Abstract. Both $M_{S}$ 8.0 Wenchuan earthquake on 12 May 2008 and $M_{S}$ 7.0 Lushan earthquake on 20 April 2013 occurred in the province of Sichuan, China. In the earthquake-affected mountainous area, a large amount of loose material caused a high occurrence of debris flow during the rainy season. In order to evaluate the rainfall intensity–duration ($I$–$D$) threshold of the debris flow in the earthquake-affected area, and to fill up the observational gaps caused by the relatively scarce and low-altitude deployment of rain gauges in this area, raw data from two S-band China New Generation Doppler Weather Radar (CINRAD) were captured for six rainfall events that triggered 519 debris flows between 2012 and 2014. Due to the challenges of radar quantitative precipitation estimation (QPE) over mountainous areas, a series of improvement measures are considered: a hybrid scan mode, a vertical reflectivity profile (VPR) correction, a mosaic of reflectivity, a merged rainfall–reflectivity ($R$–$Z$) relationship for convective and stratiform rainfall, and rainfall bias adjustment with Kalman filter (KF). For validating rainfall accumulation over complex terrains, the study areas are divided into two kinds of regions by the height threshold of 1.5 km from the ground. Three kinds of radar rainfall estimates are compared with rain gauge measurements. It is observed that the normalized mean bias (NMB) is decreased by 39 % and the fitted linear ratio between radar and rain gauge observation reaches at 0.98. Furthermore, the radar-based $I$–$D$ threshold derived by the frequentist method is $I = 10.1D^{-0.52}$ and is underestimated by uncorrected raw radar data. In order to verify the impacts on observations due to spatial variation, $I$–$D$ thresholds are identified from the nearest rain gauge observations and radar observations at the rain gauge locations. It is found that both kinds of observations have similar $I$–$D$ thresholds and likewise underestimate $I$–$D$ thresholds due to undershooting at the core of convective rainfall. It is indicated that improvement of spatial resolution and measuring accuracy of radar observation will lead to the improvement of identifying debris flow occurrence, especially for events triggered by the strong small-scale rainfall process in the study area.

1 Introduction

Rainfall-induced debris flow is a kind of ubiquitous natural hazard for mountainous areas with complex terrain. It is a geomorphic movement process which scours the sediment from steep areas into alluvial fans. The formation of rainfall-induced debris flow is generally related to three
main factors, including gravitational potential energy, abundant loose materials and meteorological events (Guzzetti et al., 2008). The gravitational potential energy remains relatively stable for a long period of time. The loose materials are normally made up of sand, unsorted silt, cobbles, gravel, boulders and woody debris (Wang et al., 2016). High-magnitude earthquake events can generate abundant loose solid material from co-seismic rock falls and landslides and deposited in gullies (Shieh et al., 2009). During the rainy season, the occurrence of debris flow after an earthquake becomes more frequent (Yu et al., 2014; Guo et al., 2016a). Both the M8.0 Wenchuan earthquake on 12 May 2008 and the M7.0 Lushan earthquake on 20 April 2013 occurred in the province of Sichuan, China, and have changed the formation conditions for debris flow. A large number of debris flows occurred from 2008 to 2014 and caused many casualties and extensive property damage.

Early warning systems (EWS) for rainfall-induced landslide and debris flow are widely implemented in many parts of the world (Baum and Godt, 2010; Glade and Nadim, 2014; Segoni et al., 2015). The performance of EWS relies highly on the updating of precipitation thresholds (Rosi et al., 2015). Furthermore, a large amount of loose materials caused by earthquake highly increases the occurrence of debris flow (Tang et al., 2009, 2012), it is necessary to revaluate the precipitation threshold. The model of rainfall intensity–duration (I−D) is widely used to represent the precipitation thresholds of triggering landslides and debris flow (Alecotti, 2004; Guzzetti et al., 2007). Some literature concluded that the I−D relationships for some of the regions were severely affected by the Wenchuan earthquake (Su et al., 2012; Cui et al., 2013; Zhou and Tang, 2014; Guo et al., 2016b). However, most of these I−D relationships are derived from rain gauge observation. This is a common technical way to estimate the I−D thresholds of debris flows using rainfall observation from the nearest rain gauge. However, the uncertainty of I−D thresholds from rain gauge observations could not be ignored. This is related to two critical limitations which probably lead to underestimation of observation of strong convective events occurring at high-altitude areas. The first limitation is the relatively sparse network density of rain gauges in the mountainous region (Marra et al., 2014); the other one is the altitude of gauge deployments, which is at low elevation for sustainability. The same limitations of rain gauge observation also exist in the mountainous regions of Sichuan. The technique of microwave remote sensing has become a necessary way for observing rainfall events in complex terrain. The radar-based quantitative precipitation estimation (QPE) has been shown to be useful for the study of debris flows, as its unique advantage of high spatial and temporal resolution. Radar observations offer the unique merit of estimating rainfall over the actual debris flow location (David-Novak et al., 2004; Chiang and Chang, 2009; Marra et al., 2014; Berenguer et al., 2015). However, there are many challenges when radar-based QPE in the mountainous area is applied to the study of debris flow. Commonly, keeping the elevation angle close to the ground and estimating the sample cut at the same height is a basic requirement for radar QPE to represent the actual rainfall distribution on the ground. The radar beam blocked by the mountain is a serious problem for the low angle observation. The radar beam angle has to be elevated to avoid the blockage. However, doing this introduces another problem: rainfall distribution at higher heights is different from that at the surface and varies greatly according to the precipitation type (Zhang et al., 2012). Errors due to radar system calibration and uncertainty in hydrometeor’s DSD (drop size distribution) also decrease the accuracy of rainfall estimates. Therefore, the combination of radar and rain gauges to provide accurate rainfall estimates in complex terrain is attracting increasingly more interest for improving warnings of future precipitation and situational awareness (Willie et al., 2017). Furthermore, debris-flow-triggering events are often related to high precipitation gradients of storms which occur for a short duration and are on a small scale (Nikolopoulos et al., 2015). Considering this, raw S-band radar reflectivity data are used to estimate rainfall and assess the impact of estimation errors on the identification of the I−D threshold over the study area.

The main aim of this study is to merge the radar QPE, thereby improving its estimation over complex terrain, and to assess the impact of rainfall estimate accuracy on the identification of I−D threshold over the study area. To do that, a series of accuracy-improving measures have been adopted including a hybrid scan mode, a vertical reflectivity profile (VPR) correction, a mosaic of reflectivity, a combination of rainfall–reflectivity (R−Z) relationship for convective and stratiform rainfall, and rainfall bias adjustment with Kalman filter (KF). Three radar rainfall estimation scenarios are evaluated with the rain gauge observations for six debris-flow-triggering rainfall events to validate the accuracy of radar estimate. I−D thresholds are identified from 519 rainfall-induced debris flow events with the frequentist method (Brunetti et al., 2010; Peruccacci et al., 2012). Another aim of this study is to understand the impact on the I−D identification due to spatial variability of rainfall observation. Rain gauge observations nearest to the debris flow within 10 km and radar observations at the rain gauge locations are used to get the I−D relationship.

2 Study domain and data

The study area is located in Sichuan in southwest China, which consists of 16 administrative districts and counties. The area of study is about 38 000 km² and occupies nearly 8% of the land area of Sichuan (see Fig. 1). This area was strongly affected by the M8.0 Wenchuan earthquake which occurred on 12 May 2008 and the M7.0 Lushan earthquake which occurred on 20 April 2013. In the following years, debris flow happened frequently. During the period from 2012
Figure 1. Location and topography of the study area. Asterisks show the location of Chengdu and Mianyang S-band weather radars which monitor the study area within 150 km (dash black circle) from the radar location. Rain gauges in the study area are marked with black triangles and mostly deployed in the valley. The two blue circle dots are the epicenter of the $M_s$ 8.0 Wenchuan earthquake on 12 May 2008 and the $M_s$ 7.0 Lushan earthquake on 20 April 2013.

Figure 2. Land use map (a) and lithology map (b) for the study area.

to 2014, the debris flow occurring in this area accounted for 58.3% of the annual debris flow events which occurred in the entire province. The area is in the transitional zone of the Qinghai–Tibet Plateau to the Sichuan Basin. Terrain changes steeply and the average altitude above sea level (a.s.l.) for this area is between 500 m and 6 km. The geological structure of the study area shows a northeast to southwest orientation. The rocks over this region are mainly comprised of volcanic rocks, mixed sedimentary rocks, siliciclastic sedimentary rocks, carbonate sedimentary rocks, acid plutonic rocks, intermediate volcanic rocks, intermediate plutonic rocks, unconsolidated sediments, metamorphic rocks, basic plutonic rocks and pyroclastic rocks. Figure 1a shows the lithological map. Quaternary deposits were distributed in the form of river terraces and alluvial fans. Owing to frequent tectonic activities, most of the gully is steeply sloped over this area, as shown in Fig. 2b. The main land use types in this region are mixed forest, cropland and grassland, as shown in Fig. 2a. Potential debris flow watersheds over the study area were extracted from morphological variables, using the logistic regression method. Berenguer et al. (2015) simplified the geomorphological variables, as the watersheds maximum height ($h_{max}$), mean slope ($s_{mean}$), mean aspect ($\theta_{mean}$) and Melton ratio (MR) are the variables with the smallest overlapping areas for assessing the susceptibility of the watersheds. The $h_{max}$, $s_{mean}$, $\theta_{mean}$ and MR were retrieved from DEM data. Combined with the debris flow occurrence over this area during 3 years, the potential susceptibility map was calculated with the logarithm regression method, as shown
Table 1. Characteristics of S-band Doppler weather radar.

<table>
<thead>
<tr>
<th>Items</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength</td>
<td>10.4 cm</td>
</tr>
<tr>
<td>Polarized mode</td>
<td>horizontal</td>
</tr>
<tr>
<td>Antenna gain</td>
<td>45</td>
</tr>
<tr>
<td>First side lobe level</td>
<td>$-29$ dBc</td>
</tr>
<tr>
<td>Peak transmitted power</td>
<td>750 kW</td>
</tr>
<tr>
<td>Noise figure</td>
<td>4 dB</td>
</tr>
<tr>
<td>Dynamic range</td>
<td>90 dB</td>
</tr>
<tr>
<td>Range resolution</td>
<td>300 m</td>
</tr>
<tr>
<td>Volume scanning elevation</td>
<td>0.5, 1.5, 2.4, 3.4, 4.3, 6.0, 9.5, 14.5, 19.5$^\circ$</td>
</tr>
<tr>
<td>Altitude above sea level of radar location</td>
<td>595 m for Chengdu site, 557 m for Mianyang site</td>
</tr>
</tbody>
</table>

The area is monitored by two well-maintained S-band Doppler weather radars (see Fig. 1). One is deployed in Chengdu city at an altitude of 596 m a.s.l. and the other one is deployed at Mianyang city at a height of 557 m a.s.l. Both of the radar systems have same system specifications, which can be seen in Table 1. The system provides radar rainfall estimates at a radial range resolution of 300 m and an angular resolution of $1^\circ$. There is a rain gauge network consisting of 551 gauges equipped at the meteorological surface station in the study areas. The number of rain gauges seems to be a lot, but most of them are deployed at the valleys. The density of rain gauges is severely scarce at the high altitude of the mountains, resulting in observation gaps where the debris flow initially takes place. The average altitude above sea level of those rain gauges is far lower than 3 km.

3 Methods

3.1 Radar-accumulated rainfall estimation methods

S-band weather radar has a unique advantage of being unaffected by attenuation, as it is subjected to Rayleigh scattering for almost all hydrometeors. However, in complex terrain conditions, S-band radar observations still face serious challenges. The main problem comes from ground clutter and severe beam blockage, resulting in inaccurate estimates of radar rainfall. A number of signal processing techniques have been developed to detect and remove clutter and anomalous propagation, including fuzzy logic, ground echo maps and Gaussian model adaptive processing (GMAP) filters (Harrison et al., 2000; Berenguer et al., 2006; Nguyen and Chandrasekar, 2013). For the radar data used in this study, ground clutter is filtered with the GMAP algorithm configured in the Vaisala Sigmet digital processor. Furthermore, in order to overcome the beam blockage and improve the rainfall estimation accuracy, radar data are corrected concerning the following issues: (i) beam shielding and hybrid scan, (ii) vertical profile of reflectivity, (iii) mosaic of hybrid scan reflectivity, (iv) combination of reflectivity rainfall relationship and (v) rainfall bias adjustment.

3.1.1 Beam shielding and hybrid scan

The hybrid scan mode is used to form the initial reflectivity field for rainfall estimate by keeping the radar main beam away from the blockage of the complex terrain (Zhang et al., 2012). In the study area, the grids with 0.36 km$^2$ resolution on the ground are aligned with radar bins of each elevation angle. The blockage coefficients of the low elevation angles at 0.5, 1.5 and 2.4$^\circ$ are calculated according to the digital elevation model (DEM), earth curvature, antenna pattern and the wave propagation model (Pellarin et al., 2002; Krajewski et al., 2006). The blockage ratio distribution of two S-band radars can be seen in Fig. 4. There is almost no topographical shielding in the near-field within a distance of 50 km from each radar. The main factor considered in the hybrid scan within 50 km is to meet the estimated rainfall from the same vertical height as much as possible. Thus the area within 20 km from radar is assigned with elevation angle of 3.4$^\circ$, the area from radar between 20 and 35 km is assigned the elevation angle of 2.4$^\circ$ and the area from radar between 35 and 50 km is assigned the elevation angle of 1.5$^\circ$. It is assigned with the elevation angle of 0.5$^\circ$ by default when there is no blockage over 50 km distance from the radar. The terrain transforms from plain land to a mountainous region over about 70 km westward of each radar. At this region the altitude rises sharply, and the elevation angle of 0.5$^\circ$ is totally obscured. Therefore, the lowest angle at which the blockage ratio does not surpass 0.5 is assigned to the aligned grid. Meanwhile, the blockage ratio is correspondingly used to compensate the energy loss of reflectivity. The final adaptive-terrain hybrid scan maps are combined as shown in Fig. 4d and h. It can be seen that most of the study area is covered by the 1.5 and 2.4$^\circ$ radar scans.

3.1.2 Vertical profile of reflectivity

Due to the hybrid scan, the radar elevation angle is raised, resulting in the majority of the observed reflectivity coming from the upper levels of precipitation profiles. This is
Figure 3. Morphology and potential debris flow watersheds map over study area: (a) slope; (b) aspect; (c) potential debris flow watersheds (gray polygon) with debris flow observation (blue circle).

Figure 4. Blockage ratio of beam shielding for the radar main lobe beam and hybrid scan map. Panels (a)–(c) represent the blockage ratio of Chengdu radar at elevations of 0.5, 1.5 and 2.4°, respectively. Panels (e)–(g) represent the blockage ratio of Mianyang radar at elevations of 0.5, 1.5 and 2.4°, respectively. Hybrid scan maps for Chengdu (d) and Mianyang (h) are merged as long as the blockage ratio is lower than 0.5.

Quite different from the actual reflectivity on the ground. It is necessary to account for the reflectivity correction at the ground level. This study adopts the VPR method to adjust the reflectivity (Zhang et al., 2012). The processing steps applied in this study are as follows: (i) convection precipitation is discriminated from stratiform based on the composite reflectivity > 50 dBz or VIL > 6.5 kg m$^{-2}$, where VIL is vertically integrated liquid water content, an estimate of the total mass of precipitation in the clouds (Amburn and Wolf, 1997). (ii) The parameterization of VPR is carried out to generate bright band top, peak, bottom heights and piecewise linear slopes $S_1$, $S_2$ and $S_3$ (see Fig. 5). (iii) Observed reflectivity is adjusted based on the parameterized VPR to piecewise extrapolate the corresponding reflectivity at the ground. Figure 5 shows a sample scatter plot of the vertical reflectivity profiles from 11:30 to 12:30 UTC+8 on 21 July 2012. Impacted by the temperature, air dynamics, particle size and phase are changed along the vertical falling. Figure 5 shows the vertical profile of reflectivity varied approximately as three piecewise linear sections. Altitude is one the critical factors affecting the atmosphere physics parameters and the performance of VPR. The areas of study are classified into two types, region types I and II, in relation to the height from the ground (≤ 1.5 km for region type I and > 1.5 km for region type II) and the distance from the radar (≤ 100 km for region type I and > 100 km for region type II). Figure 6 shows the identification results for both radars. Apart from the VPR adjustment, these two kinds of regions are assessed during the validation of radar QPE in order to understand the actual impact of distance and height of radar observations on the rainfall estimation.
3.1.3 Mosaic of hybrid scan reflectivity

Both S-band radars have common coverage areas where reflectivity data should be mosaicked to construct a large-scale sensing for rainfall events. Taking the distance and altitude as weighing parameters, the mosaic formula is defined as

$$Z_M = \frac{\sum_i w_i \times k_i \times Z_i}{\sum_i w_i \times k_i},$$

and

$$w_i = \exp\left(-\frac{d_i^2}{L^2}\right),$$

$$k_i = \exp\left(-\frac{h_i^2}{H^2}\right).$$

Here, $Z_M$ represents the mosaicked hybrid scan reflectivity, $Z_i$ is the single radar hybrid scan reflectivity, $i$ is the radar index, $w$ is weighing component for the horizontal weighting and $k$ is weighing component for the vertical weighting. The variable $d$ is the distance between the analysis grid and the radar, and $h$ is the height above the ground of the single radar hybrid scan. The parameters $L$ and $H$ are scale factors of the two weighting functions.

3.1.4 Combination of rainfall relationship

Rainfall rates are calculated from radar reflectivity by a power law empirical relationship called the $R - Z$ relationship (Austin, 1987; Rosenfeld et al., 1993) and, theoretically, the $R - Z$ relationships should be adjusted when the DSDs change over the rainfall duration. However, it is still a challenge to obtain fine spatial distribution of DSDs with changes of time over complex terrains. This study adopts the two widely verified $R - Z$ relationships defined as $Z = 300R^{1.4}$ for convective precipitation (Fulton et al., 1998) and $Z = 200R^{1.6}$ for stratiform (Marshall et al., 1955), and the rainfall type is identified during VPR processing.

3.1.5 Rainfall bias adjustment

The errors of the $R - Z$ relationship mainly come from DSD variation, radar calibration errors, etc. (Berne and Krajewski, 2013), so the rainfall biases change over time. The mean field bias correction is a method to calculate the ratio of the means of radar estimate and the rain gauge observation (Anagnostou and Krajewski, 1999; Chumchean et al., 2003; Yoo and Yoon, 2010). In this study, the bias is calculated based on hourly radar rainfall accumulation and rain gauge accumulated observation. It is defined as

$$BIAS = \frac{\sum_i r_i}{\sum_i g_i},$$

where $BIAS$ is mean rainfall bias in 1 h, $g_i$ is 1 h accumulated rainfall of rain gauge, $i$ is rain gauge index, $r_i$ is the radar-based 1 h accumulated rainfall over the $i$th rain gauge and $N$ is the total number of rain gauges. As described above, the density of rain gauge deployment over the mountainous area is relatively scarce. Therefore the precipitation measured by individual gauges at high and low altitudes may lead to overestimation and underestimation, respectively. Therefore, the KF is adopted to alleviate the measurements noise of the bias (Ahner, 1986; Chumchean et al., 2006; Kim and Yoo, 2014).

The basic steps of KF in this study are as follows.

1. State the estimate prediction:

$$BIAS_P(n) = BIAS_{KF}(n-1),$$

where $BIAS_P$ represents the bias prediction, $BIAS_{KF}$ represents the bias estimate update and $n$ is discrete time.

2. State the estimate error covariance prediction:

$$P_P(n) = F^2 \times P_{KF}(n-1) + Q,$$

where $P_P$ represents the bias estimate error covariance prediction, $P_{KF}$ represents the bias estimate error covariance update and $Q$ represents covariance function of the system error.
3. Calculate the Kalman gain:

\[ G(n) = P_p(n) \times (P_p(n) + S)^{-1}, \]  

(7)

where \( G \) represents the Kalman gain. \( S \) represents covariance function of the measurement error.

4. Update the state estimate:

\[ \text{BIAS}_{KF}(n) = \text{BIAS}_p(n) + G(n) \times [\text{BIAS}_m(n) - \text{BIAS}_p(n)], \]  

(8)

where \( \text{BIAS}_m \) represents the bias measurement.

5. Update the estimate error covariance:

\[ P_{KF}(n) = (1 - G(n)) \times P_p(n). \]  

(9)

It is assumed that the variation of the real bias within each hour is negligible, and the initial estimator for mean field radar rainfall logarithmic bias and its error variance are assumed to equal their updated values, which are, respectively, \( \text{BIAS}_{KF}(0) \) and \( P_{KF}(0) \).

### 3.2 Intensity–duration threshold identification methods

Rainfall thresholds for the possible initiation of debris flows are identified according to the \( I-D \) power law relationship (Guzzetti et al., 2007), it is defined as follows:

\[ I = \alpha D^{-\beta}. \]  

(10)

Calculating the event duration \( D \) and the average intensity \( I \) requires the start and end times of the rainfall event. The duration and intensity of each debris flow can be directly identified with the time-sequential radar rainfall estimate. These times are determined by an interval of at least 24 h, rain rates of less than 0.1 mm h\(^{-1}\) (Guzzetti et al., 2008; Murra et al., 2014) or corresponding radar reflectivity of less than 10 dBz to separate two consecutive rainfall events. The parameters of \( \alpha \) and \( \beta \) are estimated with the frequentist method (Brunetti et al., 2010).

In order to illustrate the impacts of radar rainfall estimate on \( I-D \) threshold, basic procedures of the frequentist method are applied to radar rainfall accumulation and are described below:

i. Radar-identified rainfall durations and average intensities are log transformed as \( \log(I) \) and \( \log(D) \). Both of them are fitted by least-squares method to form a linear equation as \( \log(I_f) = \log(\alpha_{50}) - \beta \log(D) \), where \( \alpha_{50} \) and \( \beta \) are the fitted intercept and slope, respectively.

ii. For each debris flow, the difference \( \delta(D) \) between the actual rainfall average intensity \( \log[I(D)] \) and the corresponding fitted intensity value \( \log[I_f(D)] \) is calculated: \( \delta(D) = \log[I(D)] - \log[I_f(D)] \).

iii. The probability density function (PDF) of the \( \delta(D) \) distribution is determined through kernel density estimation and furthermore fitted with a Gaussian function, which is defined as

\[ f(\delta) = a \times \exp\left(-\frac{(\delta - b)^2}{2c^2}\right), \]  

(11)

where \( a > 0, c > 0 \) and \( a, b \) and \( c \in R \).
iv. The threshold for expected minimum exceedance probability ($P_{mep}$) is determined by PDF function, as

$$\int_{-\infty}^{\delta_{mep}} f(\delta)d\delta = P_{mep},$$

(12)

where $\delta_{mep}$ is the intercept parameters. $\delta_{mep}$ can be resolved through Eq. (12) for given $P_{mep}$ and then the $\alpha_{mep}$ corresponding to the $P_{mep}$ is calculated as

$$\alpha_{mep} = \alpha_{50}\exp(\delta_{mep}).$$

(13)

Finally, $\alpha_{mep}$ and $\beta$ are the best-fitted parameters for exceedance probabilities $P_{mep}$.

The minimum exceedance probability is set to 5 % for this study.

4 Events, results and discussion

Six debris-flow-triggering rainfall events which occurred in the area of study between 2012 and 2014 are analyzed. Those events happened at the most severely earthquake-affected region during rainy season and triggered a total of 519 debris flows that caused casualties and extensive property damage. Table 2 summarizes the characteristics of the rainfall events. Three events occurred in August, two events occurred in July and one occurred in June. These events are deemed to be representative of the debris-flow-triggering precipitation in the region during the rainy season. The event duration and maximum rainfall accumulation are also retrieved by the rain gauge nearest to debris flow location and radar observations. The identification of the rainfall event was determined by an interval of at least 24 h, during which the rain rate is less than 0.1 mm h$^{-1}$ (Guzzetti et al., 2008; Marra et al., 2014). Table 2 indicates that the durations and rainfall accumulations identified by gauge and radar are different due to the precipitation type and density of rain gauges. The identification differences of event nos. 1, 2 and 6 between gauge and radar are not as large as event nos. 3, 4 and 5. From Fig. 7, showing radar-estimated rainfall accumulation for the six rainfall events (the improving measures described below are applied in Fig. 7), it can be seen that the precipitation of event nos. 3, 4 and 5 is dominated by convection and the strong core of rainfall regions is located in the high-altitude area where rain gauges are relatively scarce. A few debris flow events occurred in the long range, approaching radar detection edges, while the rainfall measured there was low. This may be caused by the decreasing resolution at long radial range. In following section, rainfall estimation accuracy, $I-D$, the distance and height are considered as evaluation factors to assess the radar-based rainfall estimate.

Considering the accuracy and robustness of the $I-D$ threshold of the debris flow are determined by the accuracy of rainfall observation and positioning, a series of processing steps including hybrid scan, VPR correction, a combined $R-Z$ relationship and mean bias adjustment are performed on six rainfall events to improve the accuracy of radar-based accumulated rainfall. In order to evaluate the overall performance and verify the impact on $I-D$ threshold due to rainfall accumulation accuracy, the assessment was performed on three scenarios of radar-based estimates: scenario I, the estimate from raw data of hybrid scan without VPR and bias adjustment; scenario II, the estimates with VPR adjustment after scenario I; and scenario III, the estimates with rainfall bias correction after scenario II. According to rainfall estimate evaluation, $I-D$ thresholds are derived from those scenarios and also assessed with regard to accuracy and spatial resolution.

4.1 Assessment of rainfall estimation accuracy

The accuracy of the radar-based rainfall event accumulation is assessed with the rain gauge observations. In order to perform an evaluation, a set of criteria is calculated including normalized standard error (NSE), normalized mean bias (NMB) and correlation coefficient (CORR), defined as

$$\text{NSE} = \frac{\frac{1}{N}\sum_{i=1}^{N}|r_i - g_i|}{\frac{1}{N}\sum_{i=1}^{N}g_i} \times 100 \%,$$

(14)

$$\text{NMB} = \frac{\frac{1}{N}\sum_{i=1}^{N}(r_i - g_i)}{\frac{1}{N}\sum_{i=1}^{N}g_i} \times 100 \%,$$

(15)

$$\text{CORR} = \frac{\frac{\sum_{i=1}^{N}(g_i - \bar{g})(r_i - \bar{r})}{\sqrt{\sum_{i=1}^{N}(g_i - \bar{g})^2}(\sum_{i=1}^{N}(r_i - \bar{r})^2)}}.$$

(16)

where NMB and NSE are in percent, CORR is dimensionless, $r_i$ and $g_i$ represent the rainfall accumulation from radar and gauge and $N$ is the total sampling number. The statistical criteria comparisons between rain gauges and the three radar estimate scenarios are shown in Table 3, and the scatter plot of radar-based estimates and rain gauge rainfall observations is shown in Fig. 8. The comparison of scenario I indicates that the NSE, NMB and CORR of the study areas are 50.7, −41.1 % and 0.78, respectively. The radar-based rainfall is underestimated. The linear ratio is estimated from linear regression of radar rainfall estimation and rain gauge observation, with the predefined intercept of zero. The linear ratio approximates to 1 when radar-based rainfall estimation is consistent with rain gauge observation. The linear ratio of rainfall observation between radar and gauge for scenario I is 0.51, as shown in Fig. 8a. The reason for the underestimation is the systematic bias and uncertainty of reflectivity on the ground. From the comparison of two types of regions, it can be observed that the NSE, NMB and CORR of region type I are relatively better than region type II. It is revealed that improved measures are needed for the hybrid scan estimate.

The comparison of scenario II indicates that the NSE, NMB and CORR for the study areas are 46.1, −18.6 % and
Figure 7. Images of radar-estimated rainfall accumulation for the six rainfall events (a–f). Circles represent the location of triggered debris flows. Events are shown in chronological order: (a) 9 July 2012; (b) 21 July 2012; (c) 17–18 August 2012; (d) 19 June 2013; (e) 8–12 July 2013; (f) 10, 12 July 2014.

Table 2. Characteristics of the rainfall events.

<table>
<thead>
<tr>
<th>Event no.</th>
<th>Date</th>
<th>Number of triggered debris flows</th>
<th>Event duration by rain gauge (h)</th>
<th>Event duration by radar (h)</th>
<th>Max. rainfall accumulation by rain gauge (mm)</th>
<th>Max. rainfall accumulation by radar (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9 Jul 2012</td>
<td>9</td>
<td>12</td>
<td>11</td>
<td>17.5</td>
<td>29.6</td>
</tr>
<tr>
<td>2</td>
<td>21 Jul 2012</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>29.3</td>
<td>23.6</td>
</tr>
<tr>
<td>3</td>
<td>17–18 Aug 2012</td>
<td>200</td>
<td>7</td>
<td>49</td>
<td>19.2</td>
<td>195.8</td>
</tr>
<tr>
<td>4</td>
<td>19 Jun 2013</td>
<td>15</td>
<td>5</td>
<td>12</td>
<td>55.3</td>
<td>101.8</td>
</tr>
<tr>
<td>5</td>
<td>8–12 Jul 2013</td>
<td>261</td>
<td>55</td>
<td>73</td>
<td>562.2</td>
<td>416.9</td>
</tr>
<tr>
<td>6</td>
<td>10–12 Jul 2014</td>
<td>25</td>
<td>20</td>
<td>21</td>
<td>28.5</td>
<td>17.8</td>
</tr>
</tbody>
</table>

0.80, respectively. This is an improvement over scenario I. The radar-based rainfall is also underestimated through the VPR adjustment, and the linear ratio of rainfall observation between radar and gauge is 0.76, as shown in Fig. 8b. This means rainfall biases still exist in the estimate. The NSE and CORR of region type I are also slightly better than region type II.

The comparison of scenario III indicates that the NSE, NMB and CORR of the entire study area are 44.0, 1.91% and 0.84, respectively. The linear ratio of rainfall observation between radar and gauge is 0.98, as shown in Fig. 8c, and this means the consistency between rainfall and radar observation is achieved through the KF-based bias correction.

Figure 9 shows the average and covariance of bias estimation by KF and mean field bias method for six rainfall events. The CORR and NSE improvement also verifies the efficiency of the KF for radar QPE in mountainous areas. Kalman filter-
Figure 8. Scatter plots of radar and rain gauge event–rainfall accumulations. (a) Scenario 1: radar estimate from hybrid scan. (b) Scenario 2: radar estimate from hybrid scan and VPR. (c) Scenario 3: radar estimate through the hybrid scan, VPR and bias correction.

Table 3. The comparison of radar and rain gauge for each estimate scenario.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Scenario I (hybrid scan)</th>
<th>Scenario II (VPR)</th>
<th>Scenario III (bias adjustment)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Region type I</td>
<td>Region type II</td>
<td>All study regions</td>
</tr>
<tr>
<td>NSE (%)</td>
<td>46.4</td>
<td>50</td>
<td>50.7</td>
</tr>
<tr>
<td>NMB (%)</td>
<td>−40.9</td>
<td>−42.8</td>
<td>−41.1</td>
</tr>
<tr>
<td>CORR</td>
<td>0.80</td>
<td>0.77</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Figure 9. The average and covariance of bias estimation by Kalman filter and mean field bias method for six rainfall events.

Table 4. The parameters of Gaussian fitting, which are used by the frequentist method to account for $I–D$ threshold.

<table>
<thead>
<tr>
<th>Parameters of Gaussian fitting</th>
<th>Scenario I</th>
<th>Scenario II</th>
<th>Scenario III</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>3.144</td>
<td>2.55</td>
<td>2.22</td>
</tr>
<tr>
<td>$b$</td>
<td>0.011</td>
<td>0.003</td>
<td>−0.003</td>
</tr>
<tr>
<td>$c$</td>
<td>0.1273</td>
<td>0.1578</td>
<td>0.1868</td>
</tr>
</tbody>
</table>

The radar rainfall estimates with high spatial resolution can retrieve rainfall duration and average intensity for each rainfall-triggered debris flow, so an abundant of sample data are captured to induce the $I–D$ relationship. Scattering distributions of event duration–intensity for the three radar-estimated scenarios is shown in Fig. 10. Comparisons of scatter distributions between scenarios I, II and III indicate that the average rainfall intensity and duration are incrementally increased when applying the improvement measures. The PDF estimations reveal that the number of positive differences $\delta(D)$ is more than the number of negative differences. This can be accounted for by storm triggering, which is relatively dominant. The parameters of the Gaussian function are summarized in Table 4. Parameter $a$ incrementally decreases. When applying the improvement measures, parameter $c$ has the opposite trend and parameter $b$ randomly changes in a small range around zero.

The $I–D$ threshold derived from the scenario III is $10.1D^{-0.52}$. It is higher than the other two $I–D$ thresholds derived from scenario I and scenario II, due to application of accuracy improving measuring.
4.3 Comparison with intensity–duration thresholds from rain gauge observations

In order to analyze the impact of the spatial sampling variability on identification of $I$–$D$ threshold for radar estimates and rain gauge observations, $I$–$D$ thresholds are derived from the rain gauge nearest to the debris flow and radar estimates at the corresponding co-location of the rain gauge (Marra et al., 2014). There are some same predefined conditions for comparison: (1) duration times are identified separately by two kinds of sensors, rainfall duration time is required to be more than 1 h and minimum mean rainfall rate is 0.1 mm h$^{-1}$. (2) The maximum distance from debris flow location is less than 10 km. (3) The identification of $I$–$D$ threshold is calculated from frequentist methods with exceedance probabilities of 0.5%. Firstly, the event’s rainfall accumulation is compared between rain gauge observations nearest to the location of debris flows and radar estimates at the location of the corresponding rain gauge. The scatter plot of rain gauge and radar estimates is shown in Fig. 11. The corresponding metrics are calculated. The CORR is 0.88, NMB is 17.07%, NSE is 28.32% and the linear ratio is 1.13, indicating that rainfall observations from the rain gauge nearest to the debris flow location and radar estimates at co-location have the tendency of consistency. The $I$–$D$ thresholds are derived from rain gauge and radar estimates. Scatter plots of $I$–$D$ pairs are shown in Fig. 12. The $I$–$D$ threshold estimated from rain gauges is $I = 5.1D^{-0.42}$. The other $I$–$D$ threshold estimated from radar is $I = 5.8D^{-0.41}$. Both $I$–$D$ thresholds seem slightly lower than $I = 10.1D^{-0.52}$, since the scarce gauge network did not capture the strong core of rainfall which triggered the debris flow. It is interesting to note that $I$–$D$ thresholds of both radar and rain gauge are very similar, although there are some measurement errors between them, as shown in Fig. 11.
4.4 Impact of rainfall spatial variation on intensity and duration

The accumulated rainfall, duration and rainfall intensity identified from the nearest rain gauge probably are different from the realities occurred at the debris flow location, since the rainfall varies in space especially for convective precipitation with sharp variation in short distance. The observed rainfall differences rely on the distance from the nearest rain gauge to the debris location and could be considered as rainfall spatial change. To this end, relative changes of the accumulated rainfall, duration and rainfall intensity versus distance are calculated from the comparisons with the radar-based estimate at the location of debris flow. The metrics for evaluating relative change versus distance are defined in Table 5. There are also some predefined conditions for the comparison of relative changes versus distance. (1) The radar rainfall estimations used for comparison are all from scenario III. (2) The radar rainfall estimations and duration identification at the debris flow location are considered as the referred value. (3) The maximum distance from debris flow location to the nearest rain gauge is predefined within 10 km and the distance resolution is set equal to the range resolution of 300 m of the China New Generation Doppler Weather Radar (CINRAD). (4) In order to assess the rainfall spatial variation using a multi-sensor, the radar-based estimate at the co-location of the nearest rain gauge, as well as rain gauge observations, is also compared with the radar-based estimate at the location of debris flow.

The metrics of accumulated rainfall relative change (ARRC), duration relative change (DRC) and rainfall intensity relative change (RIRC) are calculated for the nearest rain gauge and radar estimate at the co-location. The results of ARRC, DRC and RIRC versus distance are shown in Fig. 13. The main findings from the evaluation results are summarized as follows:

1. The results of ARRC, DRC and RIRC all have an enlarging tendency along with the increasing distance. The maximum ARRC, DRC and RIRC for rain gauge observations are 42.2, 41.67 and 55.88 %, respectively. The
maximum ARRC, DRC and RIRC for radar-based estimate at the co-location of the nearest rain gauge are 43.33, 41 and 45.2 %, respectively.

2. Nonlinear regression is applied for ARRC, DRC and RIRC versus distance to investigate the average tendency, as shown in Fig. 13. The regression curves of ARRC and DRC for rain gauge and radar are similar, within 10 and 4 km, respectively, indicating the observed difference as a function of distance is dominated by the natural spatial variability and the potential impact from differences in rainfall estimates coming from different sensors is secondary, especially for estimating duration.

It is clear from the above discussion that the rainfall estimation accuracy and spatial variation impact the identification of $I$–$D$ threshold. We further take the $\alpha$ and $\beta$ estimated from scenario III as a reference value and calculate the relative change of $\alpha$ and $\beta$ for each scenario, as shown in Table 6. The relative change of $\alpha$ for scenarios I, II and III is $-24.5$, $-13.8$ and $0 \%$, respectively. The relative change of $\beta$ for scenarios I, II and III is $-28.8$, $-17.3$ and $0 \%$, respectively. It is indicated that improving the accuracy of rainfall estimate is able to decrease the relative changes of $\alpha$ and $\beta$.

Concerning rainfall spatial variation, the relative change of $\alpha$ for the nearest gauge observation and radar-based estimate at the co-location is $-49.5$ and $-42.6 \%$, respectively. The relative change of $\beta$ for the nearest gauge observation and radar-based estimate at the co-location is $-19.5$ and $-21.2 \%$, respectively. The relative change of $\alpha$ is remarkably larger than the one derived from radar-based estimate on the debris flow location, but the differences of $\alpha$ and $\beta$ for rain gauges and radar-based estimate at the co-location are not significant.

### 4.5 Comparison with previous results

The $I$–$D$ threshold for the study regions is compared with other global and regional thresholds in the literature. It can be seen from Fig. 14 that the threshold obtained in this work (red in Fig. 14) falls in the range of other $I$–$D$ thresholds. The results were also compared with the rainfall thresholds previously proposed in the Wenchuan earthquake area (Tang et al., 2012; Zhou and Tang, 2014; Guo et al., 2016a). Our result lies in the middle range of them. The difference comes from the database we used, the radar data which are used to fill the observation gap of rain gauges, and the identification method of $I$–$D$ threshold that was also different due to a different exceedance probability. The $I$–$D$ threshold of this study was cross-checked with that proposed in the area at-
5 Summary

The main purpose of this paper is to evaluate the debris flow occurrence thresholds of the rainfall intensity–duration in the earthquake-affected areas of Sichuan over the rainy seasons from 2012 to 2014. The paper calculates the intensity–duration threshold from radar-based rainfall estimates, which is different from the common method of using rain gauge observation. Radar observations have high spatial resolutions sensitive to convective precipitation, which is a critical issue for rain gauge observation due to its scarcity and low-altitude deployment over mountainous areas. However, the accuracy of radar-based QPE over complex areas is affected by the terrain and remains a challenge for hydrological application. The following work was done to draw the conclusions.

a. Two S-band Doppler radars covered the study area. Radar observations for six rainfall events were processed with a series of mountain-oriented QPE algorithms, including a terrain-adapted hybrid scan, VPR correction, a reflectivity mosaic, a combination of $R-Z$ relationships a rainfall bias correction. Three types of estimation from radar are performed and compared with rain gauge observations to validate the accuracy. The results show that the combination of all correction procedures reduces the bias to 1.91% and the NSE to 44% and improves the correlation coefficient to 0.84 and the linear ratio to 0.98.

b. Intensity–duration rainfall thresholds for the triggering debris flow are calculated with a frequentist approach. The $I-D$ threshold of $I = 10.1D^{-0.52}$ is derived from the KF-corrected radar estimates. The accumulated rainfall is lower than rain gauge observations and the derived $I-D$ is also underestimated. The hybrid scan, VPR correction and combination of $R-Z$ relationship are strongly required.

c. The $I-D$ deduced from rain gauge observations nearest to the occurrence of debris flow is highly similar to the one deduced from the radar estimates at the same location as rain gauge: $I = 5.1D^{-0.42}$ and $I = 5.8D^{-0.41}$ respectively. These $I-D$ thresholds are underestimated due to the rainfall spatial variation and the noncontinuous sampling effect.

Finally, it is clear that radar-based rainfall estimates and thresholds supplement the monitoring gaps of EWS where rain gauges are scarce. A better understanding of the relationship between rainfall and debris flow initiation for earthquake-affected areas can be gained by improving the spatiotemporal resolution and low-elevation-angle coverage of radar observation, especially for monitoring the convective storm occurring at the mountains.

Data availability. The data used in this research can be made available on demand. Please contact the authors for requests.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue “Landslide early warning systems: monitoring systems, rainfall thresholds, warning models, performance evaluation and risk perception”. It is not associated with a conference.

Acknowledgements. This research was supported by the National Natural Science Foundation of China (no. 41505031), the Science and Technology Support Project of Sichuan (no. 2015SZ0214), China Meteorological Bureau Meteorological Sounding Engineering Technology Research Center funding, China Scholarship Council (no. 201508515021), and scientific research funding of CUIT (no. J201603). We thank the weather bureau of Sichuan for the data services.

The participation of VC is supported by the US National Science Foundation through the Hazard SEES program.
Edited by: Samuele Segoni
Reviewed by: three anonymous referees

References


