Assessment the risk of snow avalanche in the North of Italy

Nicolis, Orietta  
*University of Bergamo, Dept. of Information Technology and Mathematical Methods*  
Viale Marconi, 5  
Dalmine 24044, Italy  
E-mail: orietta.nicolis@unibg.it

Assunção, Renato  
*Universidade Federal de Minas Gerais, Departmento de Estatistica*  
Ave. Antônio Carlos, 6627  
Belo Horizonte, MG, 31270-901, Brazil  
E-mail: assuncao@est.ufmg.br

The study of avalanche events is particularly important to assess and predict the degree of risk involved in a given area and time. In this work we propose an alternative methodology based on a space-time point process where the intensity function indicates the limiting expected rate of occurrence of snow avalanches occurring on day t at location (x, y), conditioned on the historical information available prior to time t. Also, we use a self-exciting model to deal with unobserved random space-time effects. The model depends also on some environmental variables (degree of slope, exposure, altitude, etc.) which may be considered as covariates. To show the ability of the model in estimating and forecasting the avalanche hazard we consider the application to the digitalized Avalanche Database of the Trentino region, Italy.

1 Introduction

Avalanches are natural phenomena that in a mostly mountainous territory may significantly affect land use. In recent years the study of avalanche phenomena has attracted growing interest especially for the increase of accidents and deaths, now comparable with those related to natural disasters. This is mainly due to a wide anthropization of mountain areas which has often brought a rapid growth of recreational activities, transportation, and constructions in high-altitude areas without an adequate assessment of avalanche hazard. Hence, the analysis of avalanche activity is extremely important to prevent damage and for activities aimed at land use planning in mountain areas.

Many scientists have been studying avalanches to try to map the risk and improve predictions. To that end several statistical methods have been proposed based on different approaches. While some papers are aimed to predict the long-return period avalanche for a given avalanche path (Meunier, and Ancey, 2004; Eckert, *et al.* 2008; Eckert, *et al.* 2010), others try to find variables that are correlated with avalanche events that can be used as predictors in a statistical model (Baggi and Schweizer, 2009; Ancey, 2001). For example, Baggi and Schweizer (2009) studied the characteristics of wet-snow avalanche activity for 20 years of observations from a high alpine valley in Switzerland. From the analysis of the occurrence data in combination with meteorological and snowpack data, they found that snow depth, precipitation and air temperature have the highest correlation with avalanche activity. Ancey (2001) distinguishes tree fixed parameters related to the avalanche path given by the mean slope, the new slope and the wind. In particular, his findings can be summarized as follows: (i) the average inclination of starting zones ranges from 27 to 50 degrees; (ii) most of time, snowfall is the cause of avalanches, hence the hazard increases significantly with the increase in the depth of new snow; (iii) the wind is an additional factor which significantly influences the stability of a snowpack since it causes uneven snow redistribution (accumulation on lee slopes), which accelerates
snow metamorphism, forms cornices, whose collapses may trigger avalanches. The latter approaches are aimed to study avalanche activity on a small spatio-temporal scale. Climate change has been considered by some authors for characterizing the avalanche activity at a larger scale (Eckert et al., 2010, Jomelli and Pech, 2004). For example, while Jomelli and Pech (2004) suggest that at low altitudes, avalanche magnitude has declined since 1650 in the Massif des Ecrins, in the French Alps, Jomelli et al. (2007) found no significative correlation between the fluctuations in avalanche activity between 1978 and 2003 and large-scale atmospheric patterns, in the Maurienne Valley in France. Few models have been proposed for forecasting the risk of avalanche in a spatio-temporal framework. Straub and Grêt-Regamey (2006) proposed a Bayesian probabilistic model for spatial mapping and hazard risk assessment, based on a deterministic dynamic model combined with an explicit representation of different parameter uncertainties and Eckert et al. (2010) introduced a spatio-temporal hierarchical model inspired from spatial epidemiology to study the fluctuations of avalanche occurrence possibly resulting from climate change.

In this work we propose an approach based on space-time point processes (see Daley, and Vere-Jones, 1998 for an introduction) for modeling the avalanche risk. In particular, the intensity function of the process indicates the limiting expected rate of occurrence of snow avalanches occurring on day $t$ at location $(x, y)$, conditioned on the historical information available prior to time $t$. Also, we use a self-exciting model to deal with unobserved random space-time effects. The location $(x, y)$ represents the baricenter of the polygon which draws the shape of avalanche. For showing the effect of some covariates (such as elevation, slope, temperature, etc.) different models are proposed. Application to the digitalized Avalanche Dataset of Trentino region (Italy) illustrates the ability of the models to forecast the risk avalanche. Although this approach has not been previously applied to avalanche events, it has been used for analysis spatio-temporal analysis of earthquakes occurrences (Ogata, 1998; Brix and Diggle, 2001) and wildfire risk (Peng et al. 2005; Schoenberg et al. 2007; Brillinger et al. 2006)

The paper is organized as follows. Section 2 and Section 3 provide a description of the data set and a preliminary analysis of data, respectively. The analysis of space-time patterns is given in Section 3. In Section 4, some forecasting models for avalanches are proposed. Results and conclusions are in Section 5.

2 The data set

The data used in this work have been provided by the province of Trento through the availability of digitalized Avalanche Database (based on a permanent survey on avalanches). The database collects and documents the avalanche events since the seventies to today, including maps with the location of various sites of avalanches and documents with the description and analysis of various phenomena (dates, causes, altitude of posting, exposition, damages to people or things, etc.). For a large zone of the Province of Trento (around 40%) the database include further detailed information of avalanche events which were collected for drawing the CLPV (Map of Probable Localization of Avalanche) which shows the dangerous areas and the avalanche events which happened before a given period of time (see Fig. 1). An online version of the data set is available on the web site http://www.territorio.provincia.tn.it/. The Avalanche Database include 4693 well documented avalanche events over a period from January 1970 to January 2008 for 1108 sites distributed on the Trentino region as shown in Fig. 1 (left). However, some avalanche counts are missing or were not perfectly surveyed during the entire time period, making certain data to be nonhomogeneous in space and time. In this study we will concentrated our attention to some time periods where the permanent survey has been conducted as fully as possible.

Others data, considered in this work, which are supposed to be in a certain way connected with avalanche activities include elevations, snowpack data and meteorological variables. Elevation data
Figure 1: Example of avalanche from a CLPV map (left); avalanche sites (center); permanent areas where snow-related data are collected (right).

(In Fig. 3, left) are given by the web site http://eros.usgs.gov which provide the GTOPO30 data set with 1km resolution. GTOPO30 is a global digital elevation model (DEM) consisting of a raster grid of regularly spaced elevation values that have been primarily derived from the U.S. Geological Survey (USGS) topographic map series. Snow-pack and meteorological data for Trentino province are collected from the public meteorological centre “Meteotrentino” (www.meteotrentino.it). The database include data from 36 permanent areas where daily are registered hand made observations of temperatures, snowpack parameters and avalanche activity (Fig. 1, right).

3 Preliminary analysis of data

Figure 2: Yearly avalanche events from 1979 to 2008 (upper left); Monthly avalanche events from 1983 to 1988 (upper right); Monthly average snow depth (upper left); Monthly average temperature (bottom right).

The number of avalanche events in Trentino change significantly in time from January, 1970
to January, 2008 (Fig. 2). In particular, a large number of avalanches occurred in the period from 1979 to 1988 while it decreased significantly in the last 20 years. There is not a clear reason for this decreasing temporal behavior but it is known that this period is characterized by different meteorological conditions. While the winter 1985-1986 is remembered for its frequent snowfalls, the winter 2006-2007 is known for being the warmest in the last 60 years. For showing the relationship between avalanche events and meteorological variables we have chosen the period 1983-1988 from November to April which is characterized by a large number of avalanche events. Figs. 2 a), 2 b), and 2 c) show the monthly number of avalanches, the monthly average level of snow, and the monthly average temperature, respectively: the highest numbers of monthly avalanches are often associated to high levels of snow and cold winters. The data on the average level of snow and the monthly average temperature are missing for the month of November 1984 and 1985. Although the amount of new snow and temperature are important factors for assessing the risk of avalanches, other factors can be responsible of avalanches (such as elevation, slope, snowpack, presence of skiers, etc.). The winter season starts September 1st of a given year and ends June 30th of the following year, but the major avalanches in Trentino region occur between November and April with peaks during the months of February and April. In fact, the latter months are often characterized by snowstorms (February) and thaw (April) which contribute to increase the level of danger.

In order to better understand the dynamics of the avalanche activity, let’s analyze some of the more intense avalanche days. Figs. 3 (middle) and 3 (right) represent the space-time distribution of avalanche events in two different periods: from January 30 to February 2, 1986 and from December 9 to December 12, 1990. While in the first period the avalanche events cover all the Trentino region in the second period the number of avalanche events are concentrated in the North East part. In both periods, the avalanche activity lasted 4 days with a spatial clustered distribution for each day. This means that if an avalanche occurs in a particular site, it is likely that other avalanches occur in the neighborhood in the same day. As expected, no avalanche occurs at the same site on consecutive days. In both examples the altitude is a very important variable for the spatial distribution of avalanche events.

### 3.1 Space-time patterns

In a space-time analysis, the first task is to check how much the spatial pattern changes over time or, equivalently, how much the temporal evolution changes as we move in space. If the spatial pattern does not change in time, we can carry out a simple analysis by analyzing separately the marginal spatial and temporal patterns rather than considering them jointly. This is so because the under
temporal invariance of the spatial pattern implies that the joint pattern is simply the product of a marginal spatial pattern and a marginal temporal pattern.

Unfortunately, this is not the case in the point process representing the avalanche. To show this lack of invariance and, at the same time, to motivate a more complex model, we will use a simple kernel smoothing estimate of the point pattern intensity in three different time periods. For a purely spatial point process, the intensity function at locations $x$ is given by

$$\lambda(x) = \lim_{\|\Delta x\| \to 0} \frac{E(N(\Delta x))}{\|\Delta x\|}$$

where $|\Delta x|$ is the area of a small region $\Delta x$ centered at $x$. Our kernel estimates use a correction for edge effects and it is based on the quartic kernel function.

Figure 4 shows the point pattern of the avalanches on top of the kernel intensity estimate for data broken up into three periods: from 1981 to 1984, from 1985 to 1988, and from 1989 to 1992. It is visually distinct the changing spatial pattern. There is an increase in the NW corner and in the E and NE corners of the map.

Figure 4: Pattern 1 (left): from 1981 to 1984; Pattern 2 (middle): from 1985 to 1988; (c) Pattern 3 (right): from 1989 to 1992. On the background, we draw the kernel estimates of the intensity function using a quartic kernel function with a border effect correction.

We carried out formal statistical tests to verify this visual impression of a changing spatial pattern. Figure 5 shows the plot of the difference $D(r) = K_1(r) - K_2(r)$ between a pair of estimated Ripley’s K-functions ($K_1(r)$ and $K_2(r)$), shown as solid lines. The dashed lines show the 95% confidence bands under the hypothesis of equal underlying K-functions. The leftmost plot shows the difference between the Ripley’s $K$-functions of patterns from time periods 1 and 2. The second plot corresponds to the difference between the $K$-functions from the first and third periods. Finally, the third plot is that connected with the difference between the $K$-functions from periods 1 and 3. The first two $D(r) = K_1(r) - K_2(r)$ curves lie completely outside the envelopes giving strong statistical evidence of true difference between the underlying K-function of the initial period with the other two time periods. The third plot shows that there is not significant difference between the second and third periods. Therefore, the visual suggestion of Figure 4 that the avalanches intensity increased in some regions is justified by the data evidence.

This brings a difficulty in terms of modeling and in terms of understanding of the phenomenon. A good model should allow for a differential change in the spatial pattern as times evolves. It is not clear which covariate could drive this change other than the longitude. We wonder if the increase of human presence (traffic, occupation by households or leisure activities) increased in the period in the regions where the intensity increased. This would indicate that the intensity is likely to have no trend in time but that the recording of the avalanches increased in those regions. There is some anecdotal
Figure 5: Test for the difference between two K-functions: periods 1 versus 2 (left), periods 1 and 3 (middle), and periods 2 and 3 (right). Bands of 95% confidence are shown as dashed lines.

Figure 6: Kernel density of avalanches in time

evidence that leisure business has been initiated more intensively in this region in the 90’s and this could imply a more intense surveillance of the region.

This difficulty induced an analysis strategy. If we analyze the whole period as a single dataset we will need to come up with a model for the changing patterns as well as a model for the prediction of avalanches in steady state situations. We think that this mixing is not healthy and can lead to models good for one purpose but not for the other. Since our main objective is to propose a predictive method for avalanches under regular conditions, we decided to break the data into two periods and to analyze only the the first one, from January 01, 1980 to December 31, 1989, covering 10 years of data. This is the period that has a reasonably stable spatial pattern. It allows us to study the predictive power of our methods in a situation that is not changing due to human intervention. Later, in a second analysis to be pursued elsewhere, we will consider the time evolution of the spatial pattern, specially its increase in the NE region.

3.2 The temporal pattern

The purely temporal trend can be seen in Figure 6. The deep and seasonal valleys in this intensity function are due to the spring and summer periods in each year when there is no avalanches. It is obvious that the mid years had a higher intensity compared to the earlier and the later years. Because of the near zero intensity in the mid year months, we fitted our models using only the data from November to April in each year. Therefore, each year is composed by five months only.

3.3 Effect of elevation and slope

Some covariates associated with the avalanches risks are exogenous, in the sense that its occurrence is not causally affected by the avalanches point process. These covariates are elevation, temperature, and the slope of the terrain. We are still collecting the data for temperature so we present here a preliminary analysis.
To explore the possible influence of altitude on the avalanches intensity, we carried out a simple spatial correlation analysis. First, it should be noted that, for all practical aspects, altitude is constant during the analysis time frame. Hence, it can not be associated with any temporal changing pattern, but only with the marginal spatial pattern. For this reason, we considered all point events during the 80’s, irrespective of their time of occurrence. The left hand side in Figure 7 shows the elevation map with the events pattern superimposed. There is a clear association between the two with more events clustered on regions of high altitude.

![Figure 7: Maps of elevation (left), slope (center) and maximum nearby slope (right) with the events pattern superimposed.](image)

On this same Figure, the middle map shows a map of the slope, calculated as the square length of the gradient vector associated with the tangent plane on the elevation surface at each pixel. The point pattern is also superimposed in this map. In this case, there is less association between the events and the slope map than between the events’ intensity and elevation. The explanation is that slope per se is not the correct variable as the avalanche risk is not linearly related with the terrain inclination. There is a non-linear association as angles too close to zero or too steep present no risk of avalanches. As pointed out in section 1, the angle with the horizontal plane must be between 25 and 50 degrees in order to present some avalanches risk. Therefore, we created a binary map with areas with slopes’ angles within this (25, 50) range marked as white while pixels where the angles are outside this range are colored in green. We see more matching between the events’ pattern and this binary map than the slope per se.

3.4 Effect of time varying and endogenous covariates

Maybe the most important predictor for avalanches in the near future at a given location \( x \) is the occurrence of other avalanches in the recent past days in the immediate neighborhood of \( x \). This variable can be easily collected and be used to forecast future avalanches. Therefore, at each pixel \( x \) and time \( t \), we considered the number of other avalanches in the previous two days that occurred in a radius of 3.0 kilometers away from \( x \).

4 A forecasting model for avalanches

4.1 Conditional intensity in point processes

Any spatial-temporal point process is uniquely characterized by its conditional intensity function \( \lambda(t, x, y | \mathcal{H}_t) \) given by the limiting conditional expectation

\[
\lambda(x, y, t | \mathcal{H}_t) \lim_{\Delta t, \Delta x, \Delta y \to 0} \frac{E[N \{(t, t + \Delta t) \times (x, x + \Delta x) \times (y, y + \Delta y)\} | \mathcal{H}_t]}{\Delta t, \Delta x, \Delta y}
\]
provided the limit exists. This is a random function that depends on the prior history, \( H_t \), of the point process up to time \( t \). Technically, this history is defined as the filtration \( \{ H_t : t \geq 0 \} \), the increasing and right-continuous family of sigma-algebras determined by all events occurring up to time \( t \) plus all initial conditions. The space-time process is \( H_t \)-measurable for every \( t \geq 0 \) and it is said to be adapted to this filtration. In practice, we can simply assume that \( H_t \) represents the set \( \{ (t_i, x_i, y_i), \forall i : t_i < t \} \) of all events that occurred previous to \( t \) where \( t_i \) is the time of \( i \)-th event and \( (x_i, y_i) \) is its spatial location.

The importance of the conditional intensity function is that, if it depends on a vector-valued parameter \( \beta = (\beta_0, \beta_1, \ldots, \beta_k) \in \mathbb{R}^{k+1} \), then the likelihood based on the observed events \( (t_i, x_i, y_i) \) for \( i = 1, \ldots, n \) is given by

\[
L(\beta) = \prod_{i=1}^{n} \lambda_\beta(x_i, y_i, t_i|H_{t_i}) \exp \left( - \int \int \int \lambda_\beta(x, y, t|H_t) \, dx \, dy \, dt \right) .
\]

The second derivative matrix with respect to \( \theta \) of the log-likelihood function \( l(\beta) \) evaluated at the maximum likelihood estimator can be used in a usual way to derive confidence intervals and hypothesis tests concerning the parameter values. Therefore, statistical inference is straightforward with the main difficulties concentrated on the numerical aspects of evaluating the likelihood and its derivatives.

4.2 Conditional intensity models for avalanche risk

In this preliminary analysis, we considered a small number of models that should capture the main aspects of the avalanche dataset. One first class of models is nonparametric and has separable spatial and temporal effects. This is given by

\[
\lambda_{1a}(x, y, t|H_t) = \lambda(x, y, t) = \beta_0 + \beta_1 S(x, y) + \beta_2 T(t)
\]

or by

\[
\lambda_{1m}(x, y, t|H_t) = \exp(\beta_0 + \beta_1 S(x, y) + \beta_2 T(t))
\]

where \( \beta \) is the parameter vector to be estimated. So, one is an additive model while the other is a multiplicative model. In these models, \( S(x, y) \) is a deterministic function of the location \( (x, y) \) and it is estimated by a two-dimensional kernel smoother

\[
S(x, y) = \frac{1}{n_0} \sum_{j=1}^{n_0} K\left(\frac{x - x_{0j}}{\phi_x}\right)K\left(\frac{y - y_{0j}}{\phi_y}\right)
\]

where \( K \) is a suitable kernel function, taken as the quartic kernel in this paper. The function \( T(t) \) is a periodic with trend deterministic function, also estimated by kernel methods using the events’ times. This function is shown in Figure 6. The determinist aspect of these functions make the conditional intensity independent of the past, justifying the first equality in (3). To have an identifiable model and to avoid numerical instabilities, we centered all covariates at zero.

It is likely that this model has less predictive power than other models as it does not incorporate important additional information. For this reason, we want to use the covariates to improve this model.
At this moment, we have the elevation \( E(x) \) and the binary slope \( S(x) \), that do not vary in time and is available for all points \( x \) in the map. We need to add a factor to control for the temporal changes. This has no fixed parametric shape and we simply allow one arbitrary value for each year. As in the previous model, we centered the covariates at zero.

Hence, another class of models has an intensity varying only with the exogenous covariates and the temporal components. We again have \( \lambda(x, y, t|H_t) = \lambda(x, y, t) \) for these models, a deterministic intensity function. It is given by

\[
\lambda_2(x, y, t) = \lambda_{1a}(x, y, t) + \beta_3 E(x) + \beta_4 S(x) + \beta_5 (\text{year} - c)
\]

where \( c \) is the mid period date so the linear trend variable is centered at zero. Another version of this model is the multiplicative form where

\[
\lambda_2(x, y, t) = \lambda_{1m}(x, y, t) \exp (\beta_3 E(x) + \beta_4 S(x) + \beta_5 (\text{year} - c))
\]

We can test for the additional improvement of this model with respect to the first class of models by means of the difference between the log-likelihood maximum values of each model.

We have one additional covariate, the number of avalanche events nearby each pixel in the previous days. We made an arbitrary choice concerning the proximity thresholds using 4 days for the previous history and a radius of 5 km for the distance around each pixel. Of course, the results can be impacted by different choices and a sensitivity analysis is needed to confirm the findings presented here. The final class of models we are going to consider are those that include this history of previous avalanches events in the area near each point. The conditional intensity is a truly random function that depends on the previous occurrences. Let

\[
H(x, t) = \int \int I_{B_x(r) \times [t-\epsilon, t]}(x, y, t)N(dx, dy, dt)
\]

where \( I_A(\cdot) \) is the indicator function of the set \( A \) and \( B_x(r) \) is a small disc centered at \( x \) and with radius \( r \). That is, \( H(x, t) \) is the number of events from the point process \( N \) that are inside the three-dimensional cylinder \( B_x(r) \times [t-\epsilon, t] \). Clearly, \( H(x, t) \) is \( \mathcal{H}_t \)-measurable. Then, the models incorporating this previous history are of two types, an additive model,

\[
\lambda_3(x, y, t|H_t) = \lambda_2(x, y, t) + \beta_6 H(x, t)
\]

and its multiplicative version,

\[
\lambda_3(x, y, t|H_t) = \lambda_2(x, y, t) \exp (\beta_6 H(x, t))
\]

5 Results and conclusions

In this preliminary report, we did not fit the models (7) and (8). They require a much heavier numerical work as each time unit (day, in our case) has an associated map with the covariate \( H(x, t) \) that enters the likelihood maximization in each iterative step. We are working on this model and should have final results soon.

The results for the models 4 and 6 are in Table 5.
The difference between the log-likelihood of these models should be tested against a chi-square distribution with 2 degrees of freedom.

We expect other covariates to matter, such as temperature and the amount of snow accumulated in the soil. Both are time varying and should be useful in terms of prediction of avalanche events. We are in the process of collecting these covariates and we expect to have an extended version of this paper incorporating these additional information in the near future.

REFERENCES


