Nat Hazards (2018) 93:531–546
https://doi.org/10.1007/s11069-018-3313-0

ORIGINAL PAPER

Determination of probabilities for the generation of high-discharge flows in the middle basin of Elqui River, Chile

Iván P. Vergara Dal Pont[1,2] · Fernanda A. Santibañez Ossa[3] · Diego Araneo[1,4] · Francisco J. Ferrando Acuña[3] · Stella M. Moreiras[1,4]

Received: 24 July 2017 / Accepted: 9 April 2018 / Published online: 18 April 2018

Abstract The probabilities for the generation of hyperconcentrated flows, and debris and mud flows in the middle basin of Elqui River (Chile) are determined. The objective was achieved collecting, for a period of 14 years, the precipitation events generating high-discharge flows, as well as the larger precipitation events that did not generate this process. For each of these events, data of peak 1-h storm precipitation, temperature (representing the zero-isotherm altitude) and antecedent precipitation of 1, 5 and 10 days were collected from three meteorological stations. Initially, an ordinal logistic regression model for each antecedent precipitation was fitted, but all were discarded due to the low significance of these variables in the generation of the models. This result allowed to infer that the high-discharge flows of the region are generated mainly by runoff and not by deep-seated or shallow landslides. Subsequently, a new model with the remaining variables was performed, which was statistically validated. From this, it was considered prudent to take as thresholds for the occurrence of hyperconcentrated flows, and debris and mud flows, their respective probabilities of 50%. For these thresholds, the model had an efficiency in the prediction of high-discharge flows of 90%. Finally, the partial correlation coefficients of each significant predictor variable with respect to the dependent were calculated, establishing that the temperature has greater influence than the peak 1-h storm precipitation.

Keywords Meteorological thresholds · Geoclimatic hazard · Ordinal logistic regression · Debris flows

✉ Iván P. Vergara Dal Pont
ivergara@mendoza-conicet.gob.ar

[1] CONICET-IANIGLA, Ruiz Leal s/n, Mendoza, Argentina

[2] Universidad Nacional de Río Negro, Belgrano 526, Viedma, Argentina

[3] Universidad de Chile, Portugal 84, Santiago de Chile, Chile

[4] Universidad Nacional de Cuyo, Centro Universitario, Mendoza, Argentina

1 Introduction

The high-discharge flows (HDFs) are a hydro-sedimentological process that is subdivided depending on the concentration and size of sediments in: debris flow, mud flow, hyperconcentrated flow and water flow (Pierson 2005). These phenomena constitute a relevant natural hazard in the generation of risk situations worldwide, a scenario that, according to current studies, is increasing due to population growth and the human settlement in marginal areas (Leroy 2006; Petley 2012).

In the Elqui valley, the HDFs have caused considerable economic and human losses since the region began to inhabit in the nineteenth century (e.g., Conte Nadeau 1986; Graña Pezoa 2007; Pérez et al. 2008); however, sedimentological–palynological research indicates that these phenomena occurred with irregular frequency since the Late Pleistocene (Antinao et al. 2015; Maldonado and De Porras 2015).

Many scientists around the world have researched the thresholds needed to generate HDFs: some through physical models (Crosta and Frattini 2003; Wilson and Wieczorek 1995) and others with empirical approaches with statistical criteria (e.g., Guzzetti et al. 2008; Brunetti et al. 2010; Giannecchini et al. 2016) or without statistical criteria (Cannon et al. 2008; Coe et al. 2008). Thresholds that are set by physical models require data that are difficult to obtain at scales greater than the local (i.e., soil density, friction angle, cohesion, hydraulic conductivity, failure depth and water table depth); on the other hand, thresholds based on empirical approaches are adapted to different scales, but usually use only duration and total precipitation as variables (or intensity resulting from the division of these parameters). These latter approaches do not usually use contributing factors such as antecedent precipitation or zero-isotherm altitude (ZIA), which is important when it is below the maximum altitude of the ravine.

In the Elqui basin, there are no defined thresholds for the generation of HDFs, and throughout the Argentine-Chilean Andes they are scarce and were developed with little quantity and quality of meteorological data (Hauser 1985; Moreiras 2005). For the region of Santiago de Chile, Hauser (1985) established, based on a statistical analysis of daily rainfall data, that there is a 50% of probability that rainfall over 60 mm/24 h triggers debris and mud flows. For the Mendoza River Valley (Argentina), Moreiras (2005) established a threshold for debris flows between 6.5 and 12.9 mm/24 h using the same data types and an influence of the antecedent precipitation of 5 days for 60% of the cases.

A research of thresholds with a novel method and realized for the Andean region of Santiago de Chile, is that of Sepúlveda and Padilla (2008). In this, through a binary logistic regression, different probabilities were defined for the occurrence of debris flows, mud flows and hyperconcentrated flows. The analysis used the daily precipitation of the meteorological station with greater predictability between two, the antecedent precipitation with greater predictability and the ZIA at daily resolution. As a result, they recorded that 33% of the flows occurred show a probability greater than 50%; the authors comment that predictability could have been better if, among other reasons, there would have been greater density of meteorological stations and if they had had a higher temporal resolution.

Regarding the ways of acquiring precipitation data, Marra et al. (2014) exposed the advantages of weather radars over the rain gauge networks; however, this technology does not exist in the Argentine-Chilean Andes. Santos et al. (2015) and Lauro et al. (2017) showed that methods from passive microwave and infrared data such as the Climate Prediction Center morphing (CMORPH; Joyce et al. 2004) are useful for measuring debris flow thresholds in the Central Argentine Andes, where there are few stations and convective precipitation. However, the authors observed that these, at least in mountainous

sectors such as the region, inadequately estimate precipitation. Santos et al. (2015) also used the Weather Research and Forecasting (WRF; Powers et al. 2017) and the Eta-Meteorological Regional Program (Eta-MRP; Seluchi and Chou 2001) numerical models, obtaining even worse rainfall estimates.

In light of the aforementioned, the objective of the present paper is to fit a statistical forecasting model, which provides the probability of HDFs occurrence, from available weather information in stations of the region. If fitted and validated correctly, the interesting thing about the predictive model would be its regional scale and its simplicity of using easy-to-obtain variables. It could be used by government institutions to decide whether or not to implement the early warning system for a given weather forecast.

1.1 Study area

The studied sector is located within the middle basin of Elqui River, in central-northern Chile and 40 km from the Pacific Ocean (Fig. 1). Its area is 2000 km^2 and has an altitude range between 400 and 3900 m.

The route CH-41 communicates the region with Argentina and also with the main cities of the basin like La Serena and Vicuña. The population is 26,900 inhabitants and is increasing to an average rate of 110 inhabitants/year (Instituto Nacional de estadísticas 2017). The economy is based primarily on agriculture and tourism.

Being a mountain region, the population is usually settled in geo-climatically hazardous areas, so that HDFs frequently generate material damages in houses and roads (Fig. 2), as

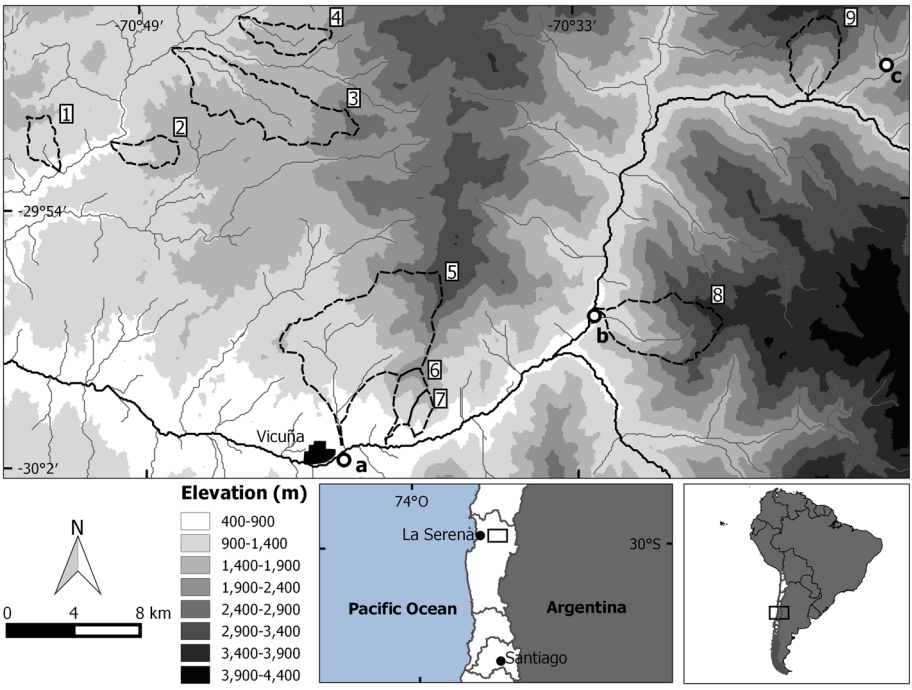

Fig. 1 Study area and its location in Chile. Activated ravines: (1) Talcoma I, (2) Talcoma II, (3) Viñita, (4) El Sauce, (5) Leiva, (6) Miraflores, (7) Cebollín, (8) Seca and (9) Culebra. Meteorological stations: **a** Vicuña, **b** Rivadavia and **c** Llanos de Huanta

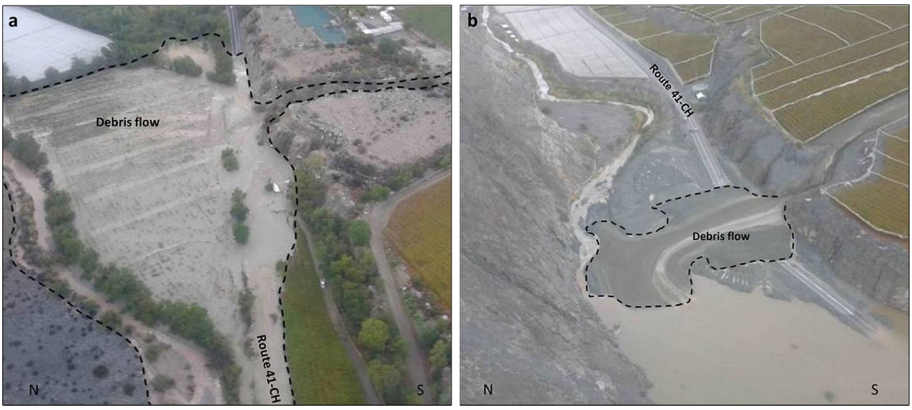

Fig. 2 Photographs as an example of the HDFs studied. **a** Debris flow of Seca Ravine in March 2015. **b** Debris flow of Culebra Ravine in March 2015

well as injuries and deaths. During the period 1997–2016, four deaths due to HDFs were recorded (Pérez 2005). This shows that the studies dedicated to understanding and preventing these phenomena are very important for the security and economy of the region.

1.2 Geology and geomorphology

The study area is located within the physiographic subunit of Western Main Cordillera, which is located between the N-S direction fault systems of Vicuña-San Felix and Baños del Toro Chollay (Aguilar et al. 2013).

The main rivers run through narrow valleys that present up to 2000 m of incision and that upstream is developed in morphological areas characterized by slopes of strong gradient and rocky outcrops. The high gradient values are concentrated in the slopes of the valleys, while the flat interfluves are dominated by low slope pediments (Aguilar et al. 2013).

The unconsolidated stratified units correspond to fluvial and alluvial deposits, which sometimes occur in terraces, and which are filling the beds of the Elqui River and the tributary ravines. The consolidated stratified units are volcanic and sedimentary rocks of Cretaceous to Jurassic age, and sedimentary triassic rocks of continental origin composed of conglomerates, sandstones and carbonaceous shales (Rauld Plott et al. 2012).

In addition, plutons of Triassic to Paleozoic age outcrop, which are distributed in strips with preferential N-S orientation and that together form the Elqui-Limarí Batholith (Rauld Plott et al. 2012).

1.3 Climate and climate variability

The climate of this sector of Chile is cold semiarid with dry summers, but above 2750 m it becomes tundra with dry summers type (Sarricolea et al. 2016). The climatic characteristics are conditioned by two main atmospheric factors: the southeastern Pacific semipermanent anticyclone and frontal systems associated with extratropical migratory depressions (Fuenzalida 1982). The average annual precipitation, for the period 1950–2000, varies between 70 and 97 mm depending on the station analyzed (Zavala and Trigos 2008), and on average 85.9% is concentrated between May and August (Vuille and

Milana 2007). The average daily potential evapotranspiration oscillates between 5.5 mm in December and January, and 1.7 mm in June (CEAZA 2017).

The region is dominated by an almost permanent temperature inversion maintained by generalized subsidence, forced by the southeastern Pacific semipermanent anticyclone. It has an average thickness of 250 m and starts between 700 and 1000 m during the warm semester, and between 500 and 700 m during the cold semester (Romero et al. 1988). This inversion, which usually inhibits the development of convective activity and rainfall, can be lifted and interrupted during occasional frontal passages that affect the region mainly in winter.

El Niño Southern Oscillation (ENSO) represents the inter-annual cycle that most influences the variability of atmospheric conditions in this region, increasing rainfall and temperatures during its positive phase (recurrence of 3–5 years; Rutllant and Fuenzalida 1991; Rosenblüth et al. 1997). The ENSO and the number of landslides present a positive correlation in the area (Ortlieb and Vargas 2015; Sepúlveda et al. 2006).

In the Elqui valley, precipitation presents a linear trend of − 0.56 mm/year for the period 1869–2004 (p = 0.001; Vuille and Milana 2007). The trend would continue during the twenty-first century according to the projections of future climate change (Vera et al. 2006). Regarding the ZIA, applying an exponential filter, Carrasco et al. (2008) measured a positive trend of 2.3 m/year (p = 0.05) in central Chile, during the period 1958–2006. Projections of future climate change indicate that the average daily temperature during the period 2071–2100, compared to the period 1961–1990, would increase for the four seasons, from 2–3 to 3–4 °C depending on the scenario of greenhouse gas emissions (Geophysics Department of the University of Chile 2006).

2 Methods

2.1 Statistical analysis

In this study, we use the logistic regression method (Hosmer and Lemeshow 2000). This multivariate statistical method has been widely used for different geoscientific applications, such as the determination of rainfall thresholds to trigger landslides (Sepúlveda and Padilla 2008; Giannecchini et al. 2016), the annual debris flows occurrence (Pavlova et al. 2014) and the production of landslide susceptibility maps (Carrara et al. 2008).

The logistic regression model allows defining a mathematical expression that relates multiple contributing factors and triggers to an occurrence probability of an HDF. The general expression for the logistic regression model is:

$$\ln\left[\frac{p}{(1-p)}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_n X_n$$

where p is probability that an event occurs, $\ln\left[\frac{p}{(1-p)}\right]$ log odds ratio, X's input predictor variables measured. β's are regression coefficients that establish the relationship between the variables.

The present paper was limited to the HDFs of type debris flow, mud flow and hyperconcentrated flow, because they are the most hazardous (Pierson 2005). Among these, the hyperconcentrated flows are those of lower energy and hazardousness (Pierson 2005).

Because HDFs of different magnitude were studied, it was decided to use the ordinal logistic regression, which allows to determine probabilities for different results of an

ordinal categorical distribution as a dependent variable. Thus, for the dependent variable of the logistic regression were assigned a value of 2 for precipitation with debris and mud flows, a value of 1 for precipitation with hyperconcentrated flows and a value of 0 for precipitation without flows.

The possible multicollinearity was evaluated through the coefficient of determination (r^2) of each independent variable with respect to all the others. The problems of multicollinearity exist when some r^2 is greater than or equal to 0.9 (Kleinbaum et al. 1988).

In order to evaluate the validity of the logistic regression model, and considering the number of events ($n = 29$), the method of leave-one-out cross-validation (Arlot and Celisse 2010) was applied. This consists of fitting the model n times removing one event at a time and calculating for each the difference between the probabilities of success, including it and excluding it from the analysis. Then, the arithmetic mean of these differences is made and its value is evaluated.

To measure the contribution of each variable to the model, the partial correlation coefficient is used (Prokhorov 2001). This parameter measures the relationship between two quantitative variables when controlling or eliminating the effect of thirds; its formula for three variables is as follows:

$$r_{YX_1 \cdot X_2} = \sqrt{\frac{\left(r_{YX_1} - r_{YX_2} r_{X_1X_2}\right)^2}{\left(1 - r_{YX_2}^2\right)\left(1 - r_{X_1X_2}^2\right)}}$$

where Y is the dependent variable, X's are the independent variables, and $r_{YX_1 \cdot X_2}$ means the partial correlation between Y and X_1 with fixed X_2. Its expression for more than three variables is:

$$r_{YX_1 \cdot X_2 X_3 \ldots X_n} = \sqrt{\frac{\left(r_{YX_1 \cdot X_3 \ldots X_n} - r_{YX_2 \cdot X_3 \ldots X_n} r_{X_1X_2 \cdot X_3 \ldots X_n}\right)^2}{\left(1 - r_{YX_2 \cdot X_3 \ldots X_n}^2\right)\left(1 - r_{X_1X_2 \cdot X_3 \ldots X_n}^2\right)}}$$

This technique allows checking the contribution of each variable independent of its order of magnitude; therefore, it allows for comparison of the weight of the variables in the model in a better way than using the regression coefficients directly.

All analyses were performed with the Excel add-in Real Statistics Resource Pack.

2.2 Variables used

Since the objective of this paper is to implement a predictive model of HDFs occurrence in an area considered homogeneous, only variables with temporal variation without spatial variation or with random spatial variation will be used. Due to this premise, variables such as anthropogenic impact by crops or structural measures, and geotechnical, geological and morphometric properties were not taken into account as predictors. As will be seen later, the climatic variables used were considered without spatial variation or with random spatial variation.

2.2.1 Precipitation

In the study area, there are three meteorological stations at 5–15 min temporal resolution, one is in the talweg of the Huanta Ravine, while the other two are in the talweg of the Elqui River (Fig. 1 and Table 1). Considering that the precipitation regime is almost homogeneous, the model was fitted using all the stations simultaneously; in this way, it was

Table 1 Name, coordinates, elevation and record of meteorological stations

Gauge	Coordinates (°)	Elevation (m)	Record
Vicuña	30.038S, 70.697W	634	02/2004–
Rivadavia	29.962S, 70.539W	900	09/2010–
Llanos de Huanta	29.827S, 70.354W	1696	07/2012–

All belong to Centro de Estudios Avanzados en Zonas Áridas (CEAZA)

assumed that the water accumulated in each meteorological station during an event depended only on the random distribution that the precipitation had.

For each precipitation event, the peak 1-h storm precipitation was used. This decision was due to the fact that the reported times of the HDFs were always shortly after the hour of maximum precipitation. In addition, this time corresponds to the maximum time of concentration among the activated ravines [data not shown and calculated from a digital elevation model at 12.5 m resolution (Alaska Satellite Facility 2017), and through the California Culverts Practice equation (Hotchkiss and McCallum 1995)]. The time of concentration is the time needed for water to flow from the most remote point in a watershed to the outlet, and it is equivalent to the moment of maximum discharge.

To measure the precipitation events that produced HDFs, the nearest station was used. The distances between the ravines and the stations had an average of 6.46 km and a range from 4.17 to 20.26 km. These values are acceptable if rainfall is considered to be generated by frontal systems with smaller spatial variations than those of convective origin. The distances between the ravines and the meteorological stations were taken from the section of the main channel that is located in the middle of the zone with average slope greater than 20°; this sector is the most suitable for the generation of HDFs according to Hungr (2005).

2.2.2 Temperature

For the generation of HDFs triggered by rain, the ZIA plays a fundamental role in determining the relationship between solid and liquid water precipitating in the ravine. If it is well below the maximum altitude of the ravine, a large amount of water will precipitate in a solid state and the generation of HDFs will be inhibited. The ZIA has an even greater influence if it is considered that the highest slopes are in the higher sectors, and the fact that it is positioned a little below the maximum altitude of the ravine can inhibit the generation of flows, to interrupt the runoff where it has greater competence.

The ZIA in the region, during precipitation events from May to September, has an average of 2300 m and a standard deviation of 250 m (Garreaud 2013). Considering that the maximum altitude of the highest activated ravine reaches 3716 m, it was decided to use this variable for predictive analysis. As a proxy for the ZIA, the average temperature during the peak 1-h storm precipitation at the Vicuña station was used. This meteorological station is the one with the longest temporal record (Table 1). The use of a meteorological station is based on the fact that its thermal variations correctly reflect the variations of the ZIA, when precipitation is present (Garreaud 2013).

Unlike precipitation data that multiple stations were used, for the temperature a single one was used since this variable has a smaller spatial variability and a greater altitudinal

variability, which makes it impossible to use several stations with different heights for the same model.

Due to the delay in the melting of ice crystals, the ZIA is no longer influential when it is 50–200 m above the maximum altitude of the highest activated ravine (Garreaud 1992; White et al. 2010). Therefore, it is important to give an upper limit to the temperature observed at the Vicuña station so that the model does not overestimate it. By the moist adiabatic lapse rate of the area that limit is 19.6 °C ($\sim$ 3900 m) and was used for a case where the average temperature during the peak 1-h storm precipitation exceeded that value (Table 2).

2.2.3 Antecedent precipitation

The antecedent precipitation was used in the predictive model since it can indirectly represent the soil saturation and/or the antecedent snow in the ravine. The first phenomenon is a contributing factor of HDFs since the more saturated the soil is during a rainy event, the less water will infiltrate and more water will drain, increasing erosion and the possibilities of generating HDFs. The second is also a factor, if there is accumulated snow in the ravine during a rain event, it can be totally or partially melted by the rain and increases the streamflow of surface runoff.

For each precipitation event, the antecedent rainfall of 1, 5 and 10 days was collected, to which the potential evapotranspiration until the corresponding event was subtracted. The latest data were calculated by Centro de Estudios Avanzados en Zonas Áridas (CEAZA) through the FAO Penman–Monteith equation (Food and Agriculture Organization 2006) and were implemented to understand how much antecedent precipitation remained in soil at the time of the event. Subsequently, a model for each corrected antecedent precipitation was performed in order to understand which is the most influential.

It is important to note that no antecedent precipitation of more days was analyzed due to the high potential evapotranspiration of the region.

2.3 Precipitation events record

The HDFs record was made through newspapers, printed and oral communications of governmental entities and field interviews with inhabitants near the ravines. The registration began in February 2004, date in which the first meteorological station was put into operation, and ended in May 2017.

The precipitation without flows was selected from the mildest situations of the precipitation with HDFs; this resulted in the peak 1-h storm precipitation having to be greater than 2 mm and the average temperature greater than 5 °C. For these precipitation events, a proportional stratified random sampling was carried out, that is to say that the stations were chosen randomly with an intent to use each the same amounts of times; however, since they have different record lengths, some were used more than others (Table 1).

In total, data were collected for: eight precipitation events with debris and/or mud flows, four with hyperconcentrated flows and seventeen which showed no response (Table 2). Two rainfall generating flows, occurred on 04/21/2004 and 06/06/2011, had to be discarded for lack of operation in the Vicuña station of the rain gauge for the first case, and the thermometer for the second.

Table 2 Precipitation events record

Date	Peak (mm)	Antecedent$_{x-d}$ (mm)			T (°C)	Activated ravines	Gauge	Distance (km)
		1	5	10				
2:15–12/05/2017	9.7	38.31	39.41	39.41	10.57	Cebollín (mf), Miraflores (df) and Leiva (df)	Vicuña	5.69
0:00–12/05/2017	6.85	25.73	25.73	25.73	11.76	Seca (df)	Rivadavia	4.91
14:15–25/03/2015	6.1	22.76	24.02	24.02	16.3	Seca (df)	Rivadavia	4.91
12:40–25/03/2015	5.7	18.1	18.1	18.1	16.73	Talcoma I (df), Talcoma II (df), Viñita (df) and El Sauce (df)	Vicuña	20.26
20:40–24/03/2015	8	6.69	6.69	6.69	16.68	Cebollín (mf), Miraflores (df) and Leiva (df)	Vicuña	5.69
14:30–24/03/2015	7.3	4.46	4.46	4.46	18.6	Culebra (df)	Llanos de Huanta	4.17
12:50–24/03/2015	4.6	3	3	3	18.1	Cebollín (mf), Miraflores (df) and Leiva (df)	Vicuña	5.69
15/01/2013	13	0	0	0	19.6	Culebra (df)	Llanos de Huanta	4.17
19/10/2015	5	0.60	12.48	12.98	10.6	Cebollín (hf), Miraflores (hf) and Leiva (hf)	Vicuña	5.69
08/08/2015	5.7	0.2	0.2	0.2	9.85	Cebollín (hf), Miraflores (hf) and Leiva (hf)	Vicuña	5.69
12/07/2015	4.58	5.95	5.95	5.95	9	Seca (hf)	Rivadavia	4.91
03/05/2005	5.84	0.51	0.51	0.51	13.1	Cebollín (hf)	Vicuña	5.69
25/07/2016	4.9	1.4	1.4	1.4	9.8	–	Rivadavia	–
03/06/2016	3.56	2.10	4.05	12.45	10.1	–	Llanos de Huanta	–
29/05/2016	5.7	6.9	6.9	6.9	12.75	–	Varillar	–
14/10/2015	6.6	10.79	10.79	10.79	5.1	–	Llanos de Huanta	–
13/09/2014	2.1	2.7	2.7	2.7	8.5	–	Rivadavia	–
12/06/2014	5.08	1.01	1.01	1.01	7.7	–	Rivadavia	–
20/07/2013	4.32	1.02	1.02	1.02	5.8	–	Llanos de Huanta	–
18/05/2013	3.1	12.49	12.49	12.49	10.8	–	Vicuña	–
15/07/2011	4.3	4.41	4.41	4.41	7.08	–	Vicuña	–
21/06/2011	3.29	2.01	11.95	11.95	7.8	–	Vicuña	–
18/06/2010	4.32	5.25	10.45	10.45	9.7	–	Vicuña	–
15/05/2010	4.32	12.12	18.87	18.87	11.3	–	Vicuña	–
28/06/2009	3.56	1.01	1.01	1.01	7.8	–	Vicuña	–
27/05/2008	4.82	1.77	1.77	1.77	9.5	–	Vicuña	–
26/07/2006	4.06	4.57	4.57	4.57	7.2	–	Vicuña	–
07/07/2006	3.8	1.27	1.27	1.27	11.5	–	Vicuña	–

Table 2 continued

Date	Peak (mm)	Antecedent$_{x-d}$ (mm)			T (°C)	Activated ravines	Gauge	Distance (km)
		1	5	10				
21/07/2004	6.61	14.72	14.72	14.72	6.1	–	Vicuña	–

For events that generated multiple flows the distance to the nearest ravine was placed; for those who have the same date, the start time was set. mf: mud flow, df: debris flow and hf: hyperconcentrated flow

3 Results

The three fitted models showed a p value lower than 0.05 for the variables temperature and peak 1-h storm precipitation, but higher for all antecedent precipitation (between 0.48 and 0.69). An explanatory variable with a p value greater than 0.05 has low predictive capacity and therefore low significance in the generation of the model (Giannecchini et al. 2016). Considering their p values, the antecedent precipitation variables were excluded, and a new model was made only with the significant variables.

The new model had a mean correct prediction probability of 0.81 (or 81%; Fig. 3). In this calculation, we averaged the probability of success that the model had in each precipitation event, without even defining probability thresholds. It was also decided to consider the probabilities for the precipitation events without HDFs since the false alarms (a type of HDF was predicted and did not occur) are also important as they reduce the credibility of the model.

The logistic regression being of the ordinal type presents two final equations when isolating the probability:

$$p_1 = \frac{1}{1 + e^{-(-1.604 \cdot Pp - 0.759 \cdot T + 16.557)}} \quad p_2 = \frac{1}{1 + e^{-(-1.604 \cdot Pp - 0.759 \cdot T + 19.523)}}$$

where p_1 is probability of not generating a HDF. p_2 is probability of not generating a debris or mud flow. Pp is peak 1-h storm precipitation (mm). T is average temperature in Vicuña station during the peak 1-h storm precipitation (°C).

When the first equation is greater than 0.5, there is more probability that no flows will occur; when the former is less than 0.5 and the second greater than 0.5, there is more probability that hyperconcentrated flows will occur; when the second is less than 0.5, there is more probability that debris or mud flows will occur.

Significant predictor variables, temperature and peak precipitation, did not show multicollinearity problems ($r^2 = 0.26$).

The model chosen was validated with the method of leave-one-out cross-validation. In Fig. 3, it is observed that for each case, the difference between the probability of the fitted event (incorporating it in the model) and the validated one (excluding it) is usually very small. The arithmetic mean of the differences for all events was 0.04 (4%; Fig. 3). This small dissimilarity between the fit and the validation allowed to consider the first as reliable and apt to predict the occurrence of HDFs.

With the certainty that the fitted model presented a correct validation and their predictor variables did not show multicollinearity problems, the thresholds for the different types of HDFs were defined. It was considered appropriate to define as thresholds for the occurrence of hyperconcentrated flows, and debris and mud flows, their respective probabilities

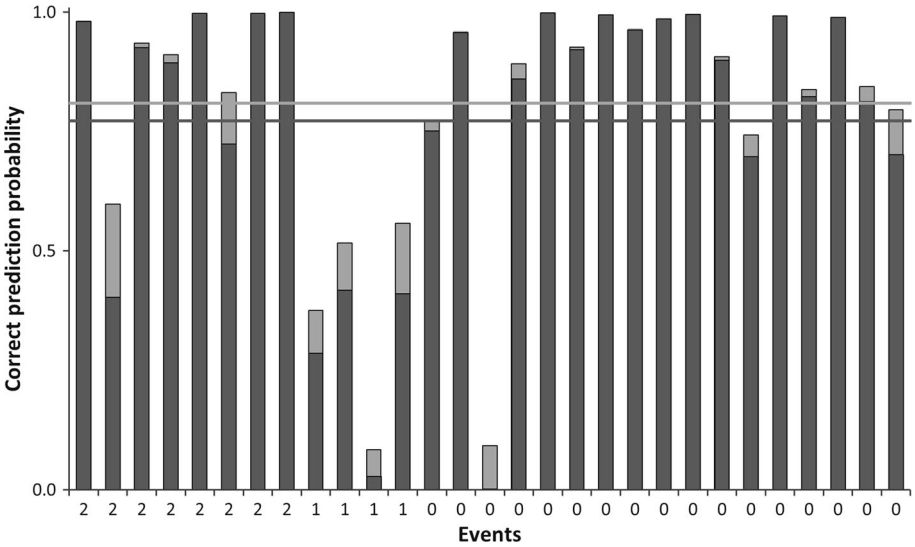

Fig. 3 Correct prediction probabilities for fitted (light gray bars) and validated (dark gray bars) precipitation events. When no (yes) the light gray bar is observed, the prediction between the fitted and validated event coincides (does not match). The dark gray (light) horizontal line is the average correct prediction probability for the validated (fitted) model

of 50% (Fig. 4). For these thresholds, the model had an efficiency in the prediction of HDFs of 90% (ratio expressed as a percentage, between the number of successful predictions and the total number of cases). In 29 cases, it generated only one false alarm and two failed alarms. (A type of flow was not predicted and this occurred.)

The model had a specific efficiency for the hyperconcentrated flows and the debris and mud flows of 50 and 100%, respectively (Fig. 4; ratios expressed as a percentage, between the number of successful predictions for these types of HDF, and the total number of these HDFs).

As expected, from the equations of the final logistic regression it is observed that the two predictor variables have a positive correlation with the HDFs. From the analysis of the partial correlation coefficients between the predictor variables selected and the occurrence of flows, it was found that the temperature has a value of 0.7 and the peak $1-$ h storm precipitation of 0.47.

4 Discussion

The statistical prediction model had a medium specific efficiency for the hyperconcentrated flows, since they are phenomena of less energy and rainfall thresholds. On the contrary, for the debris and mud flows, the model had an excellent specific efficiency since they are phenomena of higher energy and rainfall thresholds. In fact, the minimum probability of success for these HDFs was of 60% (Fig. 4).

It is considered prudent to implement the early warning system (evacuation of the exposed population, road closures, etc.) when the hyperconcentrated flow threshold is surpassed in a weather forecast. This threshold represents a conservative atmospheric

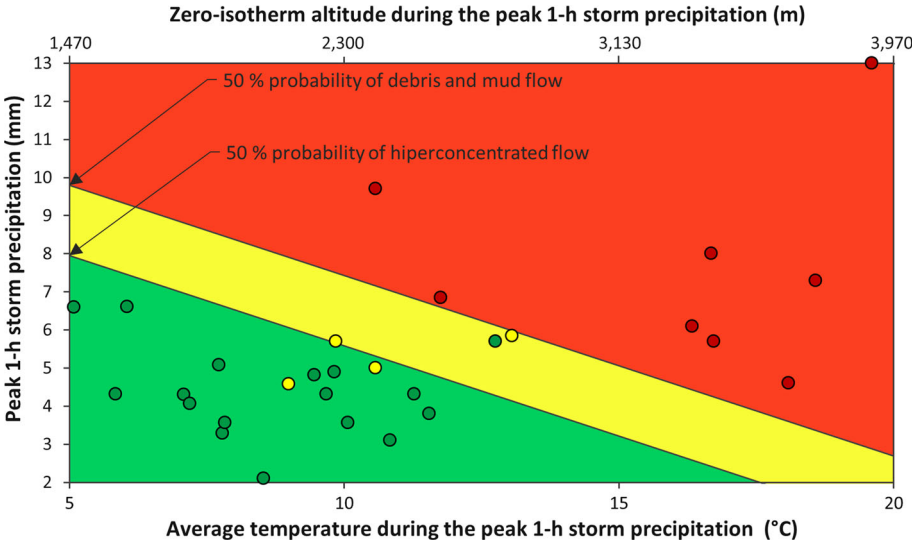

Fig. 4 Graph showing the precipitation events collected and the meteorological thresholds at 50% probability, for hyperconcentrated flows and debris and mud flows. The green circles correspond to precipitation events that did not generate HDFs, the yellow (red) circles to precipitation events that generated hyperconcentrated flows (debris and mud flows)

condition for the debris and mud flows initiation, which historically have been the most destructive HDFs. Similar probabilistic thresholds have been used in recent research (Staley et al. 2017; Giannecchini et al. 2016). The input variables of the model (peak 1 − h storm precipitation and average temperature during the peak) can be easily obtained from the traditional outputs of the regional weather forecasts. This makes the implementation of the early warning system easier.

The three predictions where the probability of success was less than 50% (wrong predictions) are probably due to the unmeasured influence of anthropogenic impact and geological, geotechnical and morphometric properties. Nevertheless, the precipitation events record suggests a preponderance of the climate over the spatial variables; due to when the rains produce HDFs, they can be observed in several ravines (Table 2).

An interesting result is the low significance of the antecedent precipitation in the generation of the models. Although this result was not expected, it coincides with other research (Coe et al. 2008; Kean et al. 2013) where it is shown that in semiarid zones, the occurrence of debris flows does not require soil saturation, because they are initiated by runoff. Considering that the HDFs can be generated without values of high moisture in the soil, it is logical that the model did not detect influence of this variable.

Berti et al. (2012) indicate that intensity-duration thresholds, for debris flows generated by deep-seated landslides, are more uncertain since they do not consider the soil moisture. Our paper shows that soil moisture can be evaluated using as a proxy the antecedent precipitation, in statistical models that allow the use of multiple predictor variables, such as the logistic regression.

The greatest influence of temperature or ZIA, regarding the peak 1 − h storm precipitation, in the generation of HDFs, may be surprising. However, it should be considered that the greatest amount of precipitation events occurs in the cold semester when the ZIA is very low; thus, only the warm semesters with positive precipitation anomaly, or the cold

semesters with positive temperature anomaly, are able to generate HDFs. Considering this important influence of the ZIA, it is probable that the positive phase of ENSO does not generate more HDFs in the region solely by increasing rainfall, as stated by Sepúlveda et al. (2006), but also generate the positive temperature anomaly described by Rosenblüth et al. (1997).

Despite the good results, the model could improve if the temporal data record was extended, and more meteorological stations were placed closer to the headwaters of ravines.

4.1 Comparison with other studies

The improved efficiency of the model presented in this paper compared to that presented in the paper of Sepúlveda and Padilla (2008) may be due to the higher temporal resolution of temperature and precipitation data. In fact, the reported times of HDFs were always shortly after the peak $1-h$ storm precipitation, indicating that daily data attenuate the real intensity. Another difference that could influence was to have used more than one meteorological station, which allowed to use the most suitable for each ravine.

Regarding research that uses intensity and duration, such as Brunetti et al. (2010), the greater probability of prediction found may be due to the different statistical method employed; however, different climates and geologies where the research were conducted make comparison difficult. What is important to note is that the weight of the partial correlation coefficient, between the ZIA (represented by temperature) and the HDFs occurrence, demonstrates the importance of this independent variable when is in proximity or below the maximum altitudes of ravines.

Sepúlveda and Padilla (2008), through the partial correlation coefficients, found that the significant explanatory variables of greater weight for HDFs are the daily precipitation, followed by the ZIA and finally by the antecedent precipitation. This result differs from that of the present study where the triggering rain has a lower weight than the ZIA (temperature), and the antecedent precipitation has no significant influence. The difference can be due to: (1) the sectors studied are different and are distant 390 km, and (2) the variables used differ in temporal resolution. The Andean region of Santiago de Chile has a lower potential evapotranspiration (less arid climate; Sarricolea et al. 2016), which would explain, for example, the different weight of the antecedent precipitation.

5 Conclusions

In this study, an ordinal logistic regression model was generated for the prediction of different HDF types, using two meteorological variables collected in three stations during 14 years. The model had a correct statistical validation and a mean correct prediction probability of 81%.

The study covered an area of 2000 km^2, with large amount of historical and prehistoric landslides. From the model, it was decided to take as thresholds for hyperconcentrated flows, and for debris and mud flows, their respective probabilities of 50%. For these thresholds, the ordinal logistic regression had an efficiency in the HDFs prediction of 90%, and a specific efficiency for the hyperconcentrated flows and the debris and mud flows of 50 and 100%, respectively. It is considered prudent to implement the early warning system when the hyperconcentrated flow threshold is surpassed in a weather forecast.

From the logistic regressions, it was found that the antecedent precipitation of 1, 5 and 10 days did not have a significant influence on the generation of HDFs. This result allowed to infer that the HDFs of the region are generated mainly by runoff and not by deep-seated or shallow landslides.

The partial correlation coefficients among each significant explanatory variable and the occurrence of flows, it was known that the ZIA is more important than the peak $1-h$ storm precipitation. This latter result suggests that the positive phase of the ENSO could not only force an increase in the HDFs in the region due to the positive anomaly that induces in the precipitation, but also to the positive anomaly that it causes in the temperature.

More studies are being conducted to improve the results and quantitatively understand how the magnitude and frequency of HDFs in the sector will change by the end of the century in the face of projected climate change.

Acknowledgements This research was supported by the Natural Hazards of the Central Andes project from the National University of Cuyo. We are grateful to Simon Higginson for reviewing the English and to Mauricio Vergara for the critical review.

References

Aguilar G, Riquelme R, Martinod J, Darrozes J (2013) Rol del clima y la tectónica en la evolución geomorfológica de los Andes Semiáridos chilenos entre los 27–32 °S. Andean Geology 40(1):79–101. https://doi.org/10.5027/andgeoV40n1-a04

Alaska Satellite Facility (2017) ALOS PALSAR radiometrically terrain-corrected. https://vertex.daac.asf.alaska.edu/. Accessed 17 July 2017

Antinao JL, Mcdonald E, Negrini R, Tiner R, Bobbit M, Kuehn S (2015) Nuevos antecedentes geocronológicos y estratigráficos en el Cuaternario del valle del Elqui, Chile. Dissertation, Chilean Geological Congress

Arlot S, Celisse A (2010) A survey of cross-validation procedures for model selection. Statist Surv 4:40–79. https://doi.org/10.1214/09-SS054

Berti M, Martina MLV, Franceschini S, Pignone S, Simoni A, Pizziolo M (2012) Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. J Geophys Res Earth Surf 117(4):1–20. https://doi.org/10.1029/2012JF002367

Brunetti MT, Peruccacci S, Rossi M, Luciani S, Valigi D, Guzzetti F (2010) Rainfall thresholds for the possible occurrence of landslides in Italy. Nat Hazards Earth Syst Sci 10(3):447–458. https://doi.org/10.5194/nhess-10-447-2010

Cannon SH, Gartner JE, Wilson RC, Bowers JC, Laber JL (2008) Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California. Geomorphology 96(3–4):250–269. https://doi.org/10.1016/j.geomorph.2007.03.019

Carrara A, Crosta G, Frattini P (2008) Comparing models of debris-flow susceptibility in the alpine environment. Geomorphology 94(3–4):353–378. https://doi.org/10.1016/j.geomorph.2006.10.033

Carrasco JF, Osorio R, Casassa G (2008) Secular trend of the equilibrium-line altitude on the western side of the southern Andes, derived from radiosonde and surface observations. J Glaciol 54(186):538–550. https://doi.org/10.3189/002214308785837002

CEAZA (2017) Red de estaciones meteorológicas del CEAZA. http://www.ceazamet.cl/. Accessed 17 July 2017

Coe JA, Kinner DA, Godt JW (2008) Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado. Geomorphology 96(3–4):270–297. https://doi.org/10.1016/j.geomorph.2007.03.017

Conte Nadeau A (1986) Vulnerabilidad a los eventos naturales catastróficos de los Valles Elqui, Limari y Choapa. Revista Geográfica de Chile Terra Australis 29:103–130

Crosta GB, Frattini P (2003) Distributed modelling of shallow landslides triggered by intense rainfall. Nat Hazard Earth Syst Sci 3(1–2):81–93

Food and Agriculture Organization of the United Nation (2006) Crop evapotranspiration. http://www.fao.org/docrep/X0490E/X0490E00.htm. Accessed 17 July 2017

Fuenzalida H (1982) A country of extreme climate. In: García H (ed) Chile: essence and evolution. Instituto de Estudios Regionales de la Universidad de Chile, Santiago de Chile, pp 27–35
Garreaud R (1992) Estimación de la altura de la línea de nieve en cuencas de Chile central. Revista Chilena de Ingeniería Hidráulica 7:21–32
Garreaud R (2013) Warm winter storms in Central Chile. J Hydrometeor 14:1515–1534. https://doi.org/10.1175/JHM-D-12-0135.1
Geophysics Department of the University of Chile (2006) Estudio de la variabilidad climática en Chile para el siglo XXI. Organización de estados Iberamericanos. http://www.oei.es/. Accessed 17 July 2017
Giannecchini R, Galanti Y, Amato GD, Barsanti M (2016) Probabilistic rainfall thresholds for triggering debris flows in a human-modified landscape. Geomorphology 257:94–107. https://doi.org/10.1016/j.geomorph.2015.12.012
Graña Pezoa F (2007) Crónicas y recuerdos de una inundación en el valle de Elqui. Comuna de Vicuña, Vicuña
Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity-duration control of shallow landslides and debris flows: an update. Landslides 5(1):3–17. https://doi.org/10.1007/s10346-007-0112-1
Hauser A (1985) Flujos de barro en la zona preandina de la Región Metropolitana: características, causas, efectos, riesgos y medidas preventivas. Revista Geológica de Chie 24:75–92
Hosmer D, Lemeshow S (2000) Applied logistic regression. John Wiley & Sons Inc., New York
Hotchkiss RH, McCallum BE (1995) Peak discharge for small agricultural watersheds. J Hydraul Eng 121(1):36–48. https://doi.org/10.1061/(ASCE)0733-9429(1995)121:1(36)
Hungr H (2005) Classification and terminology. In: Jakon M, Hungr H (eds) Defris-flow hazards and related phenomena. Springer, Chichester, pp 159–196
Instituto Nacional de estadísticas (2017) Demográficas y vitales. http://www.ine.cl/estadisticas/demograficas-y-vitales. Accessed 17 July 2017
Joyce RJ, Janowiak JE, Arkin PA, Xie P (2004) CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J Hydromet 5(3):487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2
Kean JW, McCoy SW, Tucker GE, Staley DM, Coe JA (2013) Runoff-generated debris flows: observations and modeling of surge initiation, magnitude, and frequency. J Geophys Res Earth Surf 118(4):2190–2207. https://doi.org/10.1002/jgrf.20148
Kleinbaum DG, Kupper LL, Muller KE (1988) Applied regression analysis and other multivariables methods. PWS-KENT Publishing Company, Duxbury
Lauro C, Moreiras SM, Junquera S, Vergara I, Toural R, Wolf J, Tutzer R (2017) Summer rainstorm associated with a debris flow in the Amarilla gully affecting the international Agua Negra Pass (30°20′S), Argentina. Environ Earth Sci. https://doi.org/10.1007/s12665-017-6530-z
Leroy SAG (2006) From natural hazard to environmental catastrophe: past and present. Quatern Int 158(1):4–12. https://doi.org/10.1016/j.quaint.2006.05.012
Maldonado A, De Porras ME (2015) Palinología de la zona andina del Norte Chico de Chile durante el Holoceno. Dissertation, Chilean Geological Congress
Marra F, Nikolopoulos EI, Creutin JD, Borga M (2014) Radar rainfall estimation for the identification of debris-flow occurrence thresholds. J Hydrol 519:1607–1619. https://doi.org/10.1016/j.jhydrol.2014.09.039
Moreiras SM (2005) Climatic effect of ENSO associated with landslide occurrence in the Central Andes, Mendoza Province, ArgentinaArgentinaArgentinaArgentina. Landslides 2(1):53–59. https://doi.org/10.1007/s10346-005-0046-4
Ortlieb L, Vargas G (2015) Hacia una historia de eventos lluviosos extremos en el sur del Desierto de Atacama, Norte Chico, a partir de fuentes documentales. Dissertation, Chilean Geological Congress
Pavlova I, Jomelli V, Brunstein D, Grancher D, Martin E, Déqué M (2014) Debris flow activity related to recent climate conditions in the French Alps: a regional investigation. Geomorphology 219:248–259
Pérez C (2005) Cambio climático: vulnerabilidad, adaptación y rol institucional. Universidad de La Serena, La Serena, Estudios de caso en el valle de Elqui
Pérez C, Cepeda J, Fiebig M, Pizarro J (2008) Desastres naturales y plagas en el valle del Río Elqui. In: Cepeda PJ (ed) Los sistemas naturales de la cuenca del río Elqui (Región de Coquimbo, Chile): Vulnerabilidad y cambio del clima. Ediciones Universidad de La Serena, La Serena, pp 295–333
Petley D (2012) Global patterns of loss of life from landslides. Geology 40(10):927–930. https://doi.org/10.1130/G33217.1
Pierson TC (2005) Hiperconcentrated flow—transitional process between water flow and debris flow. In: Jakob M, Hungr H (eds) Defris-flow hazards and related phenomena. Springer, Chichester, pp 159–196

Powers JG, Ahmadov R, Peckhamand SE et al (2017) The Weather Research and Forecasting Model: overview, system efforts, and future directions. Bull Am Meteor Soc 98(8):1717–1737. https://doi.org/10.1175/BAMS-D-15-00308.1

Prokhorov AV (2001) Partial correlation coefficient. In: Hazewinkel M (ed) Encyclopedia of Mathematics. Springer, New York

Rauld Plott R, et al (2012) Zonificación Áreas de Riesgo Plan Regulador Intercomunal Provincia del Elqui, Región de Coquimbo. In: Ministerio de Vivienda y Urbanismo (ed) Plan regulador Intercomunal Provincia del Elqui. Gobierno de Elqui, La Serena, pp 1–139

Romero H, Rovira A, Véliz G (1988) Geografía IV Región de Coquimbo. Instituto Geográfico Militar, Santiago de Chile

Rosenbluth B, Fuenzalida H, Aceituno P (1997) Recent temperature variations in Southern. Int J Climatol 17:67–85. https://doi.org/10.1002/(SICI)1097-0088(199701)17:1<67:AID-JOC120>3.0.CO;2-G

Rutllant J, Fuenzalida H (1991) Synoptic aspects of the central Chile rainfall variability associated with the Southern Oscillation. Int J Climatol 11:63–76

Santos JR, Norte F, Moreiras S, Araneo D, Simonelli S (2015) Predicción de episodios de precipitación que ocasionan aludes en el área montañosa del noroeste de la provincia de Mendoza. Argentina 40(1):65–75

Sarricolea P, Herrera-Ossandon M, Meseguer-Ruiz Ó (2016) Climatic regionalisation of continental Chile. J Maps 5647:1–8. https://doi.org/10.1080/17445647.2016.1259592

Seluchi M, Chou C (2001) Evaluation of two Eta Model versions for weather forecast over South America. Geofísica Internacional 40(3):219–237

Sepúlveda SA, Padilla C (2008) Rain-induced debris and mudflow triggering factors assessment in the Santiago cordilleran foothills, Central Chile. Nat Hazards 47(2):201–215. https://doi.org/10.1007/s11069-007-9210-6

Sepúlveda SA, Rebolledo S, Vargas G (2006) Recent catastrophic debris flows in Chile: geological hazard, climatic relationships and human response. Quat Int 158(1):83–95. https://doi.org/10.1016/j.quaint.2006.05.031

Staley D, Negri J, Kean J, Laber J, Tillery A, Youberg A (2017) Prediction of spatially explicit rainfall intensity-duration thresholds for post-fire debris-flow generation in the western United States. Geomorphology 278:149–162. https://doi.org/10.1016/j.geomorph.2016.10.019

Vera C, Silvestri G, Liebmann B, González P (2006) Climate change scenarios for seasonal precipitation in South America from IPCC-AR4 models. Geophys Res Lett 33(13):L13707. https://doi.org/10.1029/2006GL025759

Vuille M, Milana JP (2007) High-latitude forcing of regional aridification along the subtropical west coast of South America. Geophys Res Lett 34(23):1–6. https://doi.org/10.1029/2007GL031899

White AB, Gottas DJ, Henkel AF, Neiman PJ, Ralph FM, Gutman SI (2010) Developing a performance measure for snow-level forecasts. J Hydrometeor 11:739–753

Wilson RC, Wieczorek GF (1995) Rainfall thresholds for the initiation of debris flow at La Honda, California. Environ Eng Geosci 1(1):11–27

Zavala H, Trigos H (2008) Hidrología de la cuenca del Valle de Elqui. In: Cepeda PJ (ed) Los sistemas naturales de la cuenca del río Elqui (Región de Coquimbo, Chile): Vulnerabilidad y cambio del clima. Ediciones Universidad de La Serena, La Serena, pp 66–164