Check for updates

Sara4r: an R graphical user interface (GUI) to estimate watershed surface runoff applying the NRCS – curve number method

Rafael Hernández-Guzmán, Arturo Ruiz-Luna and Eduardo Mendoza

ABSTRACT

This paper introduces a graphical user interface (GUI) for the R software that allows the rainfall-runoff relationship to be calculated, using the curve number method. This GUI is a raster-tool whose outputs are runoff estimates calculated using land use/land cover and hydrologic soil group maps. The package allows the user to select among three different antecedent moisture conditions and includes modifications about the initial abstraction parameter. We tested this GUI with data derived from two watersheds in Mexico and the outputs were compared with those produced using a well-established GIS tool in a vector environment. The results produced by these two approaches were practically the same. The main advantages of our package are: (1) '*Sara4r*' is faster than previous vector based tools; (2) it is easy to use, even for people with no previous experience using R; (3) the modular design allows the integration of new routines; and (4) it is free and open source.

Key words | antecedent moisture condition, curve number, hydrologic modeling, runoff, watershed

HIGHLIGHTS

- An easy-to-use R-based graphical user interface to conduct surface runoff estimation was developed.
- Outputs from the newly NRCS-CN based tool are practically the same as a well-established GIS tool in a vector environment.
- Being a free and open source software, it will allow us to incorporate new routines and evolve into a more robust tool.
- Sara4r package offers more calculation speed.

Rafael Hernández-Guzmán (corresponding author)
Instituto de Investigaciones sobre los Recursos Naturales,
Universidad Michoacana de San Nicolás de Hidalgo,
Av. San Juanito Itzícuaro s/n. Col. Nueva Esperanza. CP 58330.
Morelia, Michoacán,
México
E-mail: *rhernandez.g@gmail.com*

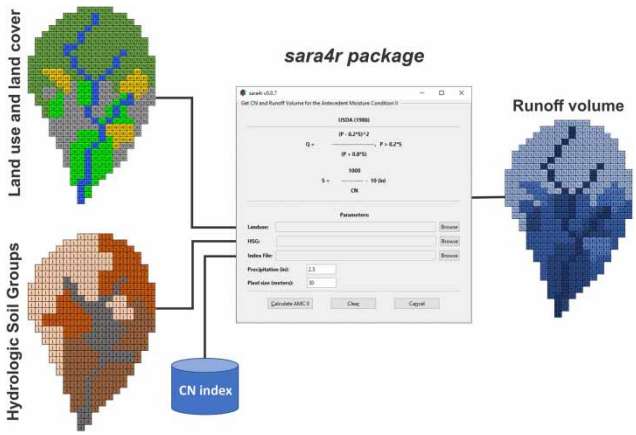
Rafael Hernández-Guzmán
Eduardo Mendoza
Laboratorio de Análisis para la Conservación de la Biodiversidad,
Instituto de Investigaciones sobre los Recursos Naturales, Universidad Michoacana de San Nicolás de Hidalgo,
Av. San Juanito Itzícuaro s/n. Col. Nueva Esperanza, CP 58330, Morelia, Michoacán,
México

Arturo Ruiz-Luna
Laboratorio de Manejo Ambiental,
Centro de Investigación en Alimentación y Desarrollo,
A.C. (CIAD – Mazatlán), Sábalo-Cerritos s/n, 82112, Mazatlán, Sinaloa,
México

doi: 10.2166/hydro.2020.087

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

GRAPHICAL ABSTRACT



INTRODUCTION

Hydrologic modeling is one of the most important fields of research within environment-related disciplines. An accurate estimation of runoff caused by rainfall is required for a comprehensive management of water resources and the analysis of possible environmental impacts and water management strategies (Chen *et al.* 2017; Fernández-Pato *et al.* 2018; Abuzied & Mansour 2019; Jiang *et al.* 2019). The understanding of the relationship between both variables is important to anticipate the potential effects of floods during extreme climate events (Verma *et al.* 2018; Beg *et al.* 2019). It is also useful to estimate the amount of available water for irrigation in agriculture (Essaid & Caldwell 2017), or the potential for energy generation in hydroelectric plants (Reichl & Hack 2017). Moreover, it helps the design of hydraulic infrastructure and the evaluation of watershed management practices (Verma *et al.* 2017). Finally, surface runoff plays an important role in biogeochemical cycles.

Several models have been developed to generalize hydrologic processes such as the estimation of direct surface runoff from rainfall data (Chen *et al.* 2017). Among them, the empiric method of the Natural Resources Conservation Service (NRCS) curve number (CN), which relates land use and land cover (LULC), antecedent runoff condition and hydrologic soil groups (HSG), among other parameters (USDA 1986), has gained wide acceptance in engineering practice. This is a simplified method, used at watershed level, to estimate direct runoff caused by a rainfall event (Hawkins *et al.* 2019), using measurable characteristics of the basin such as slope, soil types, and land cover and land uses.

Since its creation, this method has been rapidly adopted in many regions around the world regardless of land uses or climatic conditions (Soulis & Valiantzas 2012; Verma *et al.* 2017). The main reason behind its wide use is that it is easy to apply due to the fact that the main input characteristics are defined in terms of land use and soil type (Hernández-Guzmán & Ruiz-Luna 2013; Walega *et al.* 2020). Regarding this, Ponce & Hawkins (1996) have discussed the pros and cons of using it, while Verma *et al.* (2017) and Hawkins *et al.* (2019) present a critical review of the NRCS-CN concept and associated models, their advantages and limitations.

Despite its widespread use, there is no general agreement regarding the methodology to estimate the CN parameter values from measured rainfall-runoff data (Soulis & Valiantzas 2012). Many alternatives have been explored to include different Antecedent Moisture Conditions (AMC) and initial abstraction (I_a) values. There are three AMC conditions for dry (AMC I), average (AMC II) and saturated soils (AMC III), that are assigned as a function of the five-day antecedent rainfall. There are two accepted values of I_a (the initial fraction of the storm depth after which runoff begins) (Hernández-Guzmán & Ruiz-Luna 2013).

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

Recent contributions on the topic include remote sensing (RS) and geographic information system (GIS) techniques which allow the development of continuous hydrological simulation models (see Verma *et al.* 2017, for a complete list of software). Following this approach, Hernández-Guzmán *et al.* (2011) developed CN-Idris, a raster-based tool incorporating the NRCS-CN method as a module for Idrisi software. A subsequent tool named SARA (Spatial Analysis of surface Runoff in ArcGIS) was developed as a vector-based tool (Hernández-Guzmán & Ruiz-Luna 2013).

In recent years the use of the freely available R software (https://www.r-project.org/) has gained great popularity among ecologists and environmental scientists, as a tool to accomplish different analytical tasks in an integrated way. In particular, the R software is increasingly used for the analysis, modeling and visualization of spatial data (Hesselbarth *et al.* 2019), capacities that can be improved with the installation of many freely available routines (libraries).

Given the above considerations, the aim of this paper is to introduce the R package '*Sara4r*', a graphical user interface (GUI) to conduct runoff estimation based on the NRCS-CN method, in a similar way to that used by modules developed for Idrisi and ArcGIS software (Hernández-Guzmán *et al.* 2011; Hernández-Guzmán & Ruiz-Luna 2013). This paper includes a general description of the NRCS-CN method, suitable as a manual for new users, inclusive for those with little or no experience at all using R (Appendix A1).

The '*Sara4r*' package enables the calculation of surface runoff in watersheds with different conditions, as is exemplified using two case studies. The first case is for the San Nicolás – Cuitzmala watershed, an exorheic basin draining to the Pacific Ocean, mostly covered by dry and evergreen forests, with less than 1% of the area corresponding to human settlements. The second case is the endorheic Lake Cuitzeo watershed, a highly modified watershed in central Mexico. These watersheds were selected due to their biological and economic importance, respectively. On the one hand, the San Nicolás – Cuitzmala watershed is home to many endangered and threatened species. On the other hand, the lake Cuitzeo watershed holds the second largest lake in Mexico.

METHODS

The NRCS-CN method is described in detail in the National Engineering Handbook Section 4: Hydrology (NEH-4) from the United States Department of Agriculture (USDA 1986). Thus, we only briefly describe the model used to develop our interface.

The NRCS-CN method

A general form of the NRCS-CN model can be expressed as follows:

$$Q = \frac{(P - Ia)^2}{P - Ia + S} \; if \; P > Ia \tag{1}$$

$$else \; Q = 0,$$

$$Ia = \lambda S \tag{2}$$

where Q is the actual direct runoff, P is the cumulative precipitation for event, Ia is the initial abstraction, a parameter that includes rainfall losses previous to runoff, due to interception by leaves, evaporation, surface depressions and infiltration, S is the potential maximum retention and λ is the initial abstraction coefficient (dimensionless).

The original method assumes that the ratio $Ia/S = \lambda$ takes a value of 0.2, thus, the most popular form for the NRCS – CN method is as follows:

$$Q = \begin{cases} \dfrac{(P - 0.2S)^2}{(P + 0.8S)} & if \quad P > 0.2S \\ 0.0 & if \quad P \leq 0.2S \end{cases} \tag{3}$$

The parameter S depends on the soil capacity for water retention or runoff; therefore, it is possible to estimate its value using the Curve Number (CN), an empirical parameter based on the soil characteristics. It is defined by the type of soil (classified according its minimum infiltration rate), land cover and land use (such as vegetation, bare soil, and impervious surfaces), treatment (such as mechanical practices and management of cultivated agricultural lands) and hydrologic

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

conditions (poor, fair, good):

$$S = \frac{1000}{CN} - 10;\ (in) \tag{4}$$

The CN is dimensionless and has a range between 0 and 100. These conditions represent the extremes between total infiltration (runoff = 0) and totally saturated watersheds (rainfall = runoff) (USDA 1986).

To estimate CN values, the NRCS provides runoff curve number tables for different cover types, hydrologic conditions and hydrologic soil groups (HSG). The HSG is a standard soil classification (A, B, C, D) that depends on soil texture and infiltration rates. The A group includes well-drained soils with a high rate of infiltration, whereas D soils are poorly drained with a permanently high water table (USDA 1986).

Antecedent moisture condition (AMC)

As mentioned above, there are three AMC, with seasonal limits based on the five-day antecedent rainfall. In the growing season, it is less than 35.6 mm, from 35.6 to 53.3 mm, and more than 53.3 mm for the lowest runoff potential (AMC I), the average condition (AMC II) and the highest runoff potential (AMC III), respectively. Whereas for the dormant season, their values are less than 12.7 mm, from 12.7 to 27.9 mm and more than 27.9 mm for AMC I, II, and III (Feyereisen *et al.* 2008). These options are included in the '*Sara4r*' package, with the AMC II as default, and AMC I and III as optional depending on the user's knowledge about the soil conditions, considering that runoff estimates could vary depending on the AMC selected (Hernández-Guzmán & Ruiz-Luna 2013).

Values for CN II, considered the standard CN, are derived from the NEH-4 tables and, if necessary, converted to AMC I or AMC III (CN I or CN III) using the following functions (Mishra *et al.* 2008):

$$CN_I = \frac{CN_{II}}{2.2754 - 0.012754 CN_{II}} \tag{5}$$

$$CN_{III} = \frac{CN_{II}}{0.430 + 0.0057 CN_{II}} \tag{6}$$

Modified SCS-CN method (Ia/S = 0.05)

The λ value in the original method is 0.2, though it has been criticized and currently ranges from 0.02 to 0.07 (Hawkins *et al.* 2010; D'Asaro & Grillone 2012; Ling *et al.* 2020). A re-examination by Hawkins *et al.* (2002) concluded that $\lambda = 0.05$ is a better approximation and they recommend this value for the runoff calculation, using the next equation:

$$Q = \frac{(P - 0.05 S_{0.05})^2}{(P + 0.95 S_{0.05})};\ P > 0.05S \tag{7}$$

The existing handbook CN tables are based on the assumption of $\lambda = 0.20$, which is why a transfer to equivalent CNs having $\lambda = 0.05$ must be calculated using the relationship between S = 0.05 and S = 0.2 defined by Hawkins *et al.* (2002):

$$S_{0.05} = 1.33 (S_{0.2})^{1.15};\ (in) \tag{8}$$

Implementation of the 'Sara4r' package

All the above equations are included in the '*Sara4r*' package, whose Graphical User Interface (GUI; Figure 1) was codified using the RGTk2 package for R (Lawrence & Lang 2019). Taking into account that much of the spatial analysis used inside this package comprises map algebra functions

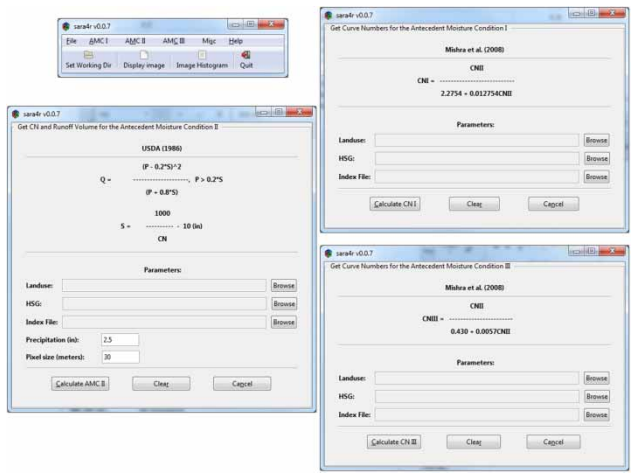


Figure 1 | Main screen of '*Sara4r*' package graphical user interface (GUI) for input data and parameters specification.

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

(i.e. raster display, overlay, reclassify), '*Sara4r*' has a dependency on the well-established 'raster' R package (Hijmans *et al.* 2020).

The basic raster layers to be input (LULC and hydrologic soils groups (HSG) maps) require geographical and geometric compatibility. Each LULC category must be codified using multiples of ten (e.g. agriculture = 10, forest = 20, bare soils = 30, etc.), while the HSGs are codified numerically (A = 1, B = 2, C = 3, D = 4). Then, the overlay of both layers produces all the possible combinations which are then stored in a new layer called 'Landsoil'.

A comma delimited file (*.csv) should be created previously with all the possible combinations in the Landsoil file, together with the CN values in the TR-55 handbook (USDA 1986). Once this information is uploaded, together with rainfall data values (inches) and pixel size (meters), CN values are assigned to every pixel in the map through a reclassification process.

The potential water retention after runoff begins (S) and the initial abstraction (I_a) parameters are estimated in the background. If precipitation value (P) is less than or equal to I_a, then Q depth (in) is codified as zero. Finally, when the pixel size (m) is specified, Qvol (m^3) is estimated. All outputs are stored in the working directory selected for the analysis, as text files in GeoTIFF format that can easily be imported into any GIS software for further analysis.

The '*Sara4r*' package is designed to output CN maps, Q depth and runoff volume, and CN and runoff volume based in a series of selectable options accessed using pull-down menus for the initial abstraction (I_a) and the AMC (Figure S1).

CASE STUDIES

Considering the two case studies, 30 m-resolution LULC maps were derived from the classification of Landsat images obtained from the USGS Global Visualization Viewer (https://glovis.usgs.gov/). The study areas were delineated based on the watershed limits and classified using the combined iterative minimum distance and hill climbing variants, implemented into the K-means unsupervised algorithm in SAGA GIS v7.5 (Conrad *et al.* 2015). The classification scheme covered the following LULC classes codified by tens: water bodies, mangrove, evergreen forest, tropical dry forest, secondary succession, crops, grasslands, littoral, and human settlements.

Soil types were defined based on the Soil vector data set scale 1: 250,000 series II (National Continuous) from the National Institute of Statistics and Geography (INEGI by its name in Spanish: Instituto Nacional de Estadística y Geografía). The polygons were classified according to the criteria of the World Reference Base for Soil Resources (WRB) (FAO 2014) and subsequently reclassified into four hydrological soil groups (HSG) using the curve number method (USDA 1986).

Rainfall data recorded on a daily basis were retrieved from meteorological stations of the Mexican National Meteorological Service (SMN), distributed in or close to the watersheds under study. The observation period is 1935–2018 with a recorded length mean equal to 30 years. As a preliminary data filter, all precipitation events $\leq$1.0 inch were removed from the databases. The number of storm events from 1980 to 2017 ranged between 108 and 246 (mean 162 events by station). A rainfall value was calculated for each watershed based on the arithmetic mean of all the stations they included.

Regarding the number of meteorological stations in Lake Cuitzeo watershed and to provide a reliably spatial representation of precipitation variation in there, we interpolated the precipitation data recorded by each station, using the inverse distance weighting (IDW) algorithm. A mean areal rainfall value was then obtained from the interpolated rainfall map.

Finally, to provide a performance measure we used these inputs to make surface runoff evaluations. We used ordered rainfall data and runoff values were compared with those obtained with a well-established GIS tool (SARA v1.0) using the same data but in a vector environment. To minimize the biasing effect of small storms, only larger storms were used ($P \geq 25.4$ mm).

Case study 1. San Nicolás – Cuitzmala watershed

The San Nicolás – Cuitzmala watershed, with a total area of about 3,900 km^2, is located between 19°19′ and 20°15′ N and 105°17′ and 104°30′ W, in west-central Mexico

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

(Figure 2). It is an exorheic watershed draining to the Pacific Ocean, characterized by a tropical sub-humid climate, with a clearly defined dry season (November to May) and also a rainy season (87% falling between June to October). The precipitation varies from 800 mm in the coastal area to 1,500 mm in the upper parts of the basin. The mean annual temperature is 25.6 °C (Flores-Díaz *et al.* 2018; Lazos-Chavero *et al.* 2018).

This region is biologically important due the heterogeneity of the landscape that promotes high alpha and beta species diversity and high levels of endemism (Suazo-Ortuño *et al.* 2018). However, the conversion of natural

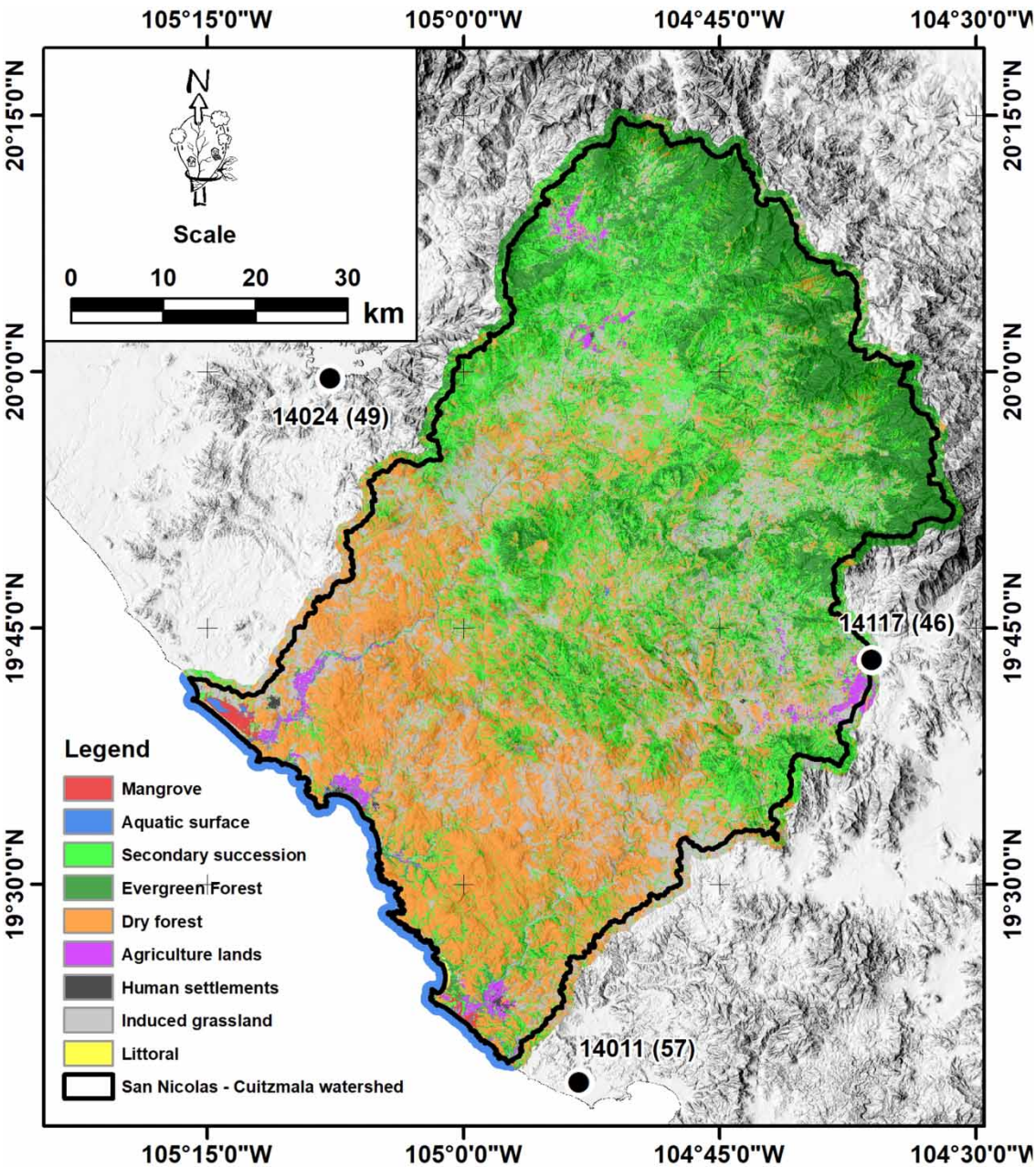


Figure 2 | Main land uses and land covers in the San Nicolás – Cuitzmala watershed, obtained through unsupervised classification of Landsat OLI image (2019). Black dots represent the meteorological stations used to estimate the P value for a typical storm event. The values in parentheses are the arithmetic mean of precipitation (mm).

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

cover into agricultural land has been recurrent, as evidenced by previous studies (Flores-Casas & Ortega-Huerta 2019; Hernández-Guzmán *et al.* 2019).

In general, this watershed is characterized by the presence in the upper part of evergreen forests covering about 18.3% of the total area, and dry forest (32.7%) distributed at the lower part of the watershed. Secondary succession (27.5%) and induced grassland (19.2%) are the main covers derived from anthropogenic activities. Regarding the soil characteristics, about 75% of the watershed is represented by the hydrologic soil group B, while 86% of the area acquired CN values between 50 and 75.

Given these considerations, '*Sara4r*' package was tested in this watershed using the antecedent moisture condition II and a precipitation value of 2 inches. The runoff estimate in the study area is about 17.75×10^6 m^3 from a unique storm event. Using the SARA v1.0 tool, outputs (17.74×10^6 m^3) are practically the same (differences <1%).

Case study 2. Lake Cuitzeo watershed

The Lake Cuitzeo watershed is an endorheic watershed of about 4,025 km^2 located in the trans Mexican Volcanic Belt in the states of Michoacán and Guanajuato, between the coordinates 19°24′ and 20°07′ N, and 100°37′ and 101°34′ W. Drainage concentrates in the lower part of the watershed, forming Lake Cuitzeo that covers an area of about 300 km^2 (Figure 3). It is the second largest lake in

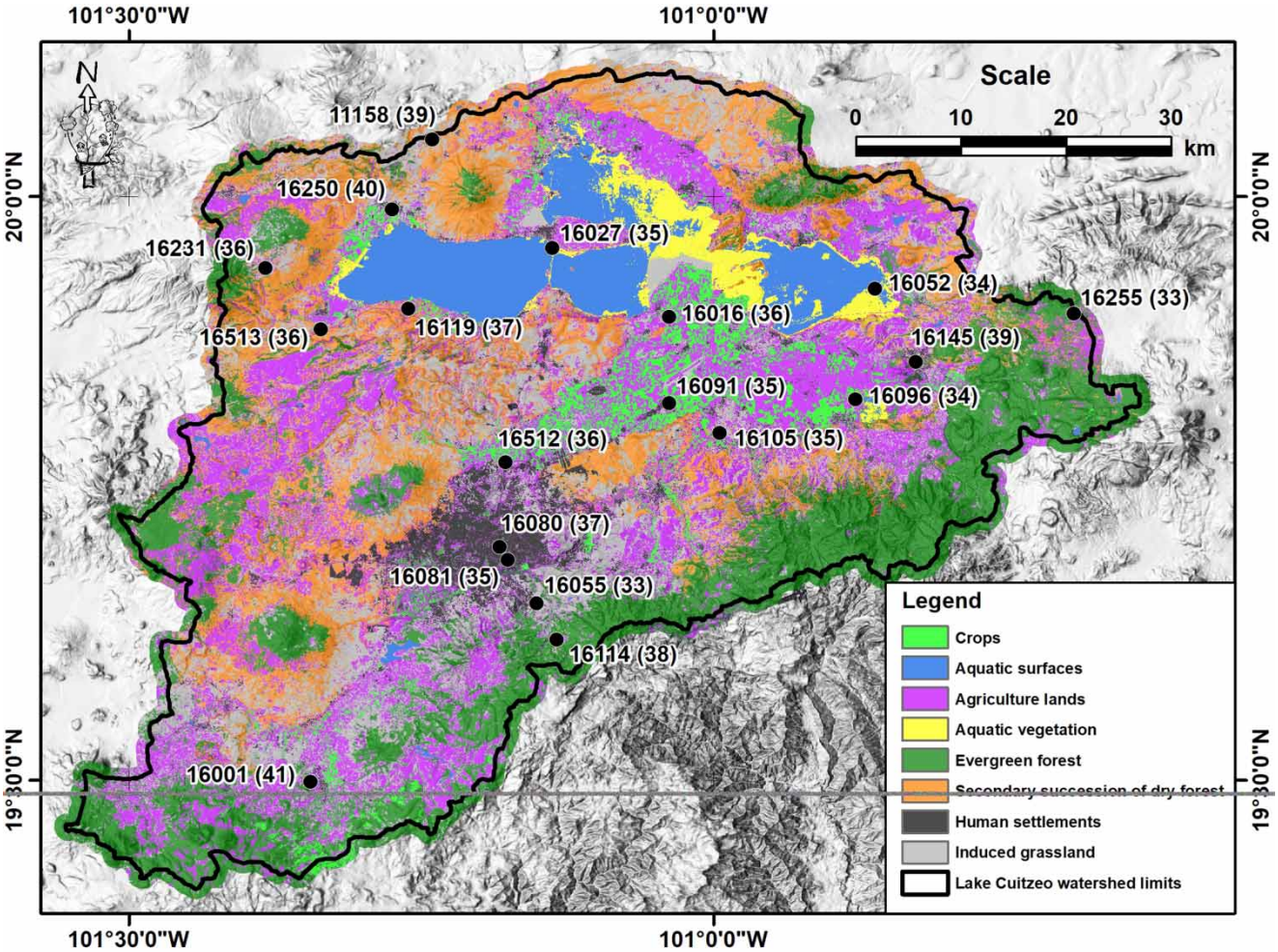


Figure 3 | Main land uses and land covers of the Lake Cuitzeo watershed obtained through unsupervised classification of Landsat OLI image (2019). Black dots represent the meteorological stations used to estimate the P value for a typical storm event. The values in parentheses are the arithmetic mean of precipitation (mm).

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

the country and classified as the most important lake area in the state of Michoacán (Mendoza *et al.* 2011). The climate is temperate with summer rains; the mean annual temperatures oscillate between 14 and 18 °C and the mean annual precipitation is about 800 mm (Correa-Ayram *et al.* 2014).

This watershed has intensive management, with agriculture and induced grassland occupying almost 50% of the study area, mainly distributed in the lower parts of the watershed. The evergreen forest is the most important natural cover (20.2%) and is distributed in the upper part of the watershed, mainly to the south and southeastern watershed. Also, secondary succession from dry forest clear-cutting is important, covering almost 17.5% of the watershed area. The hydrologic soil groups C and D occupied about 42 and 40%, respectively, while most of the watershed has a high runoff potential, with CN values oscillating between 60 and 98. Considering the AMC II and 1.4 inches as the precipitation value for a storm event, the runoff estimated in the study area is about 41.24×10^6 m^3 which, compared with those obtained using the SARA v1.0 tool (41.23×10^6 m^3), are practically the same (again, with differences <1%). The mean areal rainfall obtained through the interpolation using the IDW algorithm was 1.41 inches (±0.039) and produced a surface runoff volume of 41.78×10^6 m^3. This is almost the same as a single storm value. These values depict the homogeneous distribution of the precipitation along the watershed.

Regarding the performance measures, these are presented in Figure 4. The modeled sub-watershed has an area = 185 km^2, lower than 250 km^2, the recommended catchment size to use the NRCS-CN method. Figure 4(a) shows the relationship between simulated precipitation values and surface runoff, while the differences between the estimated surface runoff using sara4r and SARA v1.0 are presented in Figure 4(b). For all simulated cases, the runoff volumes are practically the same (<1%). After 41 runs, our package takes less than 10 seconds each (mean = 5 s) to perform the same as SARA v1.0 (mean = 87 s).

DISCUSSION

The NRCS-CN method was originally released in 1986 and since then, it has been widely applied for the estimation of surface runoff resulting from storm events occurring in small watersheds. Because of the simple structure and clearly stated assumptions, it soon became one of the most popular techniques among the engineers and the practitioners who adopted it for its use in a variety of regions, land uses and climate conditions (Chen *et al.* 2017). However, Ponce & Hawkins (1996) suggested that this model should be used cautiously in watersheds larger than 250 km^2 due to the risk of producing biased runoff estimations. Nevertheless, with some adaptations, the method was found suitable for the study of small- to large-scale watersheds, with areas ranging from 0.33 to 10,336 km^2 (D'Asaro & Grillone 2012; Lian *et al.* 2020). The watershed analyzed in this study falls within this scale range (~4,000 km^2). According to Lian *et al.* (2020), the main limitation of their study was the ignorance of the antecedent moisture condition. To overcome this limitation, our package includes options to made runoff estimations considering the three AMC as well as a modified runoff equation that incorporates the initial abstraction ratio $\lambda = Ia/S = 0.05$. This value of $\lambda = 0.05$ is recommended for agency use (Hawkins *et al.* 2010).

Regarding the utilization of the NRCS-CN method, Verma *et al.* (2017) reviewed the methodology and provided an updated list of modified CN-based hydrological models. Hawkins *et al.* (2019) list some of the more useful continuous watershed models based on curve number hydrology (e.g., SWAT, AnnAGNPS, APEX, among others). Many of them are open source software, but they require programming experience otherwise the possibility to modify or adapt the NRCS-CN method is very limited. Stanić *et al.* (2018) recently included the NRCS-CN method into the 3DNet platform to initially estimate the surface runoff for hydrological modeling. However, they argued that the structure of the 3DNet is pre-set and, differently from other GIS software, cannot be altered by the user. Although 3DNet is a novel distributed hydrologic model developed for continuous hydrologic simulations, its limitations with regard to making modifications to the CN method are overcome with our package.

In the past, similar tools that integrate the NRCS-CN method were developed for GIS platforms. These tools include ArcCN-Runoff (Zhan & Huang 2004), NRCS GeoHydro (Merkel *et al.* 2008), CN-Idris (Hernández-Guzmán *et al.*

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

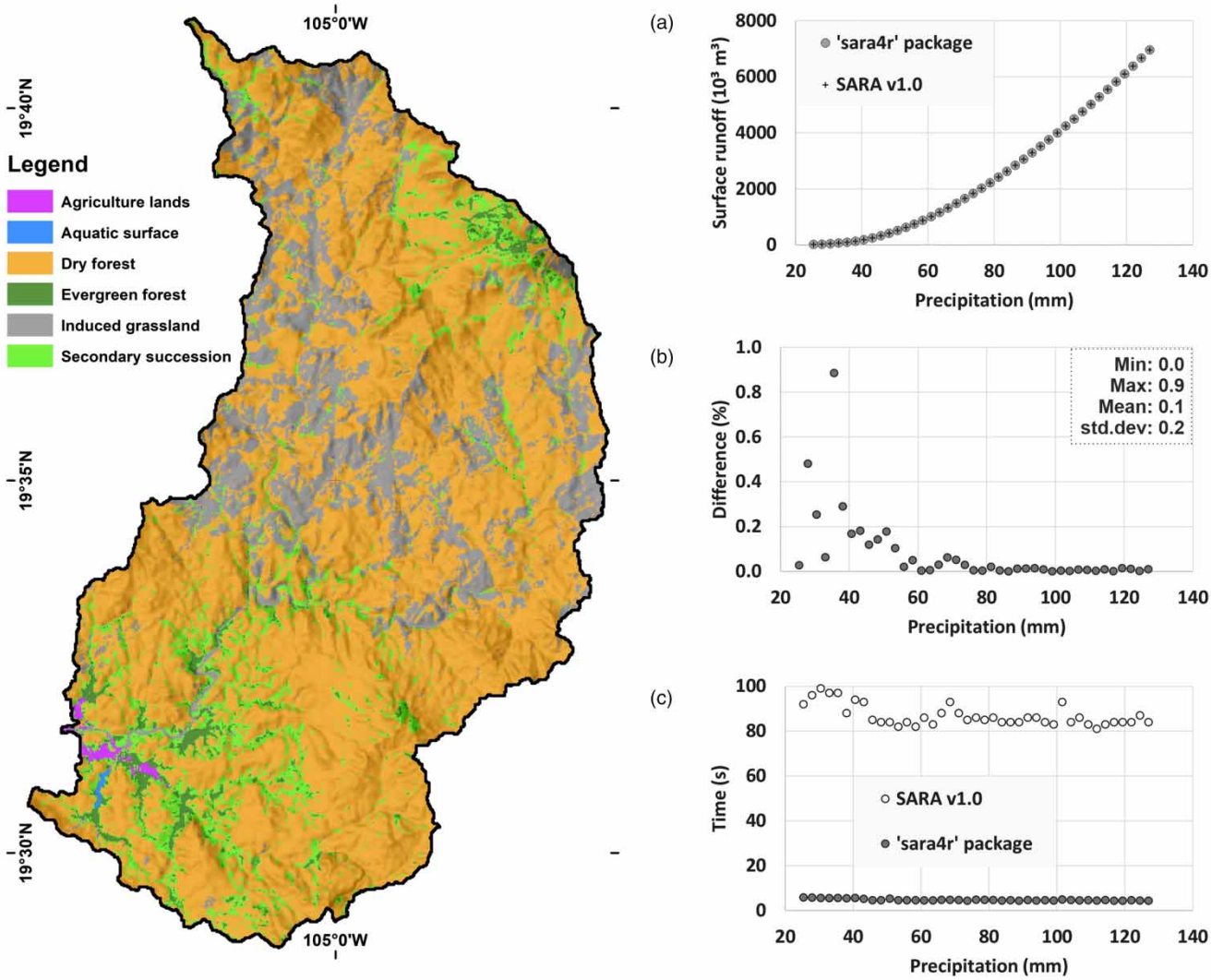


Figure 4 | (a) The relationship between simulated precipitation values and surface runoff in a modeled sub-watershed (area = 185 km^2) inside the study area. (b) Differences between the estimated surface runoff using sara4r and SARA v1.0. (c) Comparison of processing times using sara4r (raster based) and SARA v1.0 (vector based) on Intel® Core™ i7-7700HQ CPU 2.8 GHz with 16 GB RAM.

2011), and SARA v1.0 (Hernández-Guzmán & Ruiz-Luna 2013). However, these tools were developed to be used with commercial software, requiring a license for their use. Moreover, the programming language used in earlier projects (Visual Basic 6.0) has become obsolete. The '*Sara4r*' package, the same as previous developments by the authors of this article, aimed to use the NRCS-CN method to assess runoff at the watershed level, using different AMC and initial abstraction configurations, making possible the adaptation to different land use and climate conditions. In addition, it offers more calculation speed, can be upgraded based on other routines available on the cloud (e.g. parallel computing) and does not require a commercial software license.

Recently, two groups of researchers developed a global HSG (Ross *et al.* 2018) and CN maps for hydrologic modeling (Jaafar *et al.* 2019) at 250 m spatial resolution. The latter include the R script used for generating the datasets. Since our newly developed interface works in a raster-based environment within the R software, it will allow the easy inclusion of these datasets. Regarding this, one option to

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

assign CN values is using a comma-delimited file (*.csv); thus, users can select their own CN values rather than constraining them to use those in the handbook tables (USDA 1986).

Finally, the philosophy behind '*Sara4r*' is to provide an open-source software for surface runoff estimation. Its convenience lies in the fact that the original code of the GUI is free and easily accessible, and may be modified as needed.

FUTURE RESEARCH AND OPPORTUNITIES

The interface was developed in a raster-based environment, allowing the inclusion of LULC data, fundamental to define runoff rates. The use of raster images allows the calculation of a unique runoff value by pixel, making this process faster than that produced using vector data. However, in addition to the LULC and soil characteristics, we consider that many other physical aspects from watersheds need to be included, to ensure a best analysis. Considering this, it would be important that future improvements of our package include: (1) correction of effects derived from physiographic characteristics (e.g. slope); (2) inclusion of the effects of spatial and temporal variability in rainfall over large areas; (3) soil moisture conditions (i.e. modifications to AMC conditions reported in the literature); (4) variations in the λ parameter value; and (5) use of long-term climate data for a comprehensive hydrologic impact analysis.

CONCLUSION

The cases examined in this study demonstrated the applicability of the '*Sara4r*' package to analyze the surface runoff in medium size watersheds for a storm value, but open the necessity to develop routines to include a spatially interpolated precipitation map as input to improve the model performance. Their integration into the R software increases the opportunities for the '*Sara4r*' package to be used on a wide variety of platforms in which R and 'RGtk2' can be installed. The availability of the '*Sara4r*' package in the Comprehensive R Archive Network (CRAN) repository facilitates its use and testing by those interested in hydrological analyses at the watershed level, promoting the possibility of evolving to a better and more robust tool, throughout further communications among user and developers, to fix bugs or to add potential improvements. This is very important, considering that previous experience on R is not necessary, and only requires feedback from users, in both directions, regarding the runoff assessment outputs and the performance of the package.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers for helpful comments on an earlier version of the manuscript. This work was supported by the project Catedras – Mexican National Council for Science and Technology (CONACYT; Project No. 148). Landsat imagery was obtained from the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center. Author contributions: R.H. and A.R. designed the research; R.H. developed the software; R.H. and A.R. analyzed the data; R.H., A.R. and E.M. wrote the paper.

SOFTWARE AVAILABILITY

Description: Package *sara4r*
Developer: Rafael Hernández-Guzmán
Year first available: 2020
Email: rhernandez.g@gmail.com
Website: https://cran.r-project.org/web/packages/sara4r/index.html; http://hydro-geomatic-lab.com/sara4r.html
Hardware Requirement: General-purpose computer.
Software Requirement: Tested in R version 3.4 and later.
Availability and cost: This software is freely available.
Programming Language: R.
Program size: 2.84 MB

LICENSING

This software is freely available under the terms and conditions of the GNU General Public License.

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Abuzied, S. M. & Mansour, B. M. H. 2019 Geospatial hazard modeling for the delineation of flash flood-prone zones in Wadi Dahab basin, Egypt. *Journal of Hydroinformatics* **21** (1), 180–206. doi: 10.2166/hydro.2018.043.

Beg, M. N. A., Leandro, J., Bhola, P., Willems, W., Carvalho, R. F. & Disse, M. 2019 Discharge Interval method for uncertain flood forecasts using a flood model chain: city of Kulmbach. *Journal of Hydroinformatics* **21** (5), 925–944. doi: 10.2166/hydro.2019.131.

Chen, W., He, B., Ma, J. & Wang, C. 2017 A WebGIS-based flood control management system for small reservoirs: a case study in the lower reaches of the Yangtze River. *Journal of Hydroinformatics* **19** (2), 299–314. doi: 10.2166/hydro.2016.049.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V. & Böhner, J. 2015 System for automated geoscientific analyses (SAGA) v. 2.1.4. *Geoscientific Model Development* **8** (7), 1991–2007. doi: 10.5194/gmd-8-1991-2015.

Correa-Ayram, C. A., Mendoza, M. E., Pérez-Salicrup, D. R. & López-Granados, E. 2014 Identifying potential conservation areas in the Cuitzeo Lake basin, Mexico by multitemporal analysis of landscape connectivity. *Journal for Nature Conservation* **22** (5), 424–435. doi: 10.1016/j.jnc.2014.03.010.

D'Asaro, F. & Grillone, G. 2012 Empirical investigation of curve number method parameters in the Mediterranean area. *Journal of Hydrologic Engineering* **17** (10), 1141–1152. doi: 10.1061/(ASCE)HE.1943-5584.0000570.

Essaid, H. I. & Caldwell, R. R. 2017 Evaluating the impact of irrigation on surface water – groundwater interaction and stream temperature in an agricultural watershed. *Science of the Total Environment* **599–600**, 581–596. doi: 10.1016/j.scitotenv.2017.04.205.

FAO 2014 *World Reference Base for Soil Resources 2014, Update 2015*. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome. Available from: http://www.fao.org/3/i3794en/I3794en.pdf (accessed 11 May 2020)

Fernández-Pato, J., Gracia, J. L. & García-Navarro, P. 2018 A fractional-order infiltration model to improve the simulation of rainfall/runoff in combination with a 2D shallow water model. *Journal of Hydroinformatics* **20** (4), 898–916. doi: 10.2166/hydro.2018.145.

Feyereisen, G. W., Strickland, T. C., Bosch, D. D., Truman, C. C., Sheridan, J. M. & Potter, T. L. 2008 Curve number estimates for conventional and conservation tillages in the southeastern Coastal Plain. *Journal of Soil and Water Conservation* **63** (3), 12–128. doi: 10.2489/jswc.63.3.120.

Flores-Casas, R. & Ortega-Huerta, M. A. 2019 Modelling land cover changes in the tropical dry forest surrounding the Chamela-Cuixmala biosphere reserve, Mexico. *International Journal of Remote Sensing* **40** (18), 6948–6974. doi: 10.1080/01431161.2019.1597305.

Flores-Díaz, A. C., Guevara-Hernández, R., Mendoza, M. E., Langrave, R., Quevedo, A. & Maass, M. 2018 Hierarchical procedure for creating local typologies for riparian zone research and management based on biophysical features. *Physical Geography* **39** (2), 118–139. doi: 10.1080/02723646.2017.1387427.

Hawkins, R. H., Jiang, R., Woodward, D. E., Hjelmfelt, A. T. & VanMullem, J. E. 2002 Runoff curve number method: examination of the initial abstraction ratio. In *Proceedings of the Second Federal Interagency Hydrologic Modeling Conference, Las Vegas, Nevada*. U.S. Geological Survey, Lakewood, Colorado. Available from: https://acwi.gov/hydrology/mtsconfwkshops/conf_proceedings/second_fihmc_nevada.pdf (accessed 11 May 2020)

Hawkins, R. H., Theurer, F. D. & Rezaeianzadeh, M. 2019 Understanding the basis of the curve number method for watershed models and TMDLs. *Journal of Hydrologic Engineering* **24** (7), 06019003. doi: 10.1061/(ASCE)HE.1943-5584.0001755.

Hawkins, R. H., Ward, T. J., Woodward, E. & Van Mullem, J. A. 2010 *Continuing Evolution of Rainfall Runoff and the Curve Number Precedent 2nd Joint Federal Interagency Conference (Las Vegas, NV) 1–12*. Available from: https://acwi.gov/sos/pubs/2ndJFIC/Contents/10E_Hawkins.pdf (accessed 3 August 2020)

Hernández-Guzmán, R. & Ruiz-Luna, A. 2013 SARA – an enhanced curve number-based tool for estimating direct runoff. *Journal of Hydroinformatics* **15** (3), 881–887. doi: 10.2166/hydro.2013.145.

Hernández-Guzmán, R., Ruiz-Luna, A. & Berlanga-Robles, C. A. 2011 CN-Idris: An Idrisi tool for generating curve number maps and estimating direct runoff. *Environmental Modelling & Software* **26** (12), 1764–1766. doi: 10.1016/j.envsoft.2011.07.006.

Hernández-Guzmán, R., Ruiz-Luna, A. & González, C. 2019 Assessing and modeling the impact of land use and changes in land cover related to carbon storage in a western basin in Mexico. *Remote Sensing Applications: Society and Environment* **13**, 318–327. doi: 10.1016/j.rsase.2018.12.005.

Hesselbarth, M. H. K., Sciaini, M., With, K. A., Wiegand, K. & Nowosad, J. 2019 *Landscapemetrics*: an open-source R tool to calculate landscape metrics. *Ecography* **42** (10), 1648–1657. doi: 10.1111/ecog.04617.

Hijmans, R. J., van Etten, J., Summer, M., Cheng, J., Baston, D., Bevan, A., Bivand, R., Busetto, L., Canty, M., Forrest, D., Ghosh, A., Golicher, D., Gray, J., Greenberg, J. A., Hiemstra, P., Hingee, K., IMAGe Karney, C., Mattiuzzi, M., Mosher, S., Nowosad, J., Pebesma, E., Lamigueiro, O. P., Racine, E. B., Rowlingson, B., Shortridge, A., Venables, B. & Wueest, R.

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest

2020 *Raster: Geographic Data Analysis and Modeling*. R package version 2, 9–5. Available from: http://CRAN.R-project.org/package=raster.

Jaafar, H. H., Ahmad, F. A. & El Beyrouthy, N. 2019 GCN250, new global gridded curve numbers for hydrologic modeling and design. *Scientific Data* **6**, 145. doi: 10.1038/s41597-019-0155-x.

Jiang, J., Zhu, A. X., Qin, C. Z. & Liu, J. 2019 A knowledge-based method for the automatic determination of hydrological model structures. *Journal of Hydroinformatics* **21** (6), 1163–1178. doi: 10.2166/hydro.2019.029.

Lawrence, M. & Lang, D. T. 2019 *RGtk2: Bindings for Gtk 2.8.0 and Above*. Available from: https://CRAN.R-project.org/package=RGtk2

Lazos-Chavero, E., Mwampamba, T. H. & García-Frapolli, E. 2018 Uncovering links between livelihoods, land-use practices, vulnerability and forests after hurricane Jova in Jalisco, Mexico. *Forest Ecology and Management* **426**, 27–38. doi: 10.1016/j.foreco.2017.10.009.

Ling, L., Yusop, Z., Yap, W.-S., Tan, W. L., Chow, M. F. & Ling, J. L. 2020 A calibrated, watershed-specific SCS-CN method: application to Wangjiaqiao watershed in the Three Gorges area, China. *Water* **12**, 60. doi: 10.3390/w12010060.

Lian, H., Yen, H., Huang, J.-C., Feng, Q., Qin, L., Bashir, M. A., Wu, S., Zhu, A.-X., Luo, J., Di, H., Lei, Q. & Liu, H. 2020 CN-China: Revised runoff curve number by using rainfall-runoff events in China. *Water Research* **177**, 115767. doi: 10.1016/j.watres.2020.115767.

Mendoza, M. E., López-Granados, E., Geneletti, D. & Pérez-Salicrup, D. R. 2011 Analysing land cover and land use change processes at watershed level: a multitemporal study in the Lake Cuitzeo Watershed, Mexico (1975–2003). *Applied Geography* **31** (1), 237–250. doi: 10.1016/j.apgeog.2010.05.010.

Merkel, W. H., Kaaushika, R. M. & Gorman, E. 2008 NRCS geohydro – a GIS interface for hydrologic modeling. *Computer & Geociences* **34** (8), 918–930. doi: 10.1016/j.cageo.2007.05.020.

Mishra, S. K., Jain, M. K., Suresh-Babu, P., Venugopal, K. & Kaliappan, S. 2008 Comparison of AMC-dependent CN-conversion formulae. *Water Resources Management* **22**, 1409–1420. doi: 10.1007/s11269-007-9233-5.

Ponce, V. M. & Hawkins, R. H. 1996 Runoff curve number: has it reached maturity? *Journal of Hydrologic Engineering, ASCE* **1** (1), 11–19. doi: 10.1061/(ASCE)1084-0699(1996)1:1(11).

Reichl, F. & Hack, J. 2017 Derivation of flow duration curves to estimate hydropower generation potential in data-scarce regions. *Water* **9**, 572. doi: 10.3390/w9080572.

Ross, C. W., Prihodko, L., Anchang, J., Kumar, S., Ji, W. & Hanan, N. P. 2018 HYSOGs250 m, global gridded hydrologic soil groups for curve-number-based runoff modeling. *Scientific Data* **5**, 180091. doi: 10.1038/sdata.2018.91.

Soulis, K. X. & Valiantzas, J. D. 2012 SCS-CN parameter determination using rainfall-runoff data in heterogeneous watersheds – the two-CN system approach. *Hydrology and Earth System Sciences* **16**, 1001–1015. doi: 10.5194/hess-16-1001-2012.

Stanić, M., Todorović, A., Vasilić, Z. & Plavšić, J. 2018 Extreme flood reconstruction by using the 3DNet platform for hydrological modelling. *Journal of Hydroinformatics* **20** (4), 766–783. doi: 10.2166/hydro.2017.050.

Suazo-Ortuño, I., Benítez-Malvido, J., Marroquín-Páramo, J., Soto, Y., Siliceo, H. & Alvarado-Díaz, J. 2018 Resilience and vulnerability of herpetofaunal functional groups to natural and human disturbances in a tropical dry forest. *Forest Ecology and Management* **426**, 145–157. doi: 10.1016/j.foreco.2017.09.041.

USDA Soil Conservation Service 1986 Urban hydrology for small watersheds. United States Department of Agriculture. In: *Natural Resources Conservation Service. Conservation Engineering Division* (R. Cronshey, R. H. McCuen, N. Miller, W. Rawls, S. Robbins & D. Woodward, eds). Technical Release 55. 2nd edn. Washington, DC, p. 164.

Verma, S., Verma, R. K., Mishra, S. K., Singh, A. & Jayaraj, G. K. 2017 A revisit of NRCS-CN inspired models coupled with RS and GIS for runoff estimation. *Hydrological Sciences Journal* **62** (12), 1891–1930. doi: 10.1080/02626667.2017.1334166.

Verma, S., Singh, P. K., Mishra, S. K., Jain, S. K., Berndtsson, R., Singh, A. & Verma, R. K. 2018 Simplified SMA-inspired 1-parameter SCS-CN model for runoff estimation. *Arabian Journal of Geosciences* **11**, 420. doi: 10.1007/s12517-018-3736-7.

Walega, A., Amatya, D. M., Caldwell, P., Marion, D. & Panda, S. 2020 Assessment of storm direct runoff and peak flow rates using improved SCS-CN models for selected forested watersheds in the Southeastern United States. *Journal of Hydrology: Regional Studies* **27**, 100645. doi: 10.1016/j.ejrh.2019.100645.

Zhan, X. & Huang, M. L. 2004 ArcCN-Runoff: an ArcGIS tool for generating curve number and runoff maps. *Environmental Modelling & Software* **19** (10), 875–879. doi: 10.1016/j.envsoft.2004.03.001.

First received 11 May 2020; accepted in revised form 19 October 2020. Available online 3 November 2020

Downloaded from http://iwaponline.com/jh/article-pdf/23/1/76/870208/jh0230076.pdf
by guest