	Liyah	SA421	43° 11' 40.31"	16° 49' 28.29"	744
		SA425	43° 9' 58.66"	16° 41' 34.54"	392
		SA426	43° 8' 6.42"	16° 34' 10.07"	99
e	Habawnah	N404	44° 0' 0.91"	17° 47' 54.03"	2180
		N405	44° 6' 44.15"	17° 58' 13.38"	940
		N406	44° 16' 57.39"	17° 49' 38.46"	4320
		N407	44° 27' 27.59"	17° 48' 7.54"	4930
		N408	43° 52' 32.74"	17° 46' 34.27"	1360

made by using Excel solver in Microsoft Excel. The implementation of the optimized procedure in Excel solver provides  $\lambda_{best}$  and  $CN_{best}$ , where  $\lambda_{best}$  and  $CN_{best}$  are the best values of  $\lambda$  and  $CN$  for each rainfall-runoff event. Also, the same procedure is used for estimation of  $\lambda_{best}$  and  $CN_{best}$  for each sub-basin as a representative value from the recorded storms on the sub-basin. The mathematical formulation of the optimization

problem is given below. The estimated runoff  $\hat{Q}$  equation is obtained by Mishra and Singh (2013),

$$\hat{Q} = \frac{[P - 254\lambda(\frac{100}{CN} - 1)]^2}{P + 254(1-\lambda)(\frac{100}{CN} - 1)} \tag{5}$$

**Fig. 2** Sample of rainfall and runoff events at station N-404 in wadi Habawnah, South-West Saudi Arabia (event date 27 April 1985), Saudi Arabian Dames and Moore (1988)

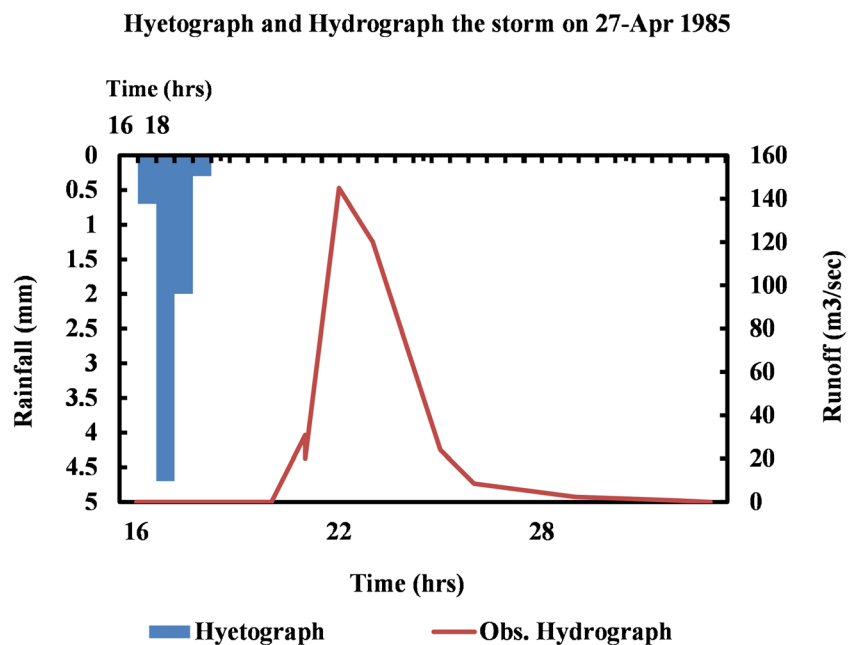
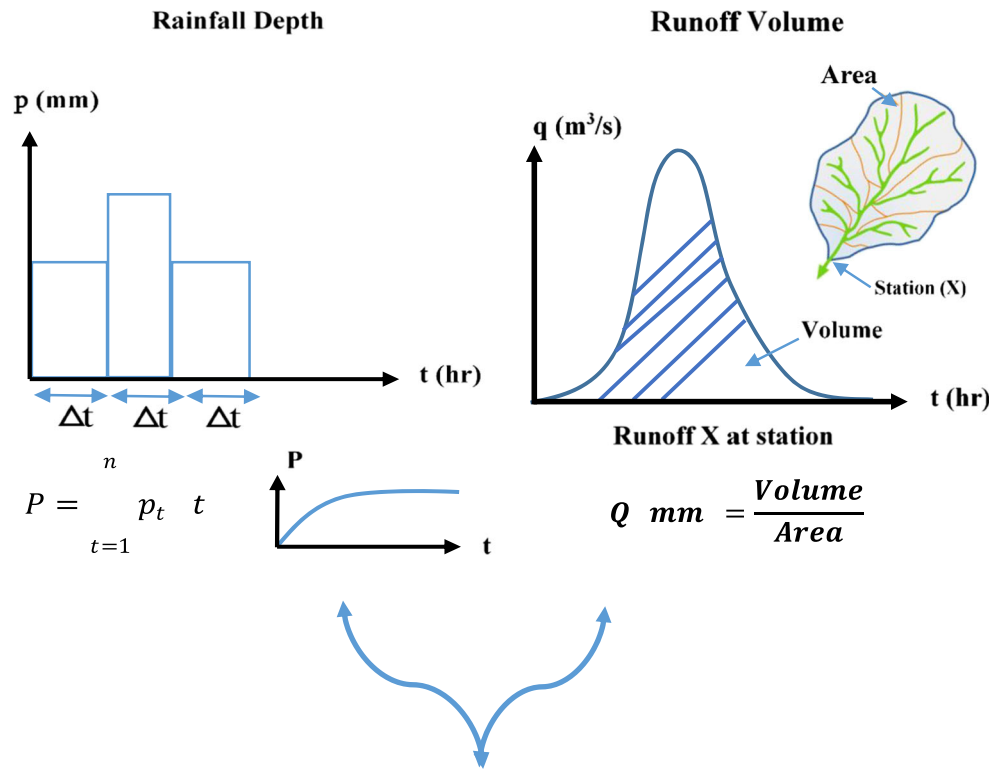


Fig. 3 Flowchart illustrates the estimation of CN from rainfall and runoff data



$$S = f(P, Q, \lambda)$$

$$S = \frac{P}{\lambda} + \frac{1 - \lambda Q - \frac{1 - \lambda^2 Q^2 + 4\lambda P Q}{2\lambda^2}}{2\lambda^2}$$

$$CN = 100 \frac{254}{S + 254}$$

The minimization problem is formulated by the objective function as Hawkins et al. (2008),

$$f(CN, \lambda) = \min \sum_{i=1}^n (Q_i - \hat{Q}_i)^2 \tag{6}$$

Where:  $Q_i$  is the observed runoff (mm),  $\hat{Q}_i$  is the estimated runoff (mm), and  $n$  is the no. of observations.

**Testing of distribution for rainfall, runoff and CN using Kolmogorov-Smirnov test (K-S test)**

The K-S test (Hamed and Rao 1999) was used to test the hypothesis based on the following:

- $H_0$  :  $z$  distributed as the specified distribution.
- $H_1$  :  $z$  is not distributed as the specified distribution.

Where,  $H_0$  is the null hypothesis,  $H_1$  is the alternative hypothesis, and  $z$  is the parameter under consideration, i.e., rainfall, runoff, CN, or 100-CN. This test compares the cumulative distribution function of a specified distribution, and the

empirical cumulative distribution function for a sample size  $n$ . The K-S test (Hamed and Rao 1999) looks at the max difference as given by,

$$K-S = \max_{i=1, \dots, n} |G(z_i, \theta) - G_n(z_i)| \tag{7}$$

Where,  $G(z_i, \theta)$  is the assumed cumulative distribution function with parameters  $\theta$ , and  $G_n(z_i)$  is the empirical cumulative distribution functions.

If the calculated value of the K-S test is greater than the tabulated value, the null hypothesis can be rejected at the level of significance  $\alpha$ .

Many distributions can represent the data, however, to find out the best out of these distributions, the Akaike Information Criterion (AIC) is computed as given by Laio et al. (2009),

$$AIC = -2 \sum_{i=1}^n \ln g(z_i) + 2N \tag{8}$$

Where,  $g(z_i)$  is the probability density function of the theoretical model evaluated at  $z_i$  and  $N$  is the number of parameters of the theoretical model. The minimum value of the AIC provides the best distribution to represent the data.

**Table 2** Summary of the Statistical Analysis of CN and  $\lambda$

Wadi	CN <sub>best</sub>			CN ( $\lambda=0.2$ )			$\overline{CN}_{best}$		
	$\lambda_{best}$	CN <sub>best</sub>	R*	$\lambda$	CN	R*	$\bar{\lambda}_{best}$	$\overline{CN}_{best}$	R*
Liyyah	0.0000	21.2	0.39	0.2	55.96	-0.15	0.0075	46.67	0.41
Habawnah	0.0000	43.6	0.21	0.2	80.55	-0.05	0.0199	65.41	0.20
Tabalah	0.0000	31.4	0.50	0.2	74.68	0.20	0.0090	52.62	0.50
AL-Lith	0.0017	29.8	0.99	0.2	63.60	0.89	0.0015	40.90	0.99
Yiba	0.0159	44.2	0.97	0.2	68.67	0.91	0.0147	54.34	0.96
Average	0.0035	34.0	0.61	0.2	68.69	0.36	0.0105	51.99	0.61
ALL-wadi	0.0000	32.1	0.86	0.2	65.39	0.75	0.0091	56.65	0.86

**Estimation of confidence intervals**

The confidence interval is widely used in statistical analysis. It describes the variability and accuracy of the parameter under study. In this paper, the confidence intervals were computed for CN. Since the CN shows gamma distribution at  $\lambda = 0.2$  and to make values comparable with the antecedent moisture condition (AMC-CN) tables, therefore, the confidence equations of the gamma distribution are given herein. The confidence intervals for the gamma distribution are expressed mathematically as (Kite 1977; and Hamed and Rao 1999),

$$CL = \overline{CN} \pm \left[ \left( \frac{\chi^2 \gamma}{4} \right) - \left( \frac{2}{\gamma} \right) \right] \sigma \tag{9}$$

$$\chi^2 = \nu \left[ 1 - \left( \frac{2}{9\nu} \right) + u \sqrt{\frac{2}{9\nu}} \right]^3 \tag{10}$$

$$\nu = \frac{8}{\gamma^2} \tag{11}$$

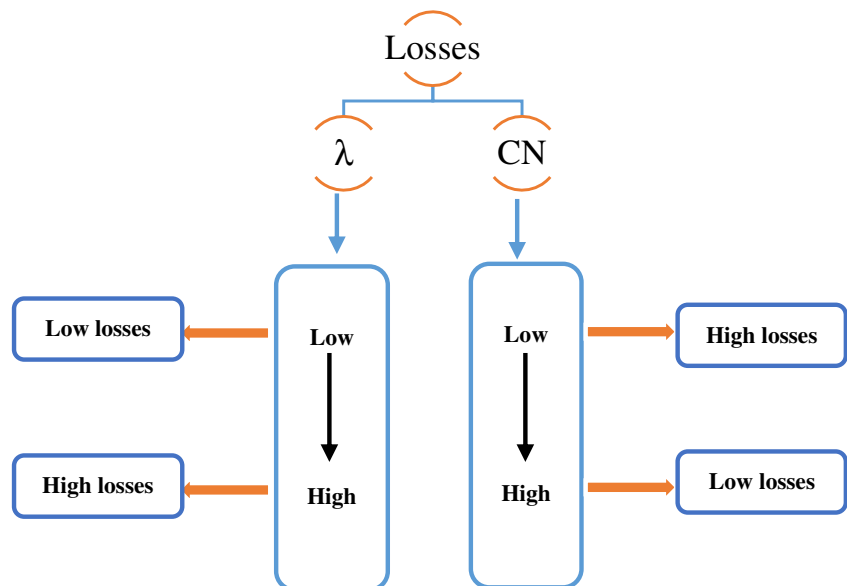
where: *CL* is the confidence intervals,  $\overline{CN}$  is the average curve number (CN),  $\chi^2$  is the chi-square value at probability, *p*,  $\gamma$  is the skewness coefficient,  $\sigma$  is the standard deviation,  $\nu$  is the degrees of freedom, and *u* is the standard normal variate for the *p* value (0.9, ..., and 0.995).

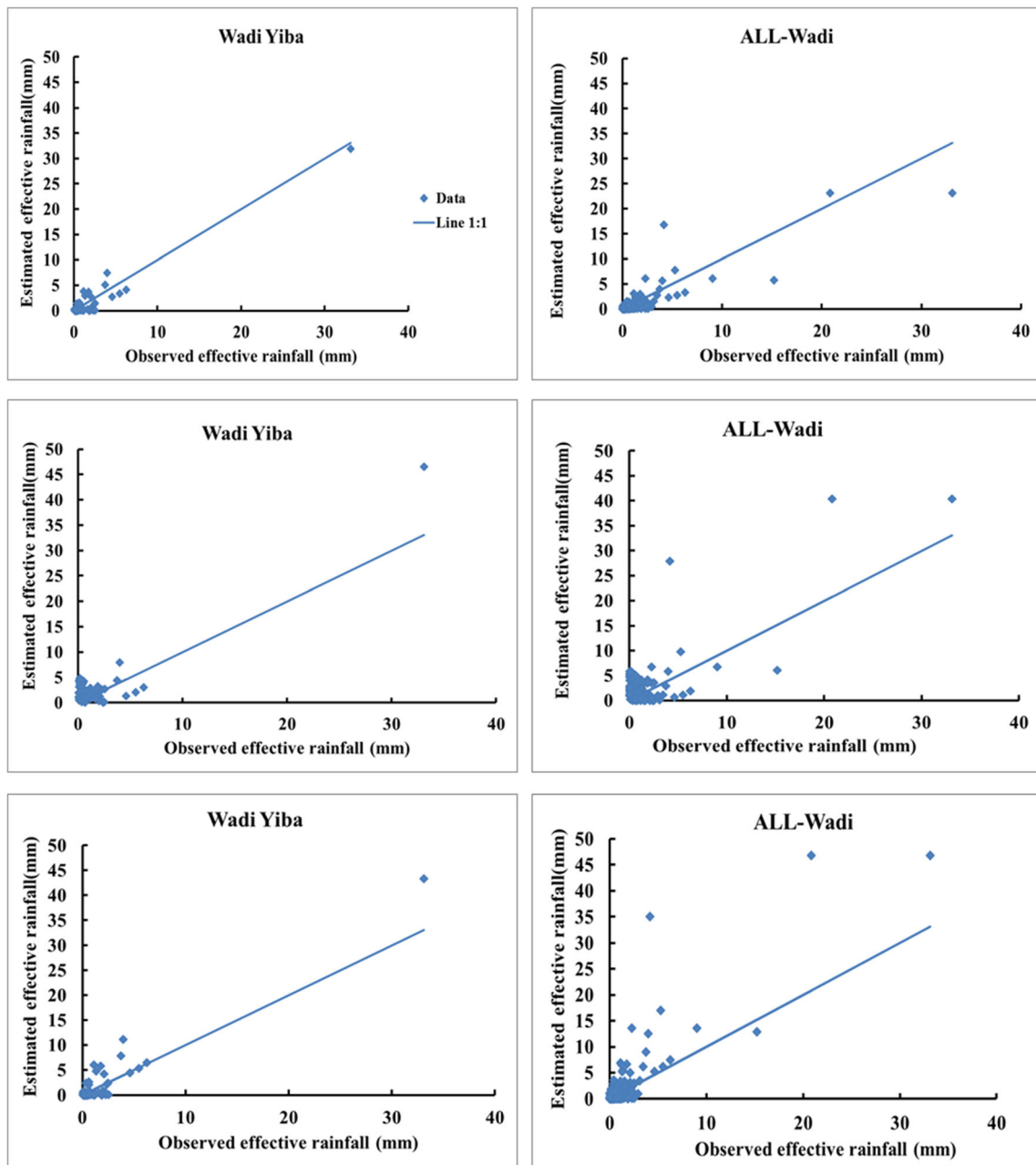
**Results and discussions**

**Analysis of CN and  $\lambda$**

Table 2 summarizes the estimated values of CN and  $\lambda$  based on the estimation methods described in “Estimation of the CN and  $\lambda$ ” It is obvious from the table (the second and third

**Fig. 4** Conceptual model of the relationship between the losses in NRCS-CN method: Abstraction losses ( $\lambda$ ) and infiltration losses (CN)





**Fig. 5** Scatter plot for the observed and estimated effective rainfall based the different estimation methods for wadi Yiba as an example and all

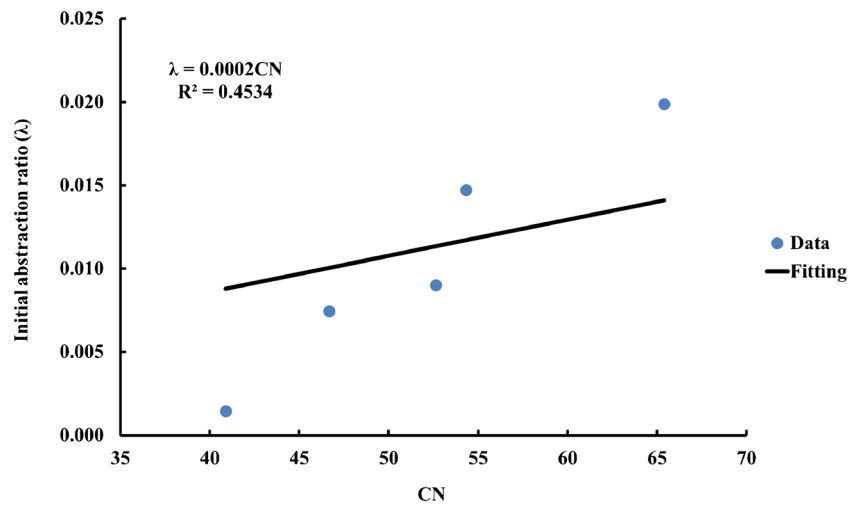
wadis: First row is Method 1 at ( $\lambda_{best}$  and  $CN_{best}$ ), second row is Method 2 at  $CN_{(\lambda=0.2)}$  and third row is Method 3 at  $\bar{\lambda}_{best}$  and  $\overline{CN}_{best}$

columns) that  $\lambda_{best}$  and  $CN_{best}$  provide the lowest estimation values which may account for unrealistically high transmission losses in the basins. The table shows  $\lambda = 0$  in three basin out of five. This also explains the reason why low CN values are obtained since most of the losses are due to infiltration and no abstraction losses which it seems unrealistic. Figure 4 shows a conceptual model for the relationship between losses in the SCS-CN method. There are two kinds of losses namely: the abstraction losses described by  $\lambda$  and the infiltration losses described by the CN. The high value of  $\lambda$  leads to high values

of CN and vice-versa for a given rainfall and runoff. In arid regions where transmission losses, due to high infiltration in a dry bed, are the main feature in arid and semi-arid regions would lead to having a low CN values and consequently low value of  $\lambda$ . In the case of  $CN_{(\lambda=0.2)}$ , which appears in the fifth and sixth columns in the table, the values of CN are the highest. The value of  $\lambda = 0.2$  does not reflect the behavior of an arid basin (i.e., incorporating transmission losses in the CN). This result is also confirmed by the work presented by Yuan et al. (2014). The third estimation method, which



**Fig. 6** Relationship between initial abstraction ratio  $\lambda$  and curve number CN for the averaging over the events (i.e.,  $\bar{\lambda}_{best}$  and  $\bar{CN}_{best}$ )



appears in the eighth and ninth columns in the table, shows the case of  $\bar{\lambda}_{best}$  and  $\bar{CN}_{best}$ . The estimated values fall in between the two above cases. This is because it is an averaging over the results of every event. The third estimation method seems to provide more realistic values for arid basins with realistic transmission losses since  $\lambda$  does not equal to zero and therefore the values of CN are not too low. It is also obvious from Table 2 that average values of  $\lambda$  and CN for all wadis are 0 and 34 for estimation method 1, 0.2, and 68.69 for estimation Method 2, and 0.0105 and 51.99 for estimation Method 3 respectively. The average correlation coefficient between estimated and observed effective rainfall is moderate (0.61) for both estimation Methods 1 and 3 while it is low for estimation Method 2 (0.36). This confirms that  $\lambda = 0.2$  is not suited for such wadis and the appropriate values of  $\lambda$  lies between 0 and 0.0105. However, the value of  $\lambda = 0.0105$  seems to be more realistic since the corresponding CN of 52 is within the acceptable field data in SCS-CN table (*National Engineering Handbook* Section 4 (NEH-4) USDA-SCS (1985)). The last row in Table 2 shows the results of the three estimation methods considering all the wadis. The results show that the estimation Methods 1 and 3 provides the highest correlation of 0.86, while, the second estimation method provides a correlation of 0.75.

Figure 5 shows scatter plots between estimated and observed effective rainfall based on the three estimation methods for wadi Yiba (left column) as a sample and for all wadi (right column). The corresponding correlation coefficient for each estimation method is presented in Table 2 for each wadi and for all wadis. Wadi Al-Lith and Yiba show the highest correlation coefficients amongst the wadis for the three estimation methods. Wadi Tabalah and Liyyah provide moderate correlation for estimation Methods 1 and 3, while poor correlation of estimation Method 2. Wadi Habawnah shows poor correlation for the three estimation methods. The worst correlation is for Method 2. The results of such analysis confirm that  $\lambda = 0.2$  is not a good choice for flood modeling in such area that covers the western part of Saudi Arabia.

Figure 6 shows a relationship between CN and  $\lambda$ . The relation confirms the conceptual model of SCS-CN loss model presented in Fig. 4. The figure shows a coefficient of determination,  $R^2$ , equals to 0.45 which is relatively moderate relationship between CN and  $\lambda$ . This relationship helps in estimating a meaningful value of  $\lambda$  based on CN in the basin and not relying on the customary value of 0.2.

**General descriptive statistics and probability distribution of P, Q, and CN**

An Excel sheet is designed to calculate the descriptive statistics for grouped and ungrouped data and also to perform hypothesis testing for fitting theoretical a probability distribution function to the data. The main statistical measures of the data are presented in Table 3. These statistics provide the arithmetic mean, standard deviation, SD, skewness coefficient,  $\gamma$ , kurtosis coefficient, median and coefficient of variation, CV, respectively as sorted in the table. The CV for rainfall is 1.13 that shows relatively high degree of variability which is common in arid regions. However, for the runoff the CV is 2.51 which shows much more variability than in rainfall. It is clear

**Table 3** Summary of the statistical analysis

	Mean	SD*	Skewness	Kurtosis	Median	CV*
Rainfall	16.85	19.02	3.28	13.80	11.03	1.13
Runoff	1.36	3.41	6.79	54.24	0.48	2.51
$CN_{best}$	57.84	24.26	-0.04	-1.08	57.18	0.42
$CN_{(\lambda=0.2)}$	85.87	10.45	-1.14	1.37	88.34	0.12
$100-CN_{best}$	42.16	24.26	0.04	-1.08	42.82	0.58
$100-CN_{(\lambda=0.2)}$	14.13	10.45	1.14	1.37	11.66	0.74

SD\*, standard deviation; CV\*, coefficient of variation

**Table 4** Summary of the hypothesis testing results at significant level ( $\alpha = 0.05$ )

	Gaussian	Log-normal	Exponential	Gamma	Beta	Gumbel
Rainfall	Reject	Accept	Accept	Reject	Reject	Reject
Runoff	Reject	Accept	Reject	Reject	Reject	Reject
CN <sub>best</sub>	Accept	Reject	Reject	Accept	Accept	Accept
CN <sub>(λ=0.2)</sub>	Reject	Reject	Reject	Reject	Accept	Reject
100-CN <sub>best</sub>	Accept	Reject	Reject	Reject	Accept	Accept
100-CN <sub>(λ=0.2)</sub>	Reject	Accept	Reject	Accept	Accept	Reject

from the table that the rainfall and runoff data are highly skewed since the coefficient of skewness is ranging between 3.28 and 6.79 respectively that is far from zero and it is negatively skewed. This observation reflects the relatively high frequency of low values of rainfall and runoff. For CN<sub>best</sub>, the CV is relatively higher than CV for CN<sub>(λ=0.2)</sub>. However, both show relatively low coefficient of variation when compared with rainfall and runoff. This reflects that the variability in CN is less than the variability in rainfall and runoff. The coefficient of variation shows low degree of variability for CN<sub>best</sub> while negligible variability is observed for CN<sub>(λ=0.2)</sub>. The skewness coefficient for CN is less than zero (i.e., it is positively skewed). This means the large values of CN has high frequency than the low CN values. This result is on the contrary of the skewness of rainfall and runoff.

The kurtosis indicates peaked distribution w.r.t. the normal distribution in both rainfall and runoff, since the values are larger than three, which is called leptokurtic. The minus sign in the table for the kurtosis is due to the subtraction of 3 to make it comparable with the normal distribution. The runoff distribution is extremely peaked when compared with the rainfall distribution since the kurtosis is 54.24 and 13.8 respectively. Sometimes, researchers prefer to use 100-CN instead of CN to perform statistical analysis, e.g., McCuen (2002). Therefore, in the current study both CN and 100-CN are considered. The mean for 100-CN is different from the mean of CN however, the SD is the same for both. The 100-CN is negatively skewed. The kurtosis is the same as for CN. The coefficient of variation, CV, of 100-CN is higher than CV for

CN. For CN<sub>(λ=0.2)</sub> and 100-CN<sub>(λ=0.2)</sub> the kurtosis is 1.37 which is low peaked, while in the cases of CN<sub>best</sub> and 100-CN<sub>best</sub> the kurtosis is less than three that is called platykurtic.

Several probability distribution functions (e.g., Gaussian, log-normal, exponential, gamma, beta, and Gumbel) are used to test the data. Tables 4, 5, and 6 display the results of the Kolmogorov-Smirnov (K-S) test. The results show that the K-S test accepts some distributions. According to Table 4, most cases give more than one distribution. The suitable distribution can be selected by comparing K-S sample values and K-S critical value. The lower value of the K-S sample than the K-S critical, the best is the distribution.

Figure 7 (first row) presents the cumulative distribution of the rainfall and runoff depths from all wadis and fitting of different theoretical distribution functions. According to the K-S test at 5% significant level, the log-normal distribution seems to fit well the data for both rainfall and runoff depths (see Tables 4 and 5). However, exponential distribution seems to fit also the rainfall depth. According the AIC criterion, the log-normal is the best as shown in Table 6.

Figure 7 (second row) and Fig. 8 (first row) present the cumulative distribution form the data and the frequency histogram of CN<sub>(λ=0.2)</sub> and CN<sub>best</sub> respectively. The K-S test at 5% significant level (Table 4) shows that Gaussian, gamma, and beta distribution can be accepted for CN<sub>best</sub>; however, beta distribution is accepted for CN<sub>(λ=0.2)</sub>. The AIC criterion shows that Beta distribution is the best for both (Table 6).

Figure 7 (third row) and Fig. 8 (second row) presents the cumulative distribution for the data and the

**Table 5** K-S test values compared with critical values

	Gaussian K-S	Log-normal K-S	Exponential K-S	Gamma K-S	Beta K-S	Gumbel K-S	K-S critical value 5%
Rainfall	–	<b>0.045</b>	0.092	–	–	–	0.107
Runoff	–	<b>0.048</b>	–	–	–	–	0.107
CN <sub>best</sub>	0.073	–	–	0.098	<b>0.042</b>	0.115	0.107
CN <sub>(λ=0.2)</sub>	–	–	–	–	<b>0.048</b>	–	0.107
100-CN <sub>best</sub>	0.073	–	–	–	<b>0.041</b>	0.118	0.107
100-CN <sub>(λ=0.2)</sub>	–	0.083	–	0.048	<b>0.035</b>	–	0.107

K-S, Kolmogorov-Smirnov. The bold face number shows the lowest K-S sample value. (–) means the distribution is rejected

**Table 6** AIC for K-S test

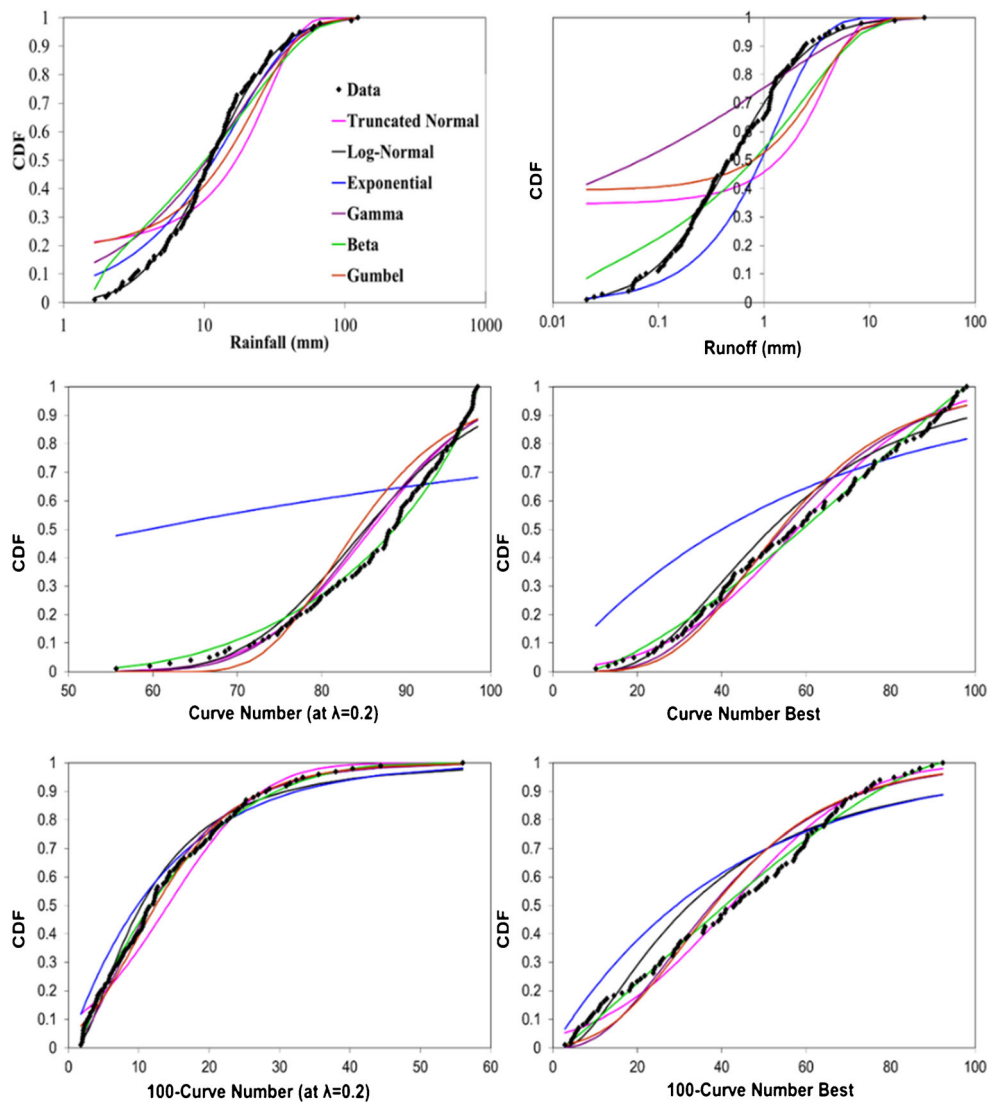
	Gaussian AIC	Log-normal AIC	Exponential AIC	Gamma AIC	Beta AIC	Gumbel AIC
Rainfall	–	<b>324.67</b>	332.39	–	–	–
Runoff	–	<b>94.42</b>	–	–	–	–
CN <sub>best</sub>	398.94	–	–	401.59	<b>387.77</b>	565.90
CN <sub>(λ=0.2)</sub>	–	–	–	–	<b>305.20</b>	–
100-CN <sub>best</sub>	399.61	–	–	–	<b>387.76</b>	629.48
100-CN <sub>(λ=0.2)</sub>	–	313.74	–	311.70	<b>304.02</b>	–

AIC, Akaike information criterion. The bold face number shows the lowest AIC value. (–): means the distribution is rejected

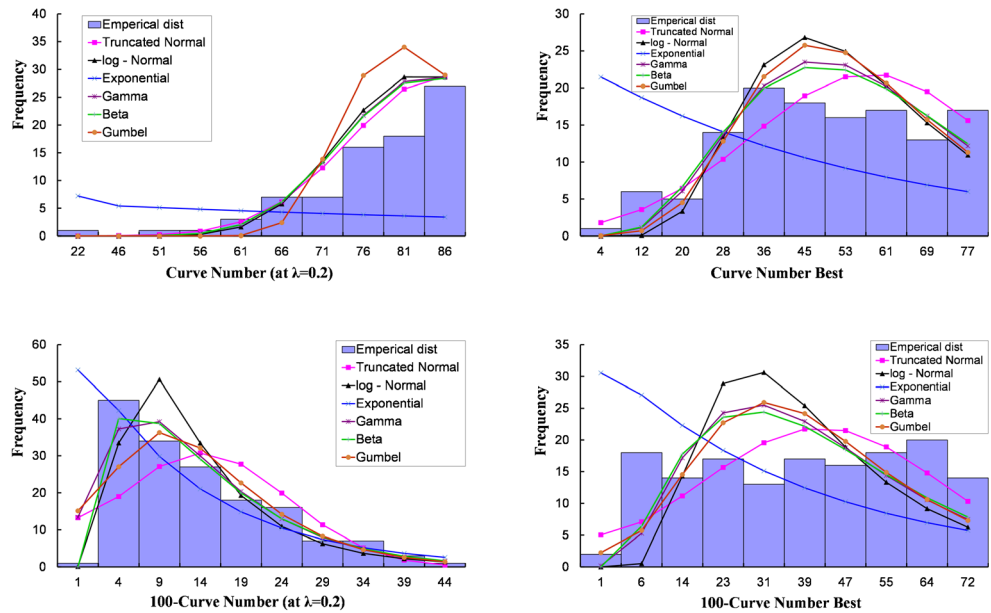
frequency histogram of 100-CN<sub>(λ=0.2)</sub> and 100-CN<sub>best</sub> respectively. Table 4 shows the results of the K-S at 5% significant level. Gaussian and beta can be accepted for 100-CN<sub>best</sub>. However, for 100-CN<sub>(λ=0.2)</sub> Lognormal,

gamma and beta can be accepted. The AIC criterion presented in Table 6 confirms that beta distribution is the best. It is worse mentioning that the acceptance of gamma distribution for 100-CN<sub>(λ=0.2)</sub> is in agreement

**Fig. 7** Fitting probability distribution function: First row, rainfall depth (left) and runoff (right), second row, CN at λ = 0.2 (left) and CN<sub>best</sub> (right), and third row, 100-CN at λ = 0.2 (left) and 100-CN<sub>best</sub> (right)



**Fig. 8** Frequency histogram of CN and 100-CN compared with different theoretical frequency distribution (CN is presented in the top row and 100-CN is presented in the bottom row) for  $\lambda = 0.2$  (left) and  $CN_{best}$  (right)



with earlier results of (McCuen 2002) for agricultural catchments in USA. However, beta seems to be the best for the study wadis in Saudi Arabia.

**Confidence intervals of CN**

The confidence bounds are presented in Table 7 for CNs ranging from 45 to 85, which is the approximate range for the sample values of CN from the five basins in SA. Limits for probabilities  $(1-\alpha)$  of 0.90, 0.95, 0.975, 0.99, and 0.995 were obtained from the cumulative distribution. Figure 9 shows limits of curve number and relationship between  $\overline{CN}$  average curve number,  $CN_U$  upper limit of CN at confidence level  $(1-\alpha)$  and  $CN_L$  lower limit of CN at confidence level  $(1-\alpha)$ . Figure 10 shows relationship between AMC CNs from *National Engineering Handbook* USDA-SCS (1963) at dry

(AMCI) and wet (AMCIII) and the upper and lower bounds of CN at different confidence levels of 0.90, 0.95, 0.975, 0.99, and 0.995 calculated in this study. From the analysis, it is found out that the AMC I and III bounds are wider than by as much as 9 CNs at confidence level 0.90. Table 8 shows the maximum difference between AMC bounds and the confidence Limits obtain in this study based on different confidence levels. Maximum at 0.90 and minimum at confidence level 0.975, 0.99, and 0.995.

**Conclusions**

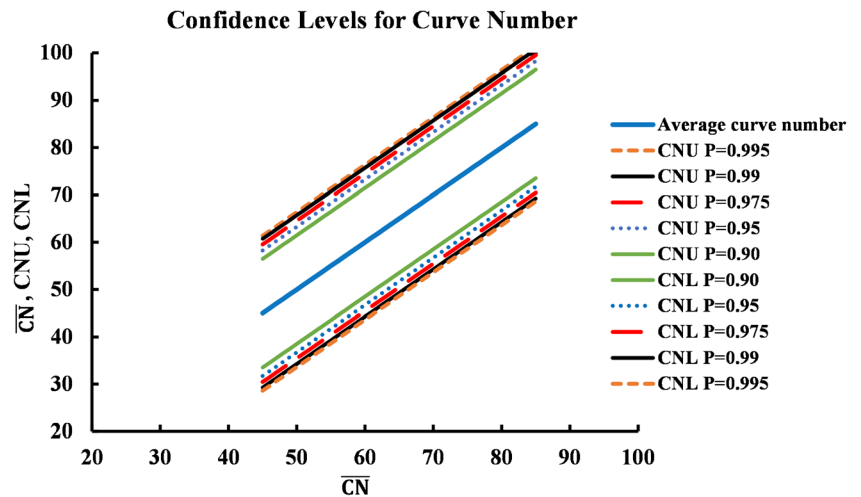
The current study leads to the following conclusions. From the statistical analysis of the data, it was found out that CN is event dependent. Therefore, the range of CN varied

**Table 7** Confidence limits for selected levels of confidence and runoff curve number

CN*	Upper confidence limit for $(1-\alpha)$ of					Lower confidence limit for $(1-\alpha)$ of				
	0.995	0.99	0.975	0.95	0.90	0.90	0.95	0.975	0.99	0.995
45	61.37	60.73	59.54	58.27	56.48	33.52	31.73	30.46	29.27	28.63
50	66.37	65.73	64.54	63.27	61.48	38.52	36.73	35.46	34.27	33.63
55	71.37	70.73	69.54	68.27	66.48	43.52	41.73	40.46	39.27	38.63
60	76.37	75.73	74.54	73.27	71.48	48.52	46.73	45.46	44.27	43.63
65	81.37	80.73	79.54	78.27	76.48	53.52	51.73	50.46	49.27	48.63
70	86.37	85.73	84.54	83.27	81.48	58.52	56.73	55.46	54.27	53.63
75	91.37	90.73	89.54	88.27	86.48	63.52	61.73	60.46	59.27	58.63
80	96.37	95.73	94.54	93.27	91.48	68.52	66.73	65.46	64.27	63.63
85	100.00	100.00	99.54	98.27	96.48	73.52	71.73	70.46	69.27	68.63

CN\* curve number, and  $(1-\alpha)$  confidence level

**Fig. 9** Limits of the curve number:  $\overline{CN}$  average curve number,  $CN_U$  is the upper limit of CN at confidence level  $(1-\alpha)$  and  $CN_L$  is the lower limit of CN at confidence level  $(1-\alpha)$

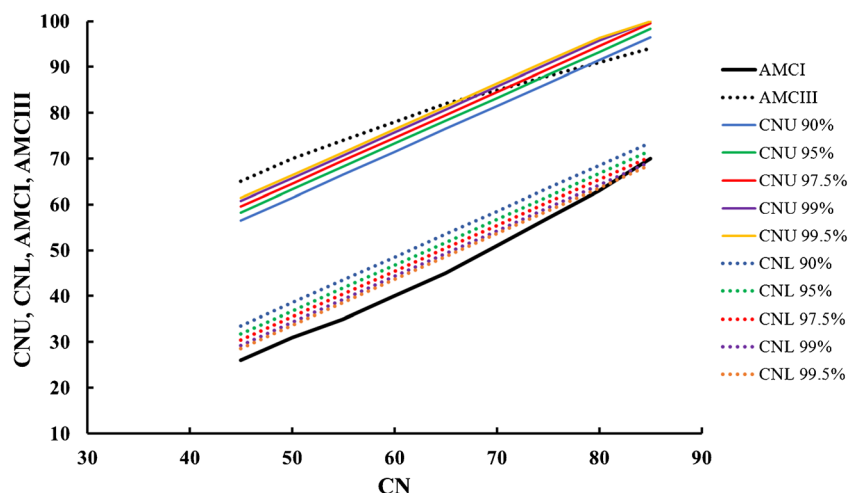


between 45 and 85 from the studied basins which deviates from the rural basins studied by McCuen (2002) that varied between 65 to 95. The reason is due to the transmission losses in the arid basins which provides lower CN values. The descriptive statistics showed a positively skewed distribution and kurtosis was leptokurtic for both rainfall and runoff. The statistical tests for rainfall and runoff exhibited log-normal distribution. The average initial abstraction ratio  $\bar{\lambda}_{best}$  was found to be 0.0105. This value is falling within the limits of the common range available in the literature (0.0–0.3). Since the average correlation coefficient between estimated and observed runoff is 0.61 at  $\lambda = 0.0105$  (based on average) and 0.86 at  $\lambda = 0.0091$  (based on all wadis) while it is 0.36 at  $\lambda = 0.2$  (based on average) and 0.75 at  $\lambda = 0.2$  (based on all wadis), therefore, the value of  $\lambda \approx 0.01$  seems to best represent the runoff in the study area. The value of  $\lambda = 0.01$  is more suited to arid region rather than the common value of 0.2, since it leads

to low value of CN that accounts for the transmission losses in the wadi. The statistical tests for  $(100-CN)$  at  $\lambda = 0.2$  showed that the gamma and beta distributions are the best for describing the variability of  $(100-CN)$ . These results are in agreement with the results of McCuen (2002) for the gamma distribution. The analysis for CN showed that the beta distribution is the best at  $\lambda = 0.2$  and  $\bar{\lambda}_{best} = 0.01$ . In this study, the confidence intervals for CN have been estimated and tabulated in Table 7 for different confidence levels ranging from 0.9 up to 0.995. These values are recommended to be used in the design of hydraulic structures in arid regions for risk-based analysis under uncertainty.

**Funding information** This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. (DF-172-155-1441). The authors, therefore, gratefully acknowledge DSR technical and financial support.

**Fig. 10** Comparison between CN confidence limits and AMCI and AMCIII. The Source for AMCI and AMCIII is the *National Engineering Handbook* USDA-SCS, (1963), AMCI: antecedent moisture condition I and AMCIII: antecedent moisture condition III



**Table 8** Max different between AMC CNs and  $CN_U$  and  $CN_L$ 

(1- $\alpha$ )	Max different of CN
0.90	9
0.95	7
0.975	6
0.99	6
0.995	6

## References

- Alagha, M. O., and Gutub, S. A., and Elfeki, A. M. (2016). Estimations of NRCS Curve Number from Watershed Morphometric Parameters: A Case Study of Yiba Watershed in Saudi Arabia. *International Journal of Civil Engineering and Technology*, 7(2), pp. 247–265.
- Baltas EA, Dervos NA, Mimikou MA (2007) Technical note: determination of the SCS initial abstraction ratio in an experimental watershed in Greece. *Hydrol Earth Syst Sci* 11(6):1825–1829
- Banasik, and Madeyski, and Wiezik, and Woodward. (1994). Applicability of Curve Number technique for runoff estimation from small Carpathian watersheds. *Development in Hydrology of Mountainous Areas* (Eds. L. Molnar, P. Miklanek & I. Meszaros), Slovak Committee for Hydrology and Institute of Hydrology Slovak Academy of Sciences, Stara Lesna, Slovakia, 125–126.
- Banasik, and Krajewski, and Sikorska, and Hejduk. (2014). Curve Number estimation for a small urban catchment from recorded rainfall-runoff events. *Archives of Environmental Protection*, 40(3), 75–86.
- Bonta JV (1997) Determination of watershed curve number using derived distributions. *J Irrig Drain Eng* 123(1):28–36
- Bosznay M (1989) Generalization of SCS curve number method. *J Irrig Drain Eng* 115(1):139–144
- Cazier, D. J., and Hawkins, R. H. (1984). Regional application of the curve number method. Paper presented at the water today and tomorrow: (pp. 710–710). ASCE
- Chen, C. (1982). An evaluation of the mathematics and physical significance of the soil conservation service curve number procedure for estimating runoff volume. Paper presented at the Proc., Int. Symp. On rainfall-runoff modeling, water resources Publ., Littleton, Colo., 387–418
- Hamed, K., and Rao, A. R. (1999). Flood frequency analysis: CRC press
- Hawkins (1973) Improved prediction of storm runoff in mountain watersheds. *J Irrig Drain Div* 99(4):519–523
- Hawkins (1993) Asymptotic determination of runoff curve numbers from data. *J Irrig Drain Eng* 119(2):334–345
- Hawkins, and Hjelmfelt Jr, A. T., and Zevenbergen, A. W. (1985). Runoff probability, storm depth, and curve numbers. *Journal of Irrigation and Drainage Engineering*, 111(4), 330–340.
- Hawkins, and Woodward, and Jiang. (2001). Investigation of the runoff curve number abstraction ratio. Paper presented at the USDA-NRCS Hydraulic Engineering Workshop, Tucson, Arizona.
- Hawkins, and Jiang, R., and Woodward, D. (2002). Application of Curve Number Method in Watershed Hydrology. Paper presented at the Presentation at American Water Resources Association Annual Meeting, Albuquerque NM.
- Hawkins, and Ward, T. J., and Woodward, D. E., and Van Mullem, J. A. (2008). Curve number hydrology: State of the practice.
- Hjelmfelt (1980) Empirical investigation of curve number technique. *J Hydraul Div* 106(9):1471–1476
- Hjelmfelt (1991) Investigation of curve number procedure. *J Hydraul Eng* 117(6):725–737
- Jiang, R. (2001). Investigation of runoff curve number initial abstraction ratio. (MS thesis, watershed management), University of Arizona, 120 pp.
- Kite, G. W. (1977). Frequency and risk analyses in hydrology. *Water Resour Publ*, 224
- Laio, F., and Di Baldassarre, G., and Montanari, A. (2009). Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research*, 45(7).
- McCuen. (1982). A guide to hydrologic analysis using SCS methods: prentice-hall, Inc.
- McCuen (2002) Approach to confidence interval estimation for curve numbers. *Hydrol Eng Amer Soc Civ Eng* 7(1):43–48
- Mishra SK, Singh VP (2002) SCS-CN method. Part 1: derivation of SCS-CN-based models. *Acta Geophysica Polonica* 50(3):457–477
- Mishra S, Singh VP (2004) Validity and extension of the SCS-CN method for computing infiltration and rainfall-excess rates. *Hydrol Process* 18(17):3323–3345
- Mishra SK, and Singh, V. (2013). Soil conservation service curve number (SCS-CN) methodology (Vol. 42): Springer Science & Business Media
- Pilgrim D, Cordery I (1993) Flood runoff. Chapter 9. In: Maidment D (ed) *Handbook of hydrology*. McGraw-hill, Inc, New York
- Ponce H (1996) Runoff curve number: has it reached maturity? *J Hydrol Eng* 1(1):11–19
- Rallison, R. E. (1980). Origin and evolution of the SCS runoff equation. Paper presented at the symposium on watershed management 1980
- Rallison RE, Cronshey RG (1979) Discussion of runoff curve numbers with varying site moisture by Richard H. Hawkins *J Irrig Drain Div* 105(4):439–439
- Rallison, R. E., and Miller, N. (1982). Past, present, and future SCS runoff procedure. Paper presented at the rainfall-runoff relationship/proceedings, international symposium on rainfall-runoff modeling held May 18–21, 1981 at Mississippi State University, Mississippi State, Mississippi, USA/edited by VP Singh
- Ramasastri, K., and Seth, S. (1985). Rainfall-runoff relationships. Rep. RN-20, National Institute of hydrology, Roorkee, Uttar Pradesh
- Rawls, W. J., and Maidment, D. R. (1993). Infiltration and soil water movement: McGraw-hill
- Saudi Arabian Dames and Moore. (1988). Representative basins study for Wadi: Yiba, Habwnah, Tabalah, Liyyah and Al-Lith (Main report) Kingdom of Saudi Arabia, Ministry of Agriculture and Water, Water Resource Development Department. Retrieved from
- Schneider LE, McCuen RH (2005) Statistical guidelines for curve number generation. *J Irrig Drain Eng* 131(3):282–290
- Soullis K, Dercas N (2007) Development of a GIS-based spatially distributed continuous hydrological model and its first application. *Water Int* 32(1):177–192
- Soullis KX, Valiantzas JD (2012) SCS-CN parameter determination using rainfall-runoff data in heterogeneous watersheds-the two-CN system approach. *Hydrol Earth Syst Sci* 16(3)
- Soullis KX, Valiantzas JD (2013) Identification of the SCS-CN parameter spatial distribution using rainfall-runoff data in heterogeneous watersheds. *Water Resour Manag* 27(6):1737–1749
- USDA-NRCS (2004) National Engineering Handbook: section 4: hydrology, soil conservation service. USDA, Washington
- USDA-SCS (1963) National Engineering Handbook Section 4: hydrology, soil conservation service. USDA, Washington
- USDA-SCS (1972) National Engineering Handbook: section 4: hydrology, soil conservation service. USDA, Washington
- USDA-SCS (1985) National Engineering Handbook Section 4: hydrology, soil conservation service. USDA, Washington
- Woodward, D. E., and Hawkins, R. H., and Jiang, R., and Hjelmfelt, J., Allen T, and Van Mullem, J. A., and Quan, Q. D. (2003). Runoff curve number method: examination of the initial abstraction ratio. Paper presented at the World Water & Environmental Resources Congress 2003.

- Yuan, Y., and Nie, W., and McCutcheon, S. C., and Taguas, E. V. (2014). Initial abstraction and curve numbers for semiarid watersheds in Southeastern Arizona. *Hydrological Processes*, 28(3), 774-783.
- Zevenbergen, A. W. (1985). Runoff curve numbers for rangeland from Landsat data. Technical rep. HL85-1, United States Dept. of agriculture-Agricultural Research Service hydrology laboratory, Beltsville, 72