ORIGINAL PAPER

An improved rainfall-threshold approach for robust prediction and warning of flood and flash flood hazards

Geraldo Moura Ramos Filho[1] · Victor Hugo Rabelo Coelho[2] ·
Emerson da Silva Freitas[1] · Yunqing Xuan[3] · Cristiano das Neves Almeida[1]

Abstract

This paper presents an improved method of using threshold of peak rainfall intensity for robust flood/flash flood evaluation and warnings in the state of São Paulo, Brazil. The improvements involve the use of two tolerance levels and the delineating of an intermediate threshold by incorporating an exponential curve that relates rainfall intensity and Antecedent Precipitation Index (API). The application of the tolerance levels presents an average increase of 14% in the Probability of Detection (POD) of flood and flash flood occurrences above the upper threshold. Moreover, a considerable exclusion (63%) of non-occurrences of floods and flash floods in between the two thresholds significantly reduce the number of false alarms. The intermediate threshold using the exponential curves also exhibits improvements for almost all time steps of both hydrological hazards, with the best results found for floods correlating 8-h peak intensity and 8 days API, with POD and Positive Predictive Value (PPV) values equal to 81% and 82%, respectively. This study provides strong indications that the new proposed rainfall threshold-based approach can help reduce the uncertainties in predicting the occurrences of floods and flash floods.

Keywords Flood hazards warning · Rainfall threshold method · Peak rainfall intensity · Antecedent precipitation index · Tolerance limits · Intermediate rainfall intensity threshold

✉ Geraldo Moura Ramos Filho
  geraldo0511@hotmail.com

✉ Yunqing Xuan
  y.xuan@swansea.ac.uk

1 Department of Civil and Environmental Engineering, Federal University of Paraíba, João Pessoa 58051-900, Brazil

2 Department of Geosciences, Federal University of Paraíba, João Pessoa 58051-900, Brazil

3 College of Engineering, Swansea University Bay Campus, Swansea SA1 8EN, UK

1 Introduction

Rainfall is an intermittent phenomenon with irregular spatiotemporal distribution and is able to cause many natural disasters (Dunkerley 2008). According to the report commissioned by the United Nations Office for Disaster Risk Reduction (UNDRR) in partnership with the Centre for Research on Epidemiology of Disasters (CRED), the last two decades (1998–2017) represent the largest number of records caused by hydrological disasters in history (Wallemacq and House 2018). The same source reported that this type of natural disaster was mainly induced by floods (accounting for 43.4% of all natural disasters) that affected more than 2 billion people in almost all countries in the world, causing more than 142,000 deaths and US$ 650 million of economic losses. Furthermore, other hydrological disasters—including storms (28.2%), landslides (5.2%), and droughts (4.8%)—accounted for more than 38% of the natural disasters that occurred between 1998 and 2017. The number of occurrences has been growing quickly mainly because of extreme weather conditions, high urbanisation rate, and inadequate response to disasters (Špitalar et al. 2014; Tsakiris 2014; Du et al. 2015). For instance, hydrological disasters induced by extreme rainfall events account for 87% of the deaths caused by natural disasters between 1991 and 2012 in Brazil (CEPED 2013).

During the last few decades, great efforts have been made for predicting and warning hydrological disasters. Complex computer models such as hydrological models have been a new challenge for many researchers and authorities to build forecasting and early warning systems (e.g. Azari et al. 2008; González-Cao et al. 2019; Li et al. 2019). While such practice remains as the mainstream approach, there are situations where empirical methods still prevail. This is because, in many places, it is impossible to carry out detailed modelling of the related physical process (e.g. landslides, floods, flash floods) either due to data availability or being too challenging to model. Among these empirical methods, the rainfall-threshold method is one of the most widely used for predicting and warning some of the hydrological disasters (Huang et al. 2015), where the thresholds are determined by the properties derived from rainfall events such as intensity, duration, and antecedent precipitation (e.g. Glade et al. 2000; Aleotti 2004; Berti et al. 2012; Papagiannaki et al. 2015; Scheevel et al. 2017; Brunetti et al. 2018; Mirus et al. 2018). One of the first studies using rainfall events to delineate thresholds was due to Caine (1980), who used rainfall peak intensity with different durations for forecasting the occurrence of rainfall-triggered landslides and debris flow in various regions of the world. The empirical methods based on rainfall thresholds have been gradually applied more for landslide warning purposes than for flooding or flash flooding (Diakakis 2012; Papagiannaki et al. 2015; Santos and Fragoso 2016). In the area of flood warning, Diakakis (2012) used rainfall intensity-duration parameters to determine the thresholds after adapting the methodologies proposed by Cannon et al. (2008) and Guzzetti et al. (2008) for landslides. In his study, two thresholds (upper and lower) are defined but large uncertainties still exist, mainly manifested by the considerable number of occurrences and non-occurrences of flooding concentrated in the region between the two thresholds. Consequently, many warning systems are frequently neglected by the community due to the large number of false alarms (Abon et al. 2012). Thus, reducing such uncertainties is crucial to minimise the costs and improve the decision-making processes (Villarini et al. 2010).

The uncertainties of the rainfall thresholds are inevitable as rainfall is not the only factor that triggers flooding and flash flooding events (Papagiannaki et al. 2015). The shortcomings of the intensity-duration thresholds are frequently mentioned in the literature, although

this type of method remains as the most widespread one in the world (Zhao et al. 2019). For instance, choosing rainfall events with short durations to build the threshold would exclude the important antecedent wetness information, whereas selecting rainfall events with long durations can mitigate this but it would also flatten the peak intensity that otherwise can be the real trigger of floods (Bogaard and Greco 2018). Other methods attempt to overcome this by using information of antecedent rainfall (e.g. Chleborad et al. 2008; Lee and Park 2016; Scheevel et al. 2017) or Antecedent Precipitation Index (API) (e.g. Glade et al. 2000; Mirus et al. 2018; Suribabu and Sujatha 2019; Zhao et al. 2019). The use of API, in contrast with the use of the antecedent rainfall, allows for the consideration of the loss of the rainfall over the past days (Suribabu and Sujatha 2019). In addition, some other studies also provided a quantitative assessment of the rainfall threshold approaches for landslides occurrences by applying probability-based methods (Berti et al. 2012). These probability-based methods allow for the definition of multiple rainfall thresholds based on different exceedance probability levels, which makes possible the establishment of various warning levels (Brunetti et al. 2010; Huang et al. 2015).

Still, when compared with the applications in predicting landslides, the number of studies using rainfall thresholds for flood and flash floods warning systems remains low and the area has been poorly explored. It is also clear that reducing the uncertainties in such applications is crucial for the effective issuance of flood warnings. The present study aims to create a rainfall threshold estimation approach for the robust prediction and warning of floods and flash floods hazards. The flood and flash flood warning system proposed in this study intends to reduce the uncertainties and to minimise false alarms observed in the region between the upper and lower thresholds by introducing an intermediate threshold derived by assessing the different interactions between the rainfall peak intensity and API, considering different evaluation metrics. Moreover, the novel inclusion of two tolerance levels in the upper and lower regions of the threshold enables a more fine-tuned flood warning level setting. The rest of the paper is structured as follows: the study area is characterised in Sect. 2; the data and the methodology are described in detail in Sect. 3; the evaluation of the results is given in Sect. 4; and the summary of the study, followed by a number of conclusive points with perspective of future studies, is presented in Sect. 5.

2 Study area

This study was carried out in São Paulo State, located in the Brazilian Southeast region with an area of 248,200 km^2 between 19°55′58″S–25°00′53″S and 50°32′15″W–47°55′36″W (Fig. 1). The state is highly urbanised with approximately 45.5 M inhabitants, reaching a level of urbanisation of 95% (IBGE 2018). The study area is divided into two zones with different physical characteristics: (1) the coastal zone which has an altitude lower than 300 m and (2) the plateau zone which comprises most of the area of the state with elevation ranging from 300 to 900 m. This topographical characteristic is an important natural factor in explaining the climate of the state of São Paulo (Setzer 1946). The coastal zone is dominated by the humid tropical climate with a mean annual temperature above 22 °C and average annual rainfall above 2000 mm. Meanwhile, the plateau zone is mainly characterised by the humid subtropical climate with an annual average temperature of 20 °C and average annual rainfall equal to 1400 mm year^{-1} (Alvares et al. 2013). The rainfall in both regions of the state is more concentrated during the austral summer, i.e. between October and March. Generally,
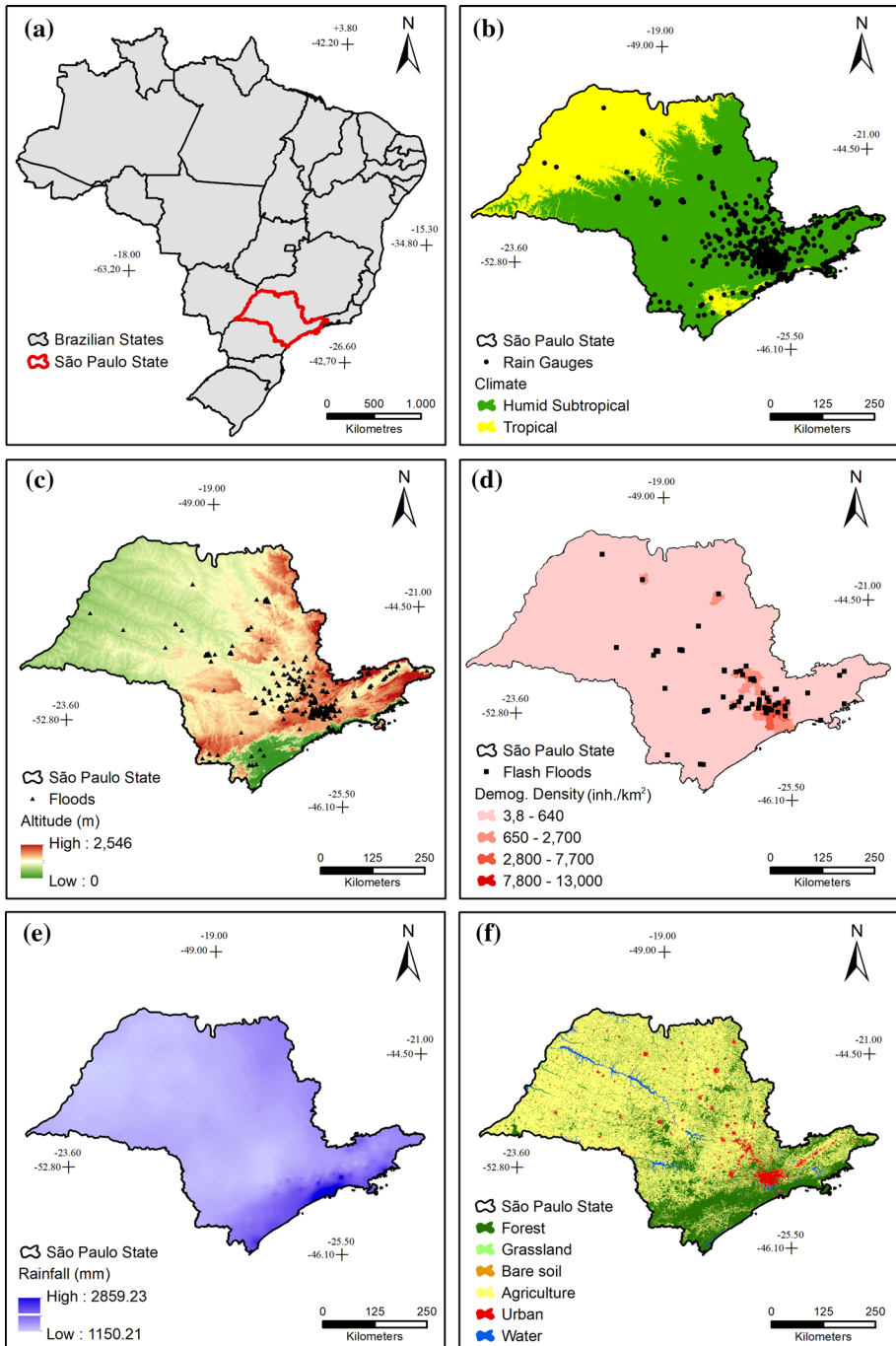
Fig. 1 **a** Map of Brazil showing the São Paulo State. **b** Rain gauges and Köppen's classification map for São Paulo State according to Alvares et al. (2013). **c** Elevation of the São Paulo State and location of the 347 flood occurrences. **d** Demographic density of São Paulo and location of the 71 flash flood occurrences. **e** Long-term (1950–1990) mean annual rainfall obtained from the meteorological stations used by Alvares et al. (2013). **f** Landsat-based land use and land cover map for 2017 provided by the MapBiomas Project (Souza et al. 2020)

April and September are the driest months in São Paulo State (Fig. 2). Approximately 70% of the study area is composed of Devonian-Cretaceous deposits of the Paraná and Bauru basins, while the remaining 30% mainly corresponds to a crystalline basement with rocks older than the Neoproterozoic Era (Garcia et al. 2018). Other sedimentary deposits (e.g. intercontinental and coastal Cenozoic basins) also compose the geology of the São Paulo State but at a small proportion.

São Paulo State is a typical hot spot frequented by landslides, floods, and soil erosion problems arising from prolonged or intense rainfall events. The occurrences of these disasters are due to the natural characteristics of the region associated with the high level of urbanisation (Tominaga et al. 2015). From 2000 to 2015, there have been more than 10,800 natural disasters recorded, causing 534 deaths and affecting approximately 971,500 people and 128,500 buildings. Out of all natural disasters recorded in São Paulo, more than 50% were caused by sudden and violent changes in the distribution or movement patterns of water (Brollo and Ferreira 2016). Moreover, São Paulo is the richest state in Brazil, with the largest number of floods and flash floods records as well as sub-daily rainfall data made available by public agencies.

3 Materials and methods

This study used a series of steps to create a robust rainfall threshold able to reduce the uncertainties of events triggering floods and flash flood occurrences, as shown in Fig. 3. Overall, the implementation of the proposed rainfall threshold approach include (a) the selection of events, (b) the application of rainfall intensity-duration parameters to define thresholds, (c) the adoption of tolerance levels to improve the rainfall intensity threshold, and (d) the implementation of an intermediate threshold relating rainfall peak intensity and API to better separate the flood and flash flood occurrences from the non-occurrences. These methodological steps are described in detail in the next items of this section.
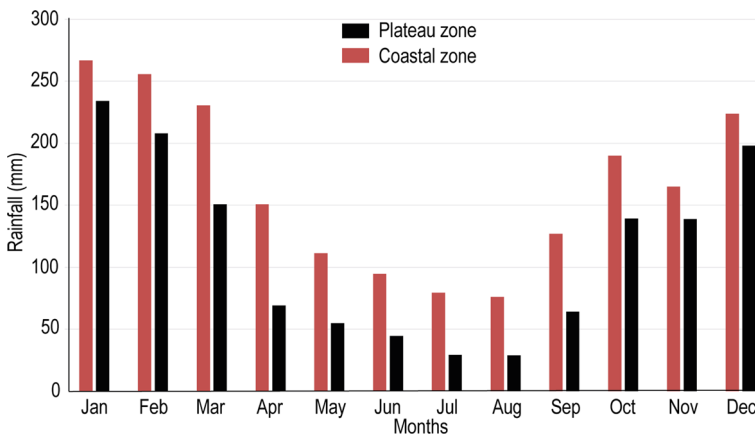

Fig. 2 Long-term (1950–1900) mean monthly rainfall for the coastal and plateau zones obtained from the meteorological stations used by Alvares et al. (2013)
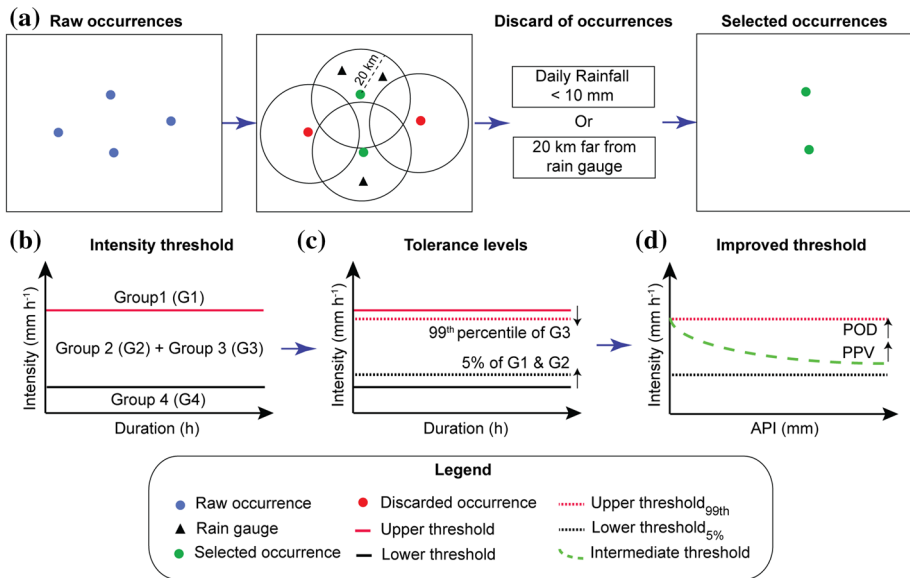
Fig. 3 Methodological chart, showing **a** the raw and selected flood and flash flood occurrences; **b** the rainfall intensity threshold approach; **c** the tolerance levels adopted to improve the rainfall intensity threshold; and **d** the improved threshold relating rainfall peak intensity and antecedent precipitation index (API)

3.1 Selection of events

3.1.1 Rainfall data

Rainfall data over the period of 1 January 2015—31 December 2017 were collected from the 732 rain gauges distributed throughout the São Paulo State. The rain gauges belong to the Brazilian National Centre for Monitoring Early Warning of Natural Disasters (CEMADEN, acronym in Portuguese), a national-wide network established by the Brazilian Government supporting the natural disasters risk management (Bacelar et al. 2020). The ground-based rainfall observation network of CEMADEN is equipped with tipping bucket gauges with a 10-min temporal resolution when it rains and 60-min temporal resolution over no-rain periods. These rainfall data were screened before use in this study. The quality-control procedure is as follows: first, a computational routine was created to select only rain gauges with less than 30 days of missing data along each of the three civil years considered in this study; then, all rain gauges meeting this first requirement were visually inspected using two standard methods, including (1) a comparison of monthly and sub-daily rainfall data of the five nearest stations was to verify large discrepancies between them; and (2) an analysis of the range of values and changes over subsequent measurements of each rain gauge to identify constant or null rainfall records that probably indicate gauge clogging. This resulted in the final 590 gauges that were selected for the whole study period (Fig. 1b). These data were then used to define the rainfall events and to calculate their respective thresholds. It is worth noting that not all 590 gauges were used every year because of the quality-check procedures adopted in this study, whereas (1) 216 rain gauges with high-quality data in 2015,

(2) 315 rain gauges with high-quality data in 2016, and (3) 355 rain gauges with high-quality data in 2017 were used.

Most rain gauges used to calculate the rainfall thresholds are located within the metropolitan regions of the São Paulo State, including (1) the Metropolitan Area of São Paulo (MASP), with an estimated density of 7689 inhabitants/km^2 and covered by 42 rain gauges, (2) the Metropolitan Area of Ubatuba (MAU), with an estimated density of 121 inhabitants/km^2 and covered by 18 rain gauges, and (3) the Metropolitan Area of Santo André (MASA), with an estimated density of 3919 inhabitants/km^2 and covered by 17 rain gauges.

3.1.2 Flood and flash flood data

Detailed information of flood and flash flood occurrences are fundamental for the analysis of their relationship with rainfall events. In this study, flood occurrences were considered as the overflow of water from a stream channel onto normally dry land in the floodplain, whereas flash flood occurrences were regarded as a rapid inland flood due to intense rainfall or a sudden flooding with short duration (Guha-Sapir et al. 2015). The inventory of these occurrences, which comprise the same period of the rainfall data, was obtained from three main sources: (1) The Integrated Storm Monitoring, Forecasting and Alerting System for the Brazilian South-Southeast Regions (SIMPAT, acronym in Portuguese), (2) The Civil Defence of the state of São Paulo, and (3) press news. Only occurrences confirmed in at least two sources of data were selected for this study. The data provided by the press news were also used to confirm and differentiate the type of occurrence (floods or flash floods) by analysing some available information such as pictures, rainfall duration, and location. In order to choose the most appropriate rain gauges, only those flood and flash occurrences that could be georeferenced (e.g. via address and coordinates) and dated were selected. This was followed by the application of two more criteria to further filter out the events/occurrences that (1) come with daily rainfall less than 10 mm near to the flood and flash flood events or (2) have the nearest rain gauge located more than 20 km from the occurrence. Although choosing only occurrences distant less than 20 km from the rain gauges as a criterion, almost 72% of the flood and flash flood events were located within 10 km from the stations.

3.1.3 Characterisation of rainfall events

In parallel with the selection of the flood and flash flood occurrences, rain gauges were chosen to define the rainfall events that better characterise the disasters. In this study, the relationship between rainfall peak intensities and the antecedent wetness conditions was assessed for the events that might or might not lead to the floods and flash floods. This assessment was performed to avoid two potential issues: (1) the inadvertent exclusion of important antecedent wetness information for rainfall events with short duration and (2) the flatness of peak intensity for rainfall events with long duration. The procedure is as follows: first, the rainfall peak intensity for all rain gauges was calculated for each day considering ten time steps (10 min, 30 min, 1 h, 2 h, 3 h, 6 h, 8 h, 10 h, 12 h, and 24 h). Thereafter, the API was tested for different time steps (1–10 days) to estimate the antecedent wetness conditions for the day before the rainfall event (Kohler and Linsley 1951), as in Eq. 1:

$$\text{API} = \sum_{t=-1}^{-i} P_t k^{-t} \tag{1}$$

where i is the number of antecedent days considered in the study, P_t is the rainfall for the day t (mm), and k is a decay rate that ranges from 0.80 to 0.98 according to Viessman and Lewis (1996). The values of API chosen in this study are within the ranging established by some well-recognised methods such as the Natural Resources Conservation Service (NRCS) method that uses 5 days of antecedent moisture condition (NRCS 1972). Some other studies also suggest values of API ranging from 2 to 6 days to characterise flooding (e.g. Tramblay et al. 2012; Froidevaux et al. 2015).

The selection of the rainfall events that better characterise the flood and flash occurrences followed largely the methodology proposed by Rossi et al. (2017), i.e. only those rainfall events with gauges having observed the most critical rainfall for the days of occurrences and situated within 20 km distance from the location where the floods or flash floods occur were selected, whereas the other rainfall events were treated as non-occurrences.

3.2 Improvements of the rainfall threshold

3.2.1 Definition of the rainfall peak intensity-duration threshold

The most representative peak of rainfall intensity was obtained by plotting peak rainfall intensities of various time intervals against their respective durations. The objective of this first step is to distinguish two clear thresholds (lower and upper) that divide the graph into three parts and four distinct groups: (1) the upper part (Group 1), which corresponds with the peak intensities that always lead to flooding or flash flooding occurrences; (2) the middle part, which contains peak intensities that may (Group 2) or may not (Group 3) lead to flooding or flash flooding events; and (3) the lower part (Group 4), which includes peak intensity values that do not lead to flooding or flash flooding. Accordingly, an analysis of the graph based on the following four criteria was also performed in this study to define the time interval of the peak rainfall intensity that better represents the flood and flash flood occurrences: (1) a higher number of occurrences above the upper threshold, (2) a higher number of non-occurrences below the lower threshold, (3) lower amplitude between the upper and lower thresholds, and (4) values of the metrics presented in Sect. 3.3.

3.2.2 Application of tolerance levels

Some studies complemented the rainfall threshold method with probabilities of occurrence to reduce the uncertainties of false alarms for hydrological (e.g. Berti et al. 2012; Huang et al. 2015; Wu et al. 2015; Santos and Fragoso 2016; Brigandì et al. 2017). Aiming to reduce the uncertainties in the middle part of the graph but without losing the characteristics of Group 1 and Group 4, two levels of tolerance (sometimes mentioned as exceedance probability) were used in this study to minimise the amplitude between the upper and lower thresholds, e.g. (1) a new lower threshold defined as the 5% of the occurrences above the lower threshold where the value 5% has also been adopted for landslide studies (e.g. Peruccacci et al. 2009, 2012, 2017; Brunetti et al. 2010; Rossi et al. 2017) and (2) a new upper threshold defined as the 99th percentile of non-occurrences above the lower threshold. The first tolerance level leaves 5% of the

empirical data points below the lower threshold, while the second tolerance level was adopted to leave a minimum number of the non-occurrences above the upper threshold.

3.2.3 Delineating the intermediate threshold

Afterwards, the API was used to analyse the occurrences and non-occurrences of the middle part of the graph after considering the two tolerance levels. The upper and lower parts of the graph were excluded from this further analysis because it is presumed that they are already well-represented by the intensity peaks. Some studies show generally a negative relationship between the antecedent conditions and a critical event rainfall, indicating that with increasingly wet conditions, less rainfall is required to trigger an occurrence (Bai et al. 2014). In this study, the middle part was outlined following the study carried out by Collins et al. (2007) for landslides, which relates rainfall intensity and API by an exponential equation to better identify occurrences and non-occurrences of events at this part of the graph, as follows:

$$I = ae^{b \, API} + c \tag{2}$$

where I is the peak intensity (mm h^{-1}), and a, b, and c are constants to be determined. The constant values were obtained by 50,000 iterations, combining: (1) 50 values of 'a' ranging from the minimum rainfall intensity to three times the maximum rainfall intensity of the occurrences; (2) 20 values of 'b' varying between -0.01 and -1; and (3) 50 values of 'c' ranging from the minimum rainfall intensity to the mean rainfall intensity of the occurrences. The best-fitted constants and the reference day for the API calculation were selected based on the optimal values of the metrics presented in Sect. 3.3.

3.3 Evaluation procedures

The performance of the upper, intermediate and lower thresholds to identify true or false alarms was evaluated using a binary classifier of the rainfall conditions that do or do not lead to flood and flash flood occurrences (Segoni et al. 2014; Turkington et al. 2014; Zhao et al. 2019). A contingency matrix consisting of four components was used for each threshold, including: (1) true positive (TP), when the threshold is exceeded and the hydrological disaster occurs; (2) false negative (FN), when the threshold is not exceeded and the hydrological disaster occurs; (3) false positive (FP), when the threshold is exceeded and the hydrological disaster does not occur; and (4) true negative (TN), when the threshold is not exceeded and the hydrological disaster does not occur. Three metrics were then applied using the contingency matrix to assess the skill score of the flood and flash flood thresholds: (1) probability of detection (POD), which measures the fraction of events that are correctly predicted; (2) false alarm ratio (FAR), which exhibits the fraction of events incorrectly predicted; and (3) positive predictive value (PPV), which shows the probability of events correctly predicted:

$$POD = \frac{TP}{TP + FN} \tag{3}$$

$$FAR = \frac{FP}{FP + TN} \tag{4}$$

$$PPV = \frac{TP}{TP + FP} \qquad (5)$$

The values of these metrics range from 0 to 100%. The optimal score for POD and PPV is close to 100%, while the perfect value for FAR is close to 0%.

3.4 Link to the colour-class warning level systems

In Brazil, national and regional disaster management agencies such as CEMADEN usually use colour-class systems to indicate different levels of risk (e.g. moderate, high, and very high). These systems generally employ classes varying from cold to warm colours to show conditions that could lead to increased risk. Similar risk information, using this colour-class system designed from multiple rainfall thresholds to link threat levels to the emergency, is also used by many disaster management agencies worldwide and scientific studies (e.g. Brunetti et al. 2010; Huang et al. 2015; Jang 2015). Based on this information, the definition of four thresholds using the methodology proposed in this study makes it possible for the implementation of probabilistic schemes for warning level systems predicting flood and flash flood occurrences, defined as follows:

1. Blue alert: rainfall events below the lower threshold that represent a low probability of occurrences when the rainfall conditions are maintained.
2. Yellow alert: rainfall events between the lower and the intermediate thresholds that represent a moderated probability of occurrences when the rainfall conditions are maintained.
3. Orange alert: rainfall events between the intermediate and the upper threshold that represent a high probability of occurrences when the rainfall conditions are maintained.
4. Red alert: rainfall events above the upper threshold that represent an extremely high probability of occurrences when the rainfall conditions are maintained.

4 Results and discussion

4.1 Characterisation of the flood and flash flood occurrences

Figure 1c, d shows the spatial distribution of the 347 and 71 occurrences of flood and flash floods, respectively, in the state of São Paulo between 1 January 2015 and 31 December 2017. Represented by separated points in the map, these occurrences were obtained from the three main sources of data described in Sect. 3.1.2. The main source of occurrences was acquired from the SIMPAT dataset, with 284 (82% of the total) floods and 58 (82% of the total) flash floods. The spatial distribution of information, collected from the different data sources, shows that a large number of floods (59%) and flash floods (55%) were concentrated in areas with population density higher than 500 inhabitants/km^2, which includes only 59 of the 645 municipalities of São Paulo State. The largest number of floods were identified in MASP (45), Bauru (12), and Sorocaba (9). On the other hand, the number of observed flash floods was higher in Bauru (8), São Paulo (6), and Campinas (5). The 240 floods and 47 flash floods occurrences considered in this study were mostly triggered during the rainy season (January–March), which represents 69% and 66% of the total, respectively. According to SIMPAT, the number of socio-economic impacts caused by the floods

and flash floods in São Paulo State during the studied period amount to more than 4310 displacements, 26 injuries, and 17 deaths.

4.2 Rainfall peak intensity-duration threshold

The results of the rainfall thresholds for floods and flash floods, without the use of tolerance levels, are shown in Fig. 4a, b. The thresholds of the upper part of the graph for floods range from 171.6 to 4.2 mm h^{-1} for the rainfall durations of 10 min and 24 h, respectively. The thresholds of the lower part of the graph for floods range from 4.7 to 1.1 mm h^{-1} for the same durations, respectively. As far as flash floods are concerned (Fig. 4b), the thresholds of the upper part of the graph presented similar values when compared to floods
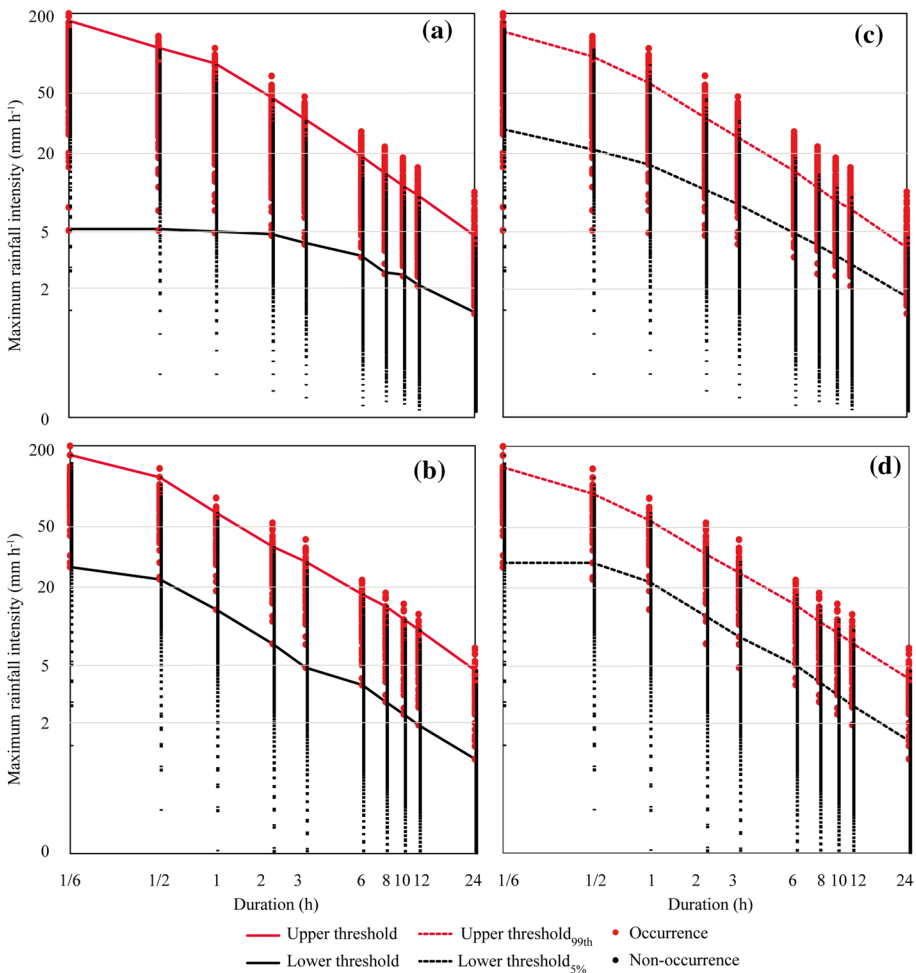
Fig. 4 Rainfall intensities peak versus rainfall duration applying the approach without the tolerance levels for **a** floods and **b** flash floods. Improved application of the methodology using the tolerance levels (99th percentile and 5%) for **c** floods and **d** flash floods. The graphs use logarithmic scale

(between 170.4 and 4.3 mm h^{-1}). Conversely, the lower part exhibited values of intensity peaks six times higher (25.2 mm h^{-1}) for shorter time steps.

It is noticeable that the peak rainfall intensity for longer durations (24 h) presents a better relationship with the eventual flood events, where the upper and lower lines of the threshold tend to be closer. Consequently, the largest number of flood occurrences and non-occurrences was registered above (below) the upper (lower) thresholds, respectively. Thus, a reduced quantity of events in the middle part of the graph, containing both occurrences and non-occurrences, was also observed. This finding differs from the study carried out by Diakakis (2012) in Greece, which found a better relationship for shorter peak intensity duration due to the upper and lower thresholds being much closer in 10 or 30 min durations than in 24 h. On the other hand, the amplitude of the middle part of the graph (the distance between the two thresholds) was similar for all time steps when flash floods were considered. However, the largest number of occurrences/non-occurrences above (below) the upper (lower) thresholds was noticed for the time steps of 1 and 2 h. Papagiannaki et al. (2015) also observed a better separation between flash floods occurrences and non-occurrences in Greece for shorter peak intensity durations, however, only when the analysis is performed on a more local scale.

Table 1 shows the evaluation metrics for predicting floods and flash floods using the two thresholds but without adding of the tolerance levels. It is noticeable that the upper threshold is a precise approach for predicting flood and flash flood occurrences, presenting FAR and PPV values equal to 0% and 100% for all time steps, respectively. However, the upper threshold is only applicable to a very limited number of occurrences, since the POD for

Table 1 Summary of evaluation metrics for the flood and flash flood thresholds, considering the approach without the use of tolerance levels

Metrics	Peak intensities (h)									
	1/6 (%)	1/2 (%)	1 (%)	2 (%)	3 (%)	6 (%)	8 (%)	10 (%)	12 (%)	24 (%)
Flood										
Upper threshold										
POD	1	7	4	7	8	12	12	13	14	17
FAR	0	0	0	0	0	0	0	0	0	0
PPV	100	100	100	100	100	100	100	100	100	100
Lower threshold										
POD	100	100	100	100	100	100	100	100	100	100
FAR	93	83	72	48	41	24	26	20	20	15
PPV	4	4	5	7	8	13	12	15	15	19
Flash flood										
Upper threshold										
POD	3	3	14	15	11	10	7	7	8	13
FAR	0	0	0	0	0	0	0	0	0	0
PPV	100	100	100	100	100	100	100	100	100	100
Lower threshold										
POD	100	100	100	100	100	100	100	100	100	100
FAR	19	13	13	15	17	12	13	13	13	13
PPV	4	7	6	6	5	7	7	7	6	7

this threshold presented low values for floods (from 1 to 17%) and flash floods (from 3 to 15%) for all time steps. The lower threshold exhibits high and low values of FAR (from 15 to 93%) and PPV (from 9 to 19%) for floods, respectively. For flash floods, reduced values of FAR (from 12 to 19%) and PPV (from 4 to 7%) are found for the lower threshold when compared to those observed for floods. These results show that the application of the rainfall peak intensity-duration threshold presents a high number of non-occurrences above the lower threshold, albeit displaying POD values equal to 100% for both type of floods. This behaviour suggests that the approach can detect most occurrences only above the lower threshold, but with a considered level of false alarms regardless of the time step adopted. Similar performance has also been observed in the application of rainfall intensity-duration thresholds for floods and landslides worldwide (e.g. Santos and Fragoso 2016; Brunetti et al. 2018; Zhao et al. 2019). For both upper and lower thresholds, the rainfall peak intensity-duration threshold without the use of tolerance levels presented better results for time steps equal to 1 and 24 h for flash floods and floods, respectively.

4.3 Tolerance levels

Figure 4c, d shows the rainfall thresholds for floods and flash floods after introducing the tolerance levels of 99th percentile for the non-occurrences below the upper and 5% of the occurrences above the lower thresholds. The two tolerance levels were defined to seek to reduce the uncertainties of the middle part of the graph. The application of the tolerance level of 99th percentile corresponded to a mean inclusion of 11 and 6 non-occurrence events of floods and flash floods above the upper threshold, respectively. However, it also brings in an increase of 14% of the number of floods and flash floods occurrences above the upper threshold. For the tolerance level at 5% percentile, the number of occurrences included below the lower threshold was 17 and 4 for floods and flash floods, respectively. However, it was also observed a considerable reduction in the number of non-occurrences of floods (63%) and flash floods (53%) in the middle part of the graph. Similarly, the study carried out by Brunetti et al. (2018) also presented a significant reduction (68%) in the number of non-occurrences for landslides above the threshold after the use of the same tolerance level. However, it is worth highlighting that, in our study, without these tolerance levels, the inclusion of occurrences/non-occurrences in the lower/upper threshold was zero.

It is observed that there is a noticeable decline of the amplitudes between the lower and upper thresholds when floods are considered using the tolerance levels, mainly for the time steps ranging from 10 min to 2 h (Fig. 4c). This reduction of the amplitude between the two thresholds predominantly occurred because of the significant rising of the lower threshold. This leads to the fact that approximately half of the non-occurrences above the lower threshold are excluded and in the meantime the number of flood occurrences above upper thresholds are included, respectively. As far as flash floods are concerned, the largest variations of the lower threshold using the tolerance levels mainly occur between the time steps 1 and 3 h, excluding more than half of the non-occurrences above the originally defined lower threshold (Fig. 4b, d). Conversely, the upper threshold for flash floods remained practically unchanged.

Table 2 shows the assertiveness of the rainfall thresholds for floods and flash floods after the use of the two tolerance levels. The results reveal a considerable improvement of POD for the upper threshold applying the tolerance level of 99th percentile for the non-occurrences, ranging now from 8 to 31% for floods and from 4 to 32% for flash floods. These outcomes obtained for POD correspond to an improvement of 14% for floods and flash

Table 2 Summary of evaluation metrics for the floods and flash floods thresholds considering the tolerance levels (99th percentile and 5%) and the intermediate thresholds

Metrics	Peak intensities (h)									
	1/6 (%)	1/2 (%)	1 (%)	2 (%)	3 (%)	6 (%)	8 (%)	10 (%)	12 (%)	24 (%)
Flood										
Upper threshold (99th)										
POD	8	13	20	24	27	27	31	31	29	31
FAR	0	0	0	0	0	0	0	0	0	0
PPV	82	87	91	92	93	93	94	94	94	94
Lower threshold (5%)										
POD	95	95	95	96	95	95	95	95	95	95
FAR	31	18	14	12	9	8	7	6	5	6
PPV	10	16	19	22	26	29	32	35	38	34
Flash flood										
Upper threshold (99th)										
POD	4	15	27	32	30	23	25	24	25	28
FAR	0	0	0	0	0	0	0	0	0	0
PPV	38	69	79	82	81	76	78	77	78	80
Lower threshold (5%)										
POD	96	97	96	96	96	97	96	96	96	96
FAR	15	8	5	4	4	6	6	6	6	7
PPV	5	10	15	18	16	13	13	13	12	11

flood compared to those acquired by the application of this methodology without the use of the proposed tolerance levels, while the FAR values remained negligible for all time steps. Overall, the PPV values after the use of the tolerance level of 1% for the upper threshold presented a slight decreasing about 9% for floods and 26% flash floods, presenting now variations above 80% and 70% for almost all intensity peaks, respectively. This fact represents a slight loss in the predictive capacity of the threshold using the tolerance level; however, a higher number of occurrences can be found. Thus, the upper threshold with the application of the tolerance level of 1% remains a robust approach for predicting the occurrences. Like the methodology without the application of the tolerance levels, the optimal scores of the metrics for floods and flash floods were observed for longer (8 h) and shorter (2 h) time steps, respectively.

The lower threshold applying the tolerance level of 5% for the flood occurrences resulted in an increase of 16% of the PPV (now ranging from 10 to 38%) and a reduction of 32% of the FAR (now ranging from 5 to 31%), when compared to the approaches without the tolerance level (Table 2). Similar increases can be observed for flash floods, with improvements of 6 and 7% for PPV and FAR rates after adopting this tolerance level, respectively. The better performance of PPV and FAR noticed for floods applying the tolerance level for the lower threshold mainly occurred because of (1) the lower values of rainfall peak intensities observed for its outbreak, and (2) the higher number of flood records included in the lower threshold (17 floods against 4 flash floods). The values of POD equal to about 95% for both floods and flash floods also indicate that almost all occurrences remain represented for all time steps after the use of the tolerance level for the lower threshold.

4.4 Intermediate threshold

This section analyses the use of an exponential equation relating rainfall intensity and API for improving the separation between occurrences and non-occurrences of the intermediate threshold, which represents the main contribution of this study. Figures 5 and 6 show the results of the application of this methodology for floods and flash floods, respectively. It is noticeable that for floods the curves were more influenced by the API for shorter time steps, especially for those equal to 10 min, 1 h, and 3 h. Accordingly, the curves for floods remained barely influenced by the API for time steps equal to 2 and 8 h. For flash floods, the curves presented good sensitivity for almost all time steps, except for 1 h. In general, the intensities for floods and flash floods tended to be constant and not dependent to API for durations higher than 1 h.

Overall, the curves generated by the exponential equations well-characterise the intermediate threshold, where the occurrences and non-occurrences can be obtained correlating rainfall intensity and API. The proposed methodology better includes the occurrences and excludes the non-occurrences for rainfall events with higher and lower values of API, respectively. Moreover, adoption of the exponential curves can help regions with a moderate probability of occurrences (yellow alert) based on limit values of API, regardless of the rainfall intensity (e.g. peak intensities of 1 to 8 h for flash floods in Fig. 6). Also, the exponential curves can determine a region capable of triggering occurrences with low values of API and intensity (e.g. peak intensity of 2 h for flood in Fig. 5).

The application of this approach for floods and flash floods, using an exponential equation for better separating the occurrences from the non-occurrences, presented considerably improved results for almost all analysed metrics and nearly all time steps considered (Table 3). The most representative result for floods was observed for longer time steps, especially for 8 h. Specifically for this time step of 8 h, the POD, FAR and PPV metrics presented values equal to 81%, 1% and 82%, respectively. Meanwhile, the time steps ranging from 1 to 12 h presented similar results for flash floods, highlighting the time steps equal to 2 (POD = 79%, FAR = 1%, and PPV = 43%) and 6 h (POD = 79%, FAR = 1%, and PPV = 44%) which presented the best metrics. Indeed, the use of methodologies considering the API to delineate thresholds has proven to be an outstanding instrument for flood and flash flood hazard predictions and warning systems.

5 Conclusions

This study improved an existing peak rainfall intensity threshold method and created a robust warning system capable of better separating the occurrences from the non-occurrences of floods and flash floods. The improvement of this new approach includes the use of two tolerance levels and the delineation of an intermediate threshold represented by an exponential curve relating rainfall intensity and API. The improvements proposed in this study helped reduce significantly many uncertainties accounted for a considerable number of occurrences and non-occurrences between the upper and lower thresholds. The application of the tolerance levels proposed in this study presented noticeable improvements for the rainfall peak intensity thresholds, with substantial reduction of false alarms after the application of a tolerance level of 5% for the lower threshold. Meanwhile, the number of occurrences above the upper threshold increased by two times after the use of a tolerance
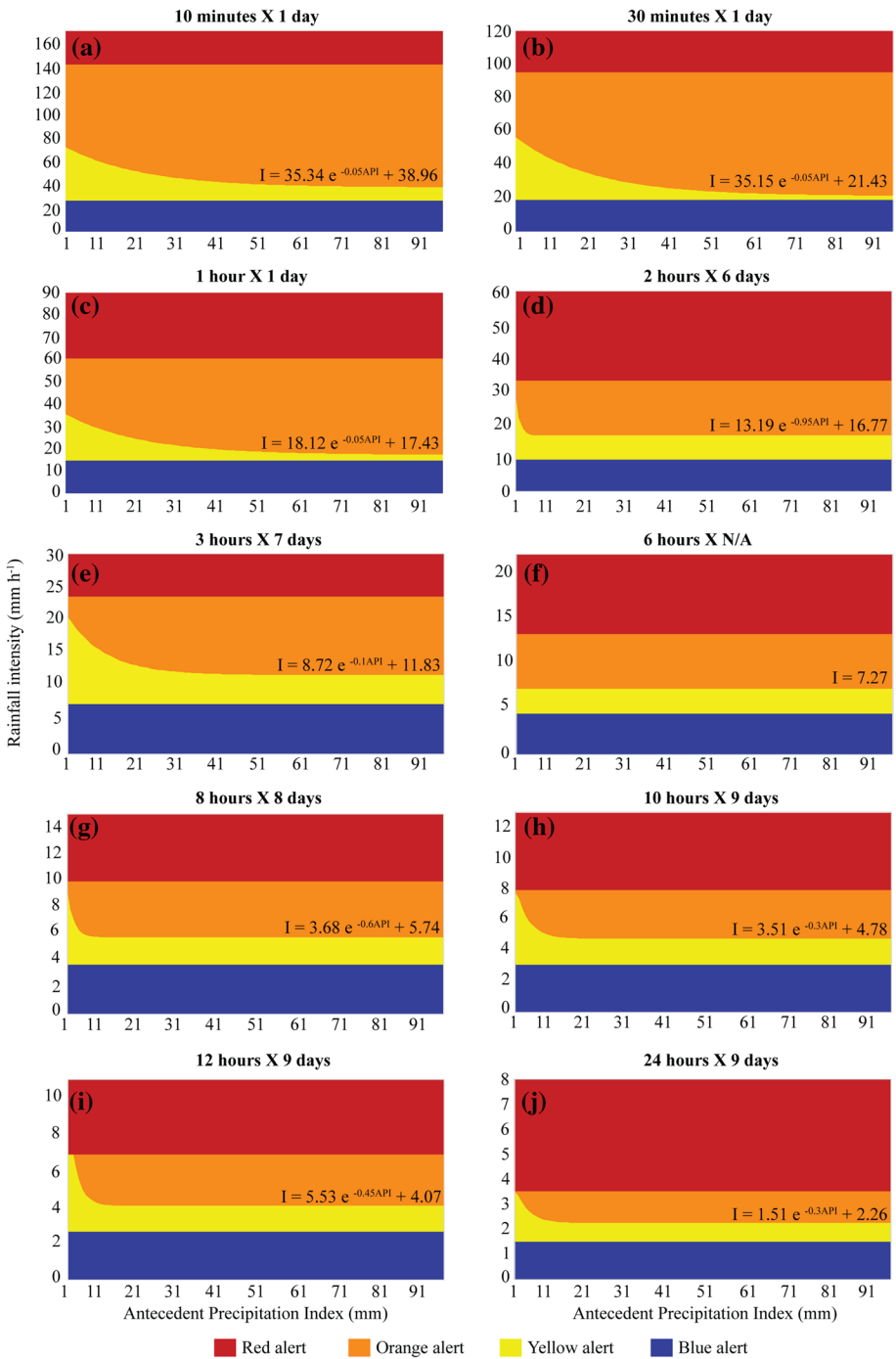
Fig. 5 Peak rainfall intensity versus antecedent precipitation index (API) graphs for each time step and delimitation of the exponential curves for warning level systems applied for floods
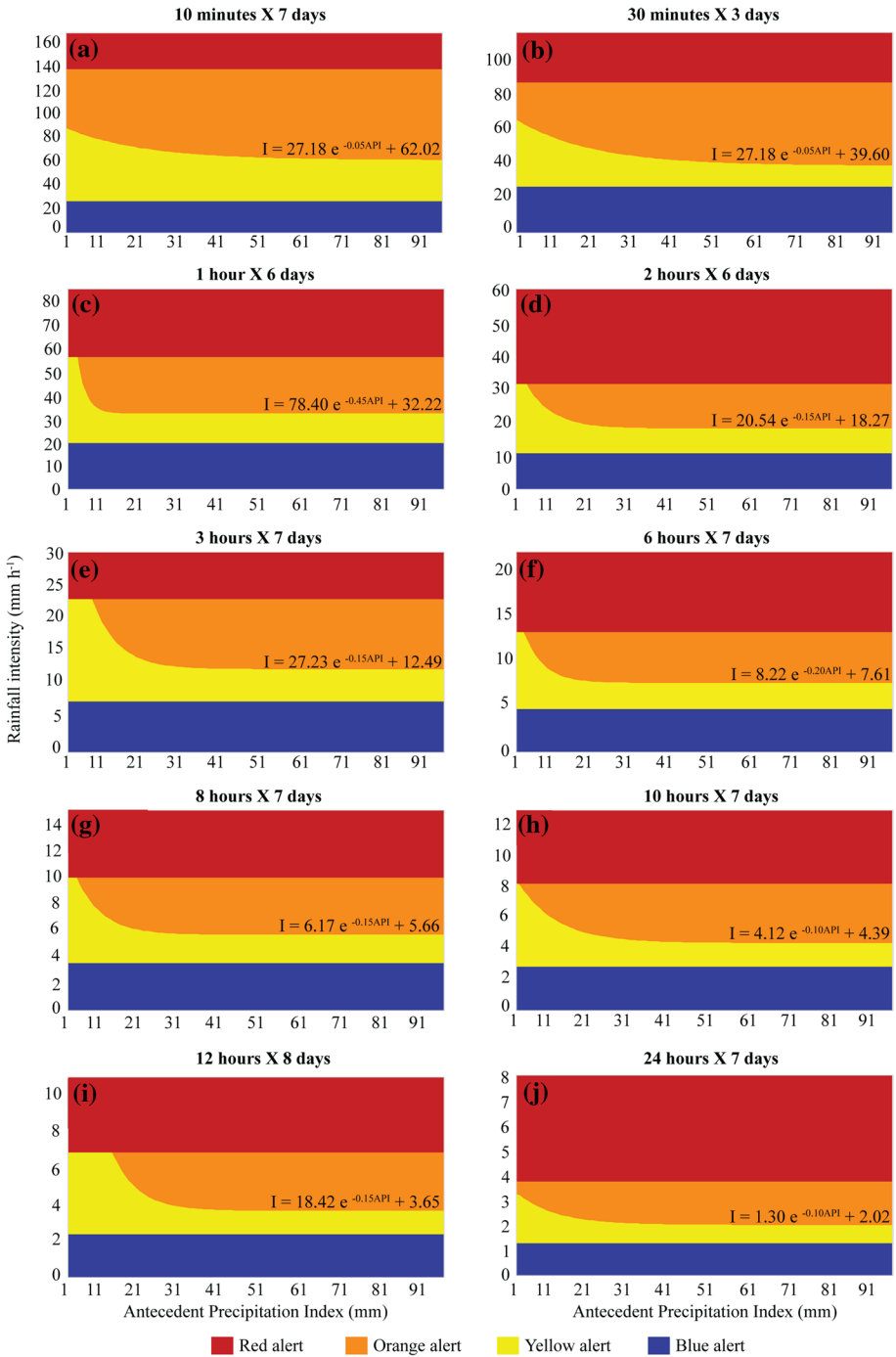
Fig. 6 Peak rainfall intensity versus antecedent precipitation index (API) graphs for each time step and delimitation of the exponential curves for warning level systems applied for flash floods

Table 3 Summary of evaluation metrics for flood and flash flood occurrences applying the improved thresholds that use rainfall intensities and antecedent precipitation index (API)

Intensities (h)×API (day)	Floods			Intensities (h)×API (day)	Flash floods		
	Metrics				Metrics		
	POD (%)	FAR (%)	PPV (%)		POD (%)	FAR (%)	PPV (%)
1/6×1	77	6	30	1/6×7	85	3	21
1/2×1	75	4	41	1/2×3	82	1	33
1×1	78	3	48	1×6	79	1	40
2×6	83	3	51	2×6	79	1	43
3×7	82	2	57	3×7	80	1	39
6×N/A	83	2	61	6×7	79	1	44
8×8	81	1	82	8×7	79	1	41
10×9	80	1	66	10×7	80	1	39
12×9	79	1	66	12×8	82	1	39
24×9	78	1	66	24×7	85	1	37

level of the 99th percentile, improving the effectiveness of the issuance of warnings. The delineation of the intermediate threshold also presented improvements for almost all time steps considered in this study, although the scores of the metrics showed a slightly worse performance for flash floods when compared to floods. This better performance noted for floods probably occurred because the higher amount of data available for this type of event when compared to those observed for flash floods. Additionally, the use of a denser rain gauge network, with stations closer to the flash flood occurrences than those used in this study, could be more effective in capturing this type of event.

Overall, the two methods proposed in this study are shown to be able to reduce the uncertainties in predicting the occurrences of floods and flash floods. It must be mentioned that a considerable amount of flood and flash flood information in São Paulo State could not be used in this study because of the poor data quality of some rain gauges and/or the lack of rain gauge coverage at all. Thus, satellite-based rainfall products with high spatiotemporal resolution could offer new opportunities for a larger-scale analysis. Further work using more rainfall properties correlated with other variables (e.g. soil moisture) could be tested to reduce even more of the uncertainties.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

Abon CC, David CPC, Tabios GQ (2012) Community-based monitoring for flood early warning system: an example in central Bicol River basin, Philippines. Disaster Prev Manag 21:85–96. https://doi.org/10.1108/09653561211202728

Aleotti P (2004) A warning system for rainfall-induced shallow failures. Eng Geol 73:247–265. https://doi.org/10.1016/j.enggeo.2004.01.007

Alvares CA, Stape JL, Sentelhas PC et al (2013) Köppen's climate classification map for Brazil. Meteorol Zeitschrift 22:711–728. https://doi.org/10.1127/0941-2948/2013/0507

Azari H, Matkan AA, Shakiba A, Pourali H (2008) Flood early warning with integration of hydrologic and hydraulic models, RS and GIS (Case study: Madarsoo Basin, Iran). In: Proceedings of the 29th Asian Conference Remote Sensing 2008, vol 3. ACRS 2008, pp 1679–1685

Bacelar LCSD, Maciel A, Angelis CF, Tomasella J (2020) Limiares de chuva deflagradores de inundações bruscas: metodologia, aplicação e avaliação em ambiente operacional. Rev DAE 68:71–86

Bai S, Wang J, Thiebes B et al (2014) Analysis of the relationship of landslide occurrence with rainfall: a case study of Wudu County, China. Arab J Geosci 7:1277–1285. https://doi.org/10.1007/s12517-013-0939-9

Berti M, Martina MLV, Franceschini S et al (2012) Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. J Geophys Res Earth Surf 117:1–20. https://doi.org/10.1029/2012JF002367

Bogaard T, Greco R (2018) Invited perspectives: hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. Nat Hazards Earth Syst Sci 18:31–39. https://doi.org/10.5194/nhess-18-31-2018

Brigandì G, Tito Aronica G, Bonaccorso B et al (2017) Flood and landslide warning based on rainfall thresholds and soil moisture indexes: the HEWS (Hydrohazards Early Warning System) for Sicily. Adv Geosci 44:78–88. https://doi.org/10.5194/adgeo-44-79-2017

Brollo MJ, Ferreira CJ (2016) Gestão de risco de desastres devido a fenômenos geodinâmicos no estado de São Paulo: Cenário 2000–2015. Instituto Geológico, São Paulo

Brunetti MT, Melillo M, Peruccacci S et al (2018) How far are we from the use of satellite rainfall products in landslide forecasting? Remote Sens Environ 210:65–75. https://doi.org/10.1016/j.rse.2018.03.016

Brunetti MT, Peruccacci S, Rossi M et al (2010) Rainfall thresholds for the possible occurrence of landslides in Italy. Nat Hazards Earth Syst Sci 10:447–458. https://doi.org/10.5194/nhess-10-447-2010

Caine N (1980) The rainfall intensity-duration control of shallow landslides and debris flows. Geogr Ann Ser A 62:23–27. https://doi.org/10.1080/04353676.1980.11879996

Cannon SH, Gartner JE, Wilson RC et al (2008) Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. Geomorphology 96:250–269. https://doi.org/10.1016/j.geomorph.2007.03.019

CEPED-Centro Universitário de Estudos e Pesquisas sobre Desastres (2013) Atlas brasileiro de desastres naturais 1991 a 2012: volume Brasil, 2nd ed. CEPED UFSC, Florianópolis

Chleborad AF, Baum RL, Godt JW, Powers PS (2008) A prototype system for forecasting landslides in the Seattle, Washington, area. GSA Rev Eng Geol 20:103–120. https://doi.org/10.1130/2008.4020(06)

Collins BD, Kayen R, Sitar N (2007) Process-based empirical prediction of landslides in weakly lithified coastal cliffs, San Francisco, California, USA. Landslides and climate change: challenges and solutions: proceedings of the international conference on landslides and climate change. https://doi.org/10.1201/noe0415443180.ch22

Diakakis M (2012) Rainfall thresholds for flood triggering. The case of Marathonas in Greece. Nat Hazards 60:789–800. https://doi.org/10.1007/s11069-011-9904-7

Du S, Shi P, Van Rompaey A, Wen J (2015) Quantifying the impact of impervious surface location on flood peak discharge in urban areas. Nat Hazards 76:1457–1471. https://doi.org/10.1007/s11069-014-1463-2

Dunkerley D (2008) Identifying individual rain events from pluviograph records: a review with analysis of data from an Australian dryland site. Hydrol Process 22:5024–5036. https://doi.org/10.1002/hyp

Froidevaux P, Schwanbeck J, Weingartner R et al (2015) Flood triggering in Switzerland: the role of daily to monthly preceding precipitation. Hydrol Earth Syst Sci 19:3903–3924. https://doi.org/10.5194/hess-19-3903-2015

Garcia MGM, Brilha J, de Lima FF et al (2018) The inventory of geological heritage of the state of São Paulo, Brazil: methodological basis, results and perspectives. Geoheritage 10:239–258. https://doi.org/10.1007/s12371-016-0215-y

Glade T, Crozier M, Smith P (2000) Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model." Pure Appl Geophys 157:1059–1079. https://doi.org/10.1007/s000240050017

González-Cao J, García-Feal O, Fernández-Nóvoa D et al (2019) Towards an automatic early warning system of flood hazards based on precipitation forecast: the case of the Miño River (NW Spain). Nat Hazards Earth Syst Sci 19:2583–2595. https://doi.org/10.5194/nhess-19-2583-2019

Guha-Sapir D, Hoyois P, Below R (2015) Annual disaster statistical review 2015 the numbers and trends Centre for Research on the Epidemiology of Disasters (CRED)

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity: duration control of shallow landslides and debris flows: an update. Landslides 5:3–17. https://doi.org/10.1007/s10346-007-0112-1

Huang J, Ju NP, Liao YJ, Liu DD (2015) Determination of rainfall thresholds for shallow landslides by a probabilistic and empirical method. Nat Hazards Earth Syst Sci 15:2715–2723. https://doi.org/10.5194/nhess-15-2715-2015

IBGE-Instituto Brasileiro de Geografia e Estatística (2018) (IBGE). https://cidades.ibge.gov.br/brasil/sp/. Accessed 16 Jul 2019

Jang JH (2015) An advanced method to apply multiple rainfall thresholds for urban flood warnings. Water (Switzerland) 7:6056–6078. https://doi.org/10.3390/w7116056

Kohler MA, Linsley RK (1951) Predicting the runoff from sorm rainfall. US Weather Bur Res Pap 34

Lee JH, Park HJ (2016) Assessment of shallow landslide susceptibility using the transient infiltration flow model and GIS-based probabilistic approach. Landslides 13:885–903. https://doi.org/10.1007/s10346-015-0646-6

Li Z, Zhang H, Singh VP et al (2019) A simple early warning system for flash floods in an ungauged catchment and application in the Loess Plateau. China Water (Switzerland). https://doi.org/10.3390/w11030426

Mirus BB, Becker RE, Baum RL, Smith JB (2018) Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. Landslides 15:1909–1919. https://doi.org/10.1007/s10346-018-0995-z

Natural Resources Conservation Service (NRCS) (1972) National engineering handbook, section 4, Hydrology, Washington, DC

Papagiannaki K, Lagouvardos K, Kotroni V, Bezes A (2015) Flash flood occurrence and relation to the rainfall hazard in a highly urbanized area. Nat Hazards Earth Syst Sci 15:1859–1871. https://doi.org/10.5194/nhess-15-1859-2015

Peruccacci S, Brunetti MT, Gariano SL et al (2017) Rainfall thresholds for possible landslide occurrence in Italy. Geomorphology 290:39–57. https://doi.org/10.1016/j.geomorph.2017.03.031

Peruccacci S, Brunetti MT, Luciani S et al (2012) Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy. Geomorphology 139–140:79–90. https://doi.org/10.1016/j.geomorph.2011.10.005

Peruccacci S, Brunetti MT, Rossi M, Guzzetti F (2009) Rainfall thresholds for the initiation of landslides in Italy. Assembly 11:2729

Rossi M, Luciani S, Valigi D et al (2017) Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data. Geomorphology 285:16–27. https://doi.org/10.1016/j.geomorph.2017.02.001

Santos M, Fragoso M (2016) Precipitation thresholds for triggering floods in the Corgo basin. Portugal Water (Switzerland). https://doi.org/10.3390/w8090376

Scheevel CR, Baum RL, Mirus BB, Smith JB (2017) Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett. Washington. US Geological Survey Open-File Report 2017–1039

Segoni S, Rosi A, Rossi G et al (2014) Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems. Nat Hazards Earth Syst Sci 14:2637–2648. https://doi.org/10.5194/nhess-14-2637-2014

Setzer J (1946) Revista Brasileira de Geografia Física. Rev Bras Geogr 08:3–26

Souza CM, Shimbo JZ, Rosa MR et al (2020) Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. Remote Sens. https://doi.org/10.3390/RS12172735

Špitalar M, Gourley JJ, Lutoff C et al (2014) Analysis of flash flood parameters and human impacts in the US from 2006 to 2012. J Hydrol 519:863–870. https://doi.org/10.1016/j.jhydrol.2014.07.004

Suribabu CR, Sujatha ER (2019) Evaluation of moisture level using precipitation indices as a landslide triggering factor-a study of Coonoor Hill Station. Climate. https://doi.org/10.3390/cli7090111

Tominaga LK, Santoro J, do Amaral R, (2015) Desastres naturais: conhecer para prevenir, 3rd edn. Instituto Geológico, São Paulo

Tramblay Y, Bouaicha R, Brocca L et al (2012) Estimation of antecedent wetness conditions for flood modelling in northern Morocco. Hydrol Earth Syst Sci 16:4375–4386. https://doi.org/10.5194/hess-16-4375-2012

Tsakiris G (2014) Flood risk assessment: concepts, modelling, applications. Nat Hazards Earth Syst Sci 14:1361–1369. https://doi.org/10.5194/nhess-14-1361-2014

Turkington T, Ettema J, Van Westen CJ, Breinl K (2014) Empirical atmospheric thresholds for debris flows and flash floods in the southern French Alps. Nat Hazards Earth Syst Sci 14:1517–1530. https://doi.org/10.5194/nhess-14-1517-2014

Viessman W, Lewis GL (1996) Introduction to hydrology, 4th edn. Harper Collins, New York

Villarini G, Krajewski WF, Ntelekos AA et al (2010) Towards probabilistic forecasting of flash floods: the combined effects of uncertainty in radar-rainfall and flash flood guidance. J Hydrol 394:275–284. https://doi.org/10.1016/j.jhydrol.2010.02.014

Wallemacq P, House R (2018) Economic losses, poverty and disasters 1998–2017. UNDRR and CRED, Geneva

Wu SJ, Hsu CT, Lien HC, Chang CH (2015) Modeling the effect of uncertainties in rainfall characteristics on flash flood warning based on rainfall thresholds. Nat Hazards 75:1677–1711. https://doi.org/10.1007/s11069-014-1390-2

Zhao B, Dai Q, Han D et al (2019) Antecedent wetness and rainfall information in landslide threshold definition. Hydrol Earth Syst Sci Discuss. https://doi.org/10.5194/hess-2019-150

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.