Review of the paper NHESS-2015-157

General comment

This contribution to the identification of rainfall that cause flash floods in a mountainous region of Colombia could be potentially interesting for NHESS but need substantial revision in data analysis, presentation of the results and structure of the paper.
Specific comments refer to the issues that, in my opinion, deserve revision.

Specific comments

Abstract

The abstract would become more readable if the acronyms (CV, CC, DesIventar, AR, AAR) were removed. The seasonal distribution of floods is not a relevant scientific result and does not deserve to be mentioned here.

2 Methodology

2.1 Description of the study area and data

This section, in addition to describing the study area, presents an analysis of the changes of precipitation-related variables in the interval 1980-2012. The outcomes of this analysis would better reported in a results section.
The description of the study area and the analysis of variations of rainfall from 1980 to 2010 are not part of methodology (general title of section 2). A separate section could be proposed for the description of study area.
The authors do not provide links between the analysis of rainfall variations in the most recent decades and the derivation of rainfall thresholds for flash flood occurrence, and do not discuss possible links between rainfall variations and flash flood occurrence. One could expect that that, due to increase in rainfall amounts and intensity occurred in the last decades (Table 1), also the number of floods has increased. Actually, table 2 do not show such an increase (see Fig. 1 of this review).
Since the main focus of this paper is the derivation of rainfall thresholds for flash floods, the authors could consider expunging the section on variations of rainfall properties.
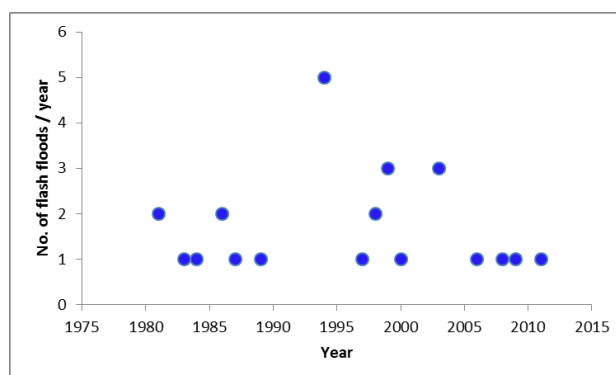

Fig. 1. No. of flash floods per year from table 2 of NHESS-2015-157.

Page 4099, lines 15-16
"In order to explore temporal characteristics of precipitation on local scale, (Fig. 1 and table 1)."
Something is missing in this sentence.

Page 4100, lines 10-11
"where the hydrologic characteristics…": not clear

Page 4100, line 19
"For this reason…" Do the authors mean that Cali River was selected because hydrometeorological data are available?
Does the "Gauge station" shown in Fig. 2 correspond to the streamgauge of Bocatoma? I would suggest using uniform terminology (streamgauge or limnigraphic station). What is the drainage area upstream of Bocatoma station (the area of 215 km^2 reported at page 4100, line 23 seems to refer to the entire Cali River basin)? The values of peak discharge reported in table 2 are of small significance unless they are associated to the basin area.

Page 4101, lines 11-15
Nineteen raingauges include also several stations outside the catchment upstream of the Bocatoma streamgauge. Since the selection of floods has been done based on discharge data from this station (page 4101, lines 5-10), it would better to analyse only rainfall data recorded within this catchment. Due to the strong gradient of rainfall the cause flash floods, rainfall data recorded outside the catchment could be not representative of the flash floods that were observed at the streamgauge station.

Page 4101, lines 9-10: "The event must be seasonal, continuous and accurate". This sentence does not clarifies the criteria adopted for event selection.

2.2 Definition of critical rain thresholds

The paper by Guzzetti et al. (2008) deals with critical rainfall thresholds for landslides, not for flash floods. Other studies cited in this section also focus on landslides and debris flows: their extension to flash floods is far from being straightforward. Since flash floods and landslides / debris flows are very different processes, citing these papers as previous studies on rainfall thresholds for flash flood occurrence is not acceptable.

How data from 19 raingauges were processes? Were individual rainfall data from each raingauge analysed? Or the average value over the catchment resulting from the interpolation of rainfall height recorded at each raingauge was used?

The right part of Fig. 2 should be briefly described as it can be unclear to international readers.

3 Analysis and results

I suggest considering always the same time interval (1980-2012) in data analysis: this would ensure homogeneity in the assessment of seasonal distribution of rainfall and flash flood and in derivation of rainfall thresholds. Although data from DesInventar cover a longer time interval (e.g., Fig. 3 and text from 4102 line 22 to 4103, line 9), data before 1980 do not contribute to the definition of rainfall thresholds, which is the main topic of this paper.

The description of the seasonal distribution of flash floods (page 4103, lines 10-15) could be shortened: it would be sufficient to say that most events take place from April to June. The seasonal distribution of selected events partly disagrees from that of Fig. 3 (as an example, the frequency of floods in June is only slightly lower than in May in table 2, whereas Fig. 3 indicates a marked decrease in the number of floods from May to June).

Page 4103, line 15: "Of the total number of events found…" More details are required on event selection (e.g.: how many floods were in the record?). See also my comment on event selection in the methodology section (Page 4101, lines 9-10).

Minimum return period is 3 years and occurs 14 times in a time series covering 30 years. A comment on the return period of the studied floods would be very useful: how they were estimated, how long is the time series considered, etc..

4104, lines 6-7.
The values reported in the upper right corner of figs. 4-7 represent the probability of detection (POD) of flash floods. This is an important parameter, but it is not sufficient to evaluate the performance of the proposed thresholds. The false alarm rate (FAR), i.e. the number of rainstorm that lie above the threshold but did not cause flash floods should also be reported. POD and FAR can be combined to compute the critical success rate (CSI), a more comprehensive index of performance (Schaefer, 1990).

Have the authors considered only rainfall (AR and AAR) related to rainstorms that caused flash floods or the analysis involved also rainstorms that did not trigger flash floods?

How the thresholds were defined? In a merely empirical way or by means of a statistical procedure? The second option would reduce subjectivity and favour extension of the approach to other areas.

Reference

Schaefer, J.T., 1990. The critical success index as an indicator of warning skill. Weather and Forecast, 5, 570-575.