Nat. Hazards Earth Syst. Sci. Discuss., 3, C934–C960, 2015
www.nat-hazards-earth-syst-sci-discuss.net/3/C934/2015/

Natural Hazards and Earth System Sciences
Discussions
Open Access

Interactive comment on "Predicting storm triggered debris flow events: application to the 2009 Ionian-Peloritan disaster (Sicily, Italy)" *by* M. Cama et al.

M. Cama et al.

mariaelena.cama@unipa.it

Received and published: 10 June 2015

Please find below our responses to the two anonymous referees. As it is evident, a lot of modifications have been done and a fifth co-author involved. We attach the full file with track changes. We look forward for any further comments.

REVIEWER 1 The contribution "Predicting storm triggered debris flow events: application to the 2009 Ionian-Peloritan disaster (Sicily, Italy)" by M. Cama and co-Authors is good, clear, well-written and potentially publishable. The Authors present a susceptibility model, based on step-wise binary logistic regression, for the pre-

diction of rainfall-induced debris flows. The method is applied on a well-known catastrophic case study occurred in Sicily (southern Italy) in 2009. The topic address scientific questions within the scope of NHESS. The discussion about the non-linearity of the process is widely shareable. The theoretical background is well-argued. The readability of the whole paper is notable, with a fluent English. Study area, materials and methods are well described. Results, discussion and conclusions are well organized. Maybe, the section named "Discussion and conclusions" could be split into two different sections. RESPONSE: Ok, done.

The same for the "Introduction" section, which could be split into a proper Introduction and a review of the huge literature about susceptibility analysis (some other references should be added). RESPONSE: Ok, done.

GENERAL COMMENTS 1)The only downside of the work concerns the definitions of the indexes obtained from the contingency table, and the related ROC analysis (Page 1748, Lines 8-19, and Figure 9). This aspect is often subject to misunderstandings. Several Authors have addressed the problem in different fields, e.g. Wilks (1995), Fawcett (2006), Rossi et al. (2010); Staley et al. (2012), Gariano et al. (2015). In particular, the corrigendum paper written by Barnes et al. (2009) seems clarifying the topic and limiting confusion and misinterpretations. I have a list of comments and suggestions about this topic:

Authors should write the Contingency Table in a better way, as proposed by Barnes et al. (2009). In the way they have written it, it's difficult to understand which are the false negatives and the false positives. Authors should note that the correct form the write the "1-specificity" index (also note as False Positive Rate or Probability Of False Detection) is (FP / (FP + TN)): cf. e.g., Barnes et al. (2009), Rossi et al. (2010), Gariano et al. (2015). I made some calculations from the numbers reported in Figure 9 (which I suggest to change into one or more tables), thus I'm not so sure that the skill scores were calculated correctly by the Authors. But this can be a consequence of the way in which they have written the contingency table. The position of False Positives and False Negatives is difficult to understand. Thus, I would like to ask the authors to rewrite better the table and to check calculations. I didn't understand why the numbers reported in the contingency table are with decimal places and are not integers. There should be not reported the numbers of cells? Please explain.

RESPONSE: Ok, we have modified the label in the roc plot and the text according to your request. As regards the contingency and decimal number of fig. 9, this was due to the TANAGRA output table layout which puts the hypothesized classes in columns and the true classes in rows (a transposed style respect to the one proposed by Barnes et al. (2009). We have now split fig. 9 into three tables and rearranged the format of the contingency table according to your advice. Besides, in order to avoid decimal number we directly reported the total (not the mean on the 100 cross validation replicates) TP, FP, TN and FN cases. The decimal format for the number of cases was due to the fact that instead to report the whole number in 100 replicates, we computed the average value of TP, TN, FP and FN. We agree with you it was quite confusing and we changed the option.

I suggest include in the analysis also the Hanssen-Kuipers discriminant (cf. Peres and Cancelliere, 2014; Gariano et al., 2015). This index gives a measure of the accuracy both for events and non-events.

RESPONSE: We have computed the HK index for 2007 and 2009 models chronovalidation. We obtained below 0.5 maximum HKs and we discuss the point in the discussion section. We add now a table showing the results.

---

2)Since the susceptibility analysis concerns rainfall-induced phenomena, I suggest to include in the analysis also another predictor (another variable) referred to the rainfall.

As an example, an index of the measure of the spatial and temporal variability of rainfall maxima (daily, hourly). Some works (e.g., Minder et al., 2009; Iovine et al., 2014) have addressed the problem including such an index in the analysis. Alternatively, Authors could made a more detailed analysis of antecedent rainfall conditions, e.g. comparing the conditions preceding the failures with the same time intervals in other periods without events.

RESPONSE: We actually do not consider the rainfall data spatially significant enough for a 8m resolution model, so to be included in a spatial stochastic model. The density of rain gauges is so low that we would test the effect of the regionalization equation we select rather than the effect of tru rainfall intensity. In fact according to Minder et alii: ". . ..an important role for small-scale rainfall features in determining where slide are triggered on the storm time scale. Yet, if the rainfall from such convective cells is distributed randomly across a region from storm to storm they will have no net influence on the pattern of susceptibility over climatological time scales. For spatial variations in mountain rainfall to influence the climatological pattern of landslide susceptibility they must be both large and persistent enough. Whether this is the case on small (10 km or less) scales remains an open question." Moreover, "In landslide susceptibility studies, information on 10 km–scale spatial variability of rainfall is very seldom considered in long-term susceptibility analysis, in part because mountain rainfall patterns have not been well observed or understood on those scales. However, in recent years it has become clear that large variations in precipitation occurring on spatial scales of 10 km or less are a persistent and predictable feature of mountain climates in a variety of regions". At the same time, we have added a synthetic view of the antecedent rainfall events which stroke the area, describing also their consequences in terms of landslide activations.

---

3) For what concerns debris flows, an important issue to be known is the area in which they can develop and the maximum distance they can reach. This is an important issue for civil protection purposes. Thus, the Authors should specify that the proposed

method can investigate only the source (triggering) areas of those phenomena. Propagation channels, alluvial fans, and invaded areas could be predicted by using different causal factors/models.

RESPONSE: Ok, we clearly specify this point in the introduction, methods and recall it in the conclusions. We limited our modelling to susceptibility assessment, without any insights into velocity, volumes and runout distances.

---

4) Some of the Authors of this manuscript have already addressed the topic in several previous work. I suggest to highlight differences and improvement of this work with respect to the previous ones. RESPONSE: Ok, we now highlight the differences between this research and the other whose results we published, in the background section.

SPECIFIC COMMENTS I suggest to use all over in the text “calibration set” instead of “training set”, and “validation set” instead of “test set”.

RESPONSE: Ok, done for all cases, with the exception of those sentences which also contain the word validation, for which the substitution would have sound odd).

Page 1736, Line 19: “According to Koppen classification: : :”; I suggest to include a reference (e.g., Koppen, W., 1948). RESPONSE: Ok, done.

Page 1742, Line 10-11: please explain better the following sentence “: : : the odd ratios (OR), which is calculated by simply exponentiating the _”. This sentence seems a bit ambiguous if compared to the definition of “odds ratio” proposed by Stephenson (2000). Probably, they are two different indexes. Thus, I would ask the Authors to explain and define better the “odd ratios”. RESPONSE: Ok, the odds (p/1-p) for the logistic regression is exp(a+bx)/1+exp(a+bx). If the x variable increases by one

unit we have that the new odds is exp[a+(x+1)b]/1+exp[a+(x+1)b] = [exp a exp bx exp b]/[exp a exp bx] = exp b. It is so demonstrated (e.g., Hosmer ad Lemeshow, 2000; Menard, 2001), for example in a one-predictor model, that OR = exp b1. By exponentiating the regressed coefficient of the predictor x1, we have a measure of the increasing/decreasing (depending on the sign of “b”) of the likelihood for Y=1 due to an unitary variation of it. It’s quite long to report this and we have only added a reference to Hosmer and Lemeshow for allowing the reader to search for details.

Page 1743, Line 16-17: Authors state “The iterative calculation stops when the addition of any of the left variables does not meaningful increase the performance of the model”. Please explain better “meaningful”; is there a threshold value? If yes, please explain it. RESPONSE: Ok, it is now specified what we intended with meaningful by directly referring to the significance of chi.square test on -2LL.

Page 1746, Line 21: Please rewrite better the following sentence “Training and test landslides can be obtained by Chung and Fabbri (2003)”. RESPONSE: Ok, done.

Page 1749: Numbers reported in the text do not always meet them reported in Figure 9. Please check. RESPONSE: Ok, corrected.

Page 1767: Please change Figure 9 into one or more Tables. RESPONSE: Ok, done.
____ âĂČ

TECHNICAL CORRECTIONS

Page 1733, Line 22: I suggest to change “databanks” into “databases”. RESPONSE: Ok, corrected.

Page 1735, Line 23: "Brenning et al., 2005"; in the reference list there is only "Brenning". RESPONSE: Ok, corrected.

Page 1735, Line 24: "Guzzetti et al., 2005"; in the reference list is "Guzzetti et al.,2006". RESPONSE: Ok, corrected.

Page 1742, Line 5: I suggest to replace "_n" with "_1 ,_2 ..._n". RESPONSE: Ok, corrected.

Page 1742, Line 13: please correct "_ s". RESPONSE: Ok, corrected.

Page 1742, Line 15: please correct "_ n". RESPONSE: Ok, corrected.

Page 1744, Line 3: change "the use a regular grid" into "the use of a regular grid". RESPONSE: Ok, corrected.

Page 1747, Line 10: "Bai et al., 2010"; in the reference list is "2009". RESPONSE: Ok, corrected.

Page 1753, Line 6: Please correct "inside a 10 km2 are there are". RESPONSE: Ok, corrected.

Page 1753, Line 12: "risk prospective" or "risk perspective"? RESPONSE: Ok, corrected.

Page 1759, Figure 1: Red and blue points indicating the rain gauge in Figure 1a are scarcely visible. I suggest to try to use a different background. The same for Fgure 1b. RESPONSE: Ok, corrected.

Moreover, please rewrite the caption with more details. RESPONSE: Ok, corrected.

Page 1760, Caption of Figure 2: I suppose that "16 October" should be "26 October". RESPONSE: Ok, corrected.

Page 1764, Figure 6: Please use the same format for the legends in figures a) and b). RESPONSE: Ok, corrected.

Page 1771, Figure 13: Please note that "TP-rate" and "FP-rate" have not been defined in the text. Moreover, I suggest to do not use rounded lines to draw ROC curves. RESPONSE: Ok, corrected.

Page 1774, Figure 16: Please use the same scale for horizontal and vertical axes. RESPONSE: Ok, corrected.

REVIEWER 2 The manuscript of Cama and co-authors presents a landslide susceptibility assessment of a 10 km2 area in Sicily (Italy) based on two landslide inventories observed after two heavy rain storms in 2007 and 2009. The topic of this study fits very well into the scope of NHESS and the corresponding special issue. The data analysis is (according to my limited expertise of stochastic modelling) state-of-the-art and presented in a comprehensible way. However, I have major concerns about the innovation of this manuscript. Regional landslide susceptibility assessment based on inventories of observed landslides and stochastic modelling has been done in the past 10 to 20 years worldwide. Thanks to the increased availability, quality and spatial resolution of GIS-data and thanks to new stochastic models such analysis have become more and more sophisticated in recent years. This work of Cama and co-authors is another nice example of such studies. It uses a very interesting data set – including a relatively large amount of observed hillslope debris flows – which (to my knowledge) has not yet been exploited for this purpose. The data set is particularly interesting because it includes two storm events that occurred in the same region and within only two years. I think, this is the major added value of this manuscript.

On the other hand, I must say that the presented analysis and the conclusions from this work are very similar to previous studies. For example, I notice an astounding similarity to a paper by Von Ruette et al., 2011, Geomorphology, 133, 11-22, which has the same methodology, a very similar way to validate the model, the same way to present the results – and very similar conclusions. I wonder, what do we actually learn from this new study which is different from the Von Ruette-study? Of course, this is a different catchment with different geomorphological conditions leading to a slightly different set of predictors. But I really don't see new general insights either with regard to the usefulness of the method or with regard to governing landscape predictors. If this paper shall be published in NHESS it requires – at least – a clear statement of the new lessons learned (from this data set) compared to previous studies. Hand in hand with this comment goes my suggestion that the introduction needs to present and discuss previous (similar) studies much more extensively than what has been done so far. This is important for the reader to understand in what way the present study of Cama addresses a new (open) question. RESPONSE: Ok, actually we think our paper give new insights in a so large topic such us the stochastically modelling of landslide phenomena aimed at producing predictive models and maps. In fact, we produced one of the few papers dealing with extreme rainfall triggered multiple debris flows events, exploring the effects of different triggering input on the same area, which corresponds to the very real topic for planners: what could I have predicted in 2007 with respect to the future landslides which has been then verified in 2009? We would expect that the performance demonstrated in this, should be the same for present day landslide-calibrated models with respect to the real future phenomena. Von Ruette et al. (2011), similarly to Lombardo et al. (2014) never apply a chrono-validation scheme. They calibrate their model with 2005 landslides and test it with the 2002 landslides of three other catchments, one of which actually is very close (but different) to the calibration one. Besides, should we consider the validation of the latter as a chrono-validation, they limit their modeling to a backward validation procedure. In

our paper, the very same area is analysed and forward and backward chrono-validation compared, with respect also to self-validated performances. Moreover, we would not consider the case study of Von Ruette as an extreme multiple debris flow event (which is thetyoe that could pose limits in chronovalidation models): they write about shallow landslides (debris flow?), but our 2009-case (616 in 2009, for a 10 km2 wide area) is quite out of the range of landslide frequencies and densities of their databases. To explore the effect of the different intensities of the trigger events, we used two inventories having very different (one order of magnitude!) cases. On the contrary, Von Ruette et alii moves between very similar cases. We now illustrate the point in the background, by enlarging the set of previous papers we refer to (as you suggested; see also rev. 1); besides, recalls to this point have been added in the conclusions section.

My final general comment concerns the language of the paper. Although the text is generally understandable I think that a careful language check by a native English speaker would be necessary to avoid formulations that seem to be wrong or complicated. For example, the authors speak about "more and more diffused databanks" (on page 1733, line 21) or "the operative validity of such expectation" (abstract line 6), which sounds odd to me. RESPONSE: Ok, we re-submitted the final version to a new mother-tongue professional English reviewer).

SPECIFIC COMMENTS - Page 1738, lines 13-15: the authors explain that the term "debris flow" is most appropriate for the observed landslides; but subsequently, they often use the term "landslide" (e.g. in the header of chapter 3.1 or on page 1741, lines 1-6). I suggest to be consistent throughout the entire manuscript in using the term "debris flow" or – even better – "hillslope debris flow". RESPONSE: Ok, done.

According to the text on page 1749, line 21 and subsequent, the final models for 2007 and 2009 included a different number and set of predictors. The significance of each of the predictors is only discussed in the text, but I'm missing a table summarizing the contribution and significance of each of the predictors to the model. Such a table would be a basis to make a comparison with other sets of predictors found in other studies. RESPONSE: Ok, actually we gave this information in figs. 10, 11, where you can see ranks, frequencies of selection (a) and beta coefficients, for the two models. According to your suggestion, we've now more discussed the role of factors.

- I think that the Figures 15 and 16 showing the difference in assessed susceptibility from the 2007 and the 2009 data set is the most interesting result and worth of a slightly extended discussion. For example, what could be the reason and the consequences of these differences? To what degree could the fact that antecedent soil conditions of the 2009-event may be influenced by the 2007 –event explain these differences? RESPONSE: Ok, we discuss in more details the point in the discussion section.

- Overall, the number of figures could be reduced. 8 Figures only showing the used data and describing the area and the events is a little bit too much. Fig. 9 is actually a Table and not a Figure. RESPONSE: Ok, Fig. 9 is now a table, while figs 2 and 3 are merged.

- Abstract, page 1732, line 4: "a past known landslide scenario" sounds incorrect to me. "past event" would be correct; scenario is future-oriented. RESPONSE: Ok, changed.

- Figures 2 and 3: the term "Hyetograph" is (to my knowledge) not correct; these figures show "time series" of precipitation, and not "distributions" of precipitation. RESPONSE: Ok, changed.

- page 1738, line 19: the main events were "preceded"

(not "anticipated") by rainfall events RESPONSE: Ok, changed.

- page 1739, line 15: It doesn't require (without s) RESPONSE: Ok, changed

---

- page 1746, line 21: ": : : can be obtained as proposed by (or as demonstrated by) Chung and Fabbri" RESPONSE: Ok, corrected.

- figure caption Fig 6: ": : : containing 73 debris flows (not "phenomena") RESPONSE: Ok, corrected.

Please also note the supplement to this comment:
http://www.nat-hazards-earth-syst-sci-discuss.net/3/C934/2015/nhessd-3-C934-2015-supplement.pdf

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., 3, 1731, 2015.

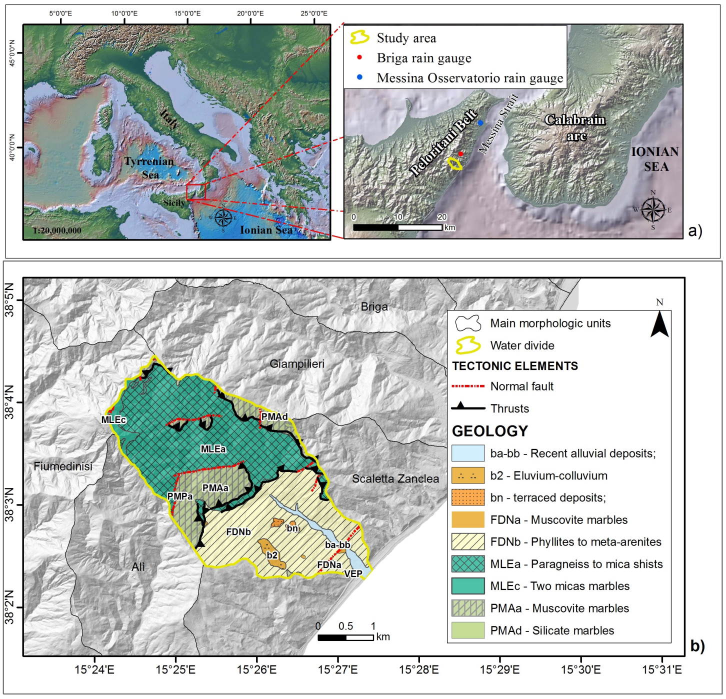

Fig. 1. Setting of the Study area

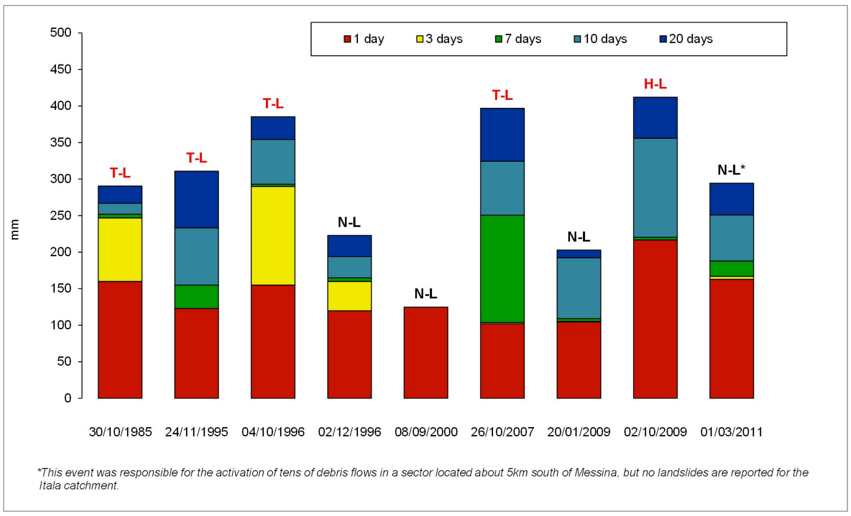

Fig. 2. Bar plot showing the cumulative rainfall in mm respectively for 1 day, 3, 7, 10 and 20 days.

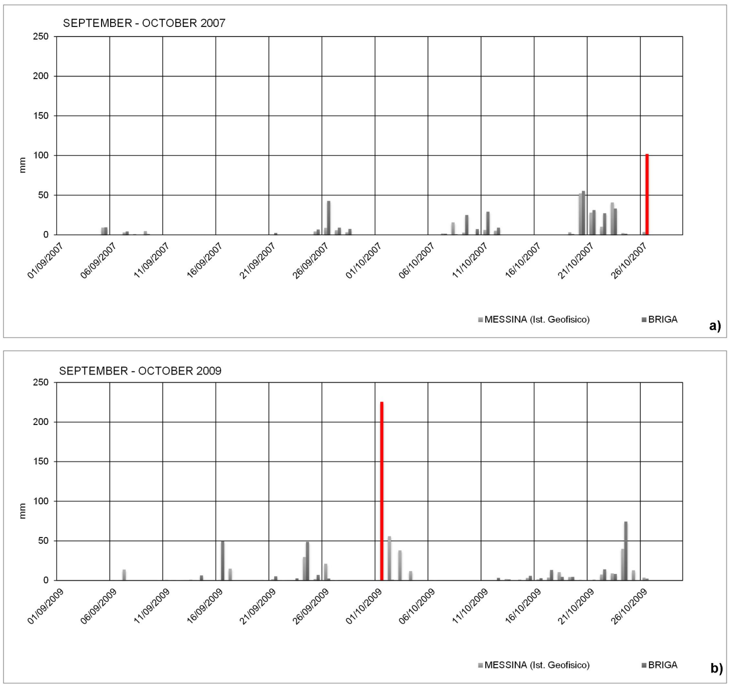

Fig. 3. Time series of 2 months precipitations: a) October 2007; b) October 2009.

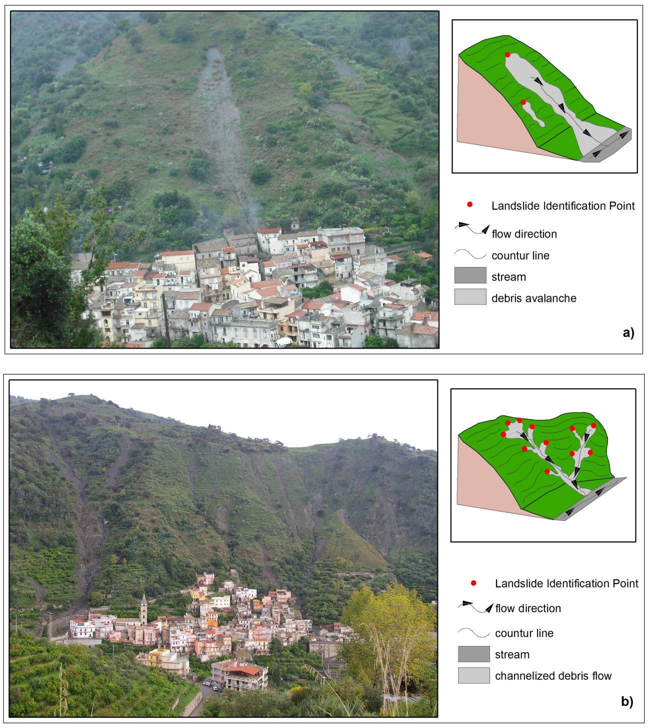

Fig. 4. Overview of the area hit by the 2009 event: a) Guidomandri village: debris avalanches are observable on the triangular facets parallel to the coast; b) Itala village: channelized debris flows crossing

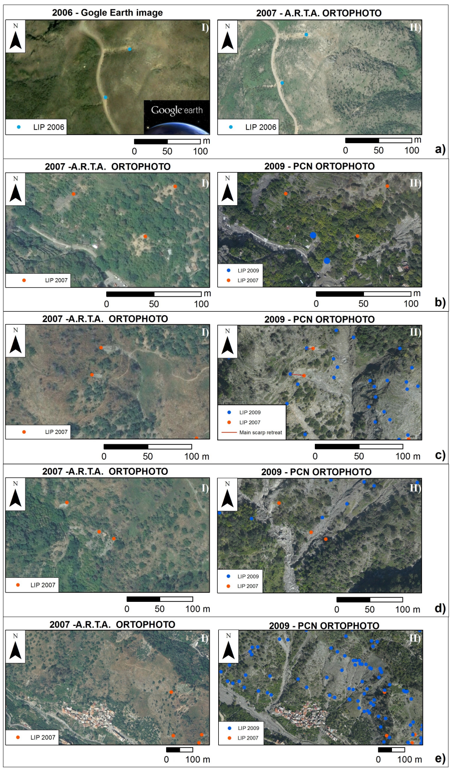

Fig. 5. Comparison of morphologies between two different images resulting in five different cases: a) debris flows recognized on the 2007 orthophoto but activated before the 2007 event; b) debris flows activa

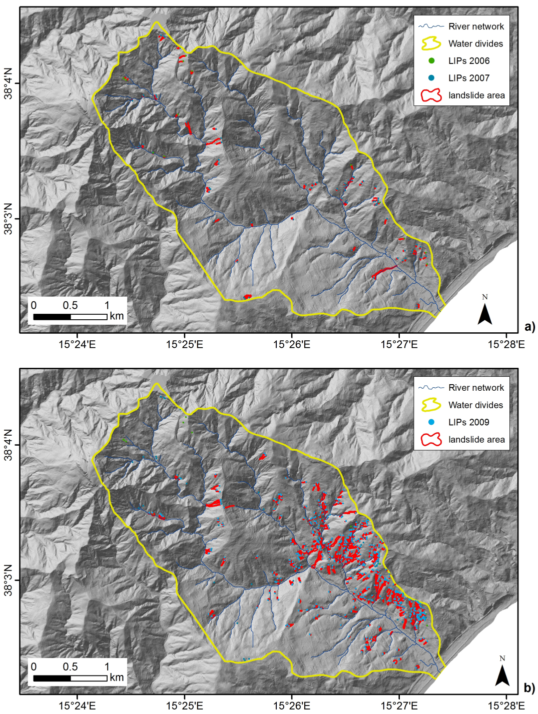

Fig. 6. Debris flow event inventories: a) 2007 inventory containing 73 debris flows; b)2009 event inventory containing 616 debris flows.

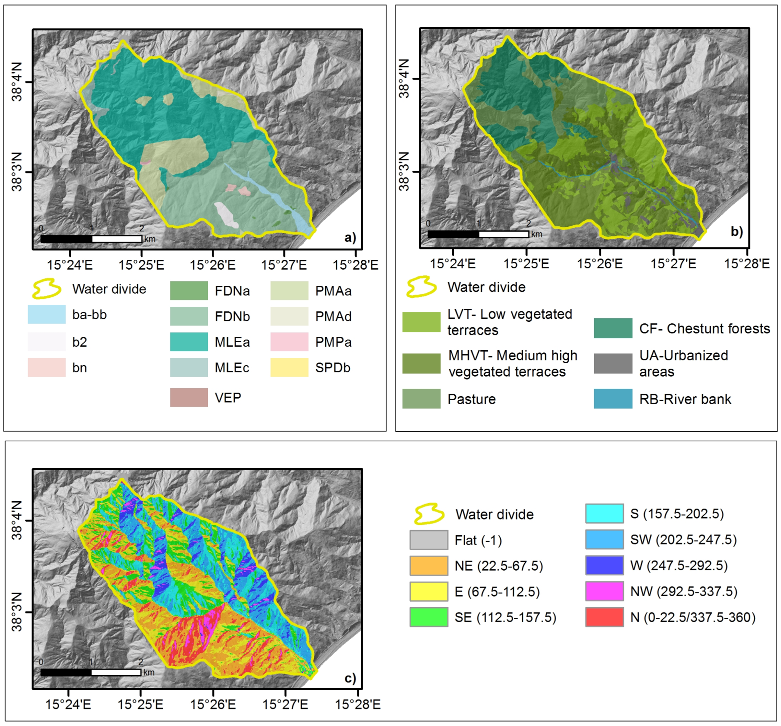

Fig. 7. Discrete variables: a) outcropping lithology (GEO; see Figure 1 for description); b) land use (USE); c)aspect (ASP).

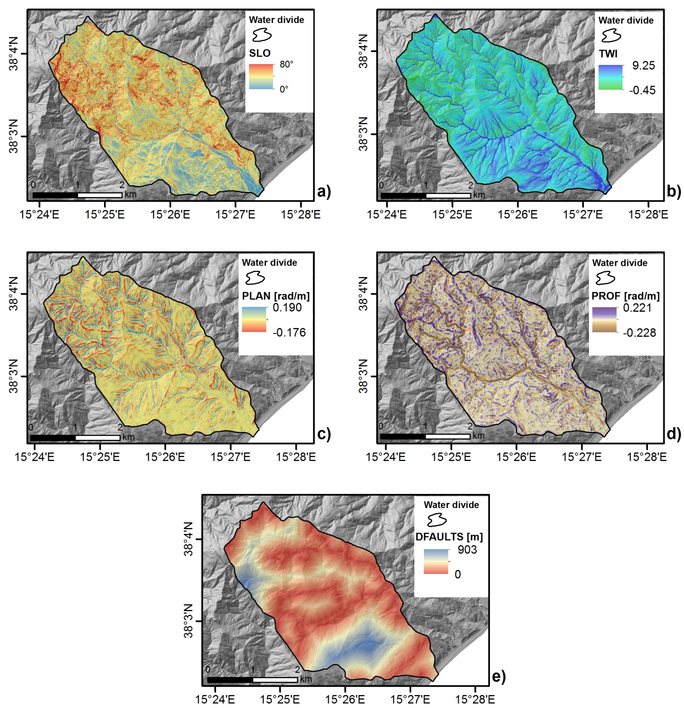

Fig. 8. Continuous variables: a) slope; b) topographic wetness index; c) plan curvature; d) profile curvature; distance from tectonic elements.

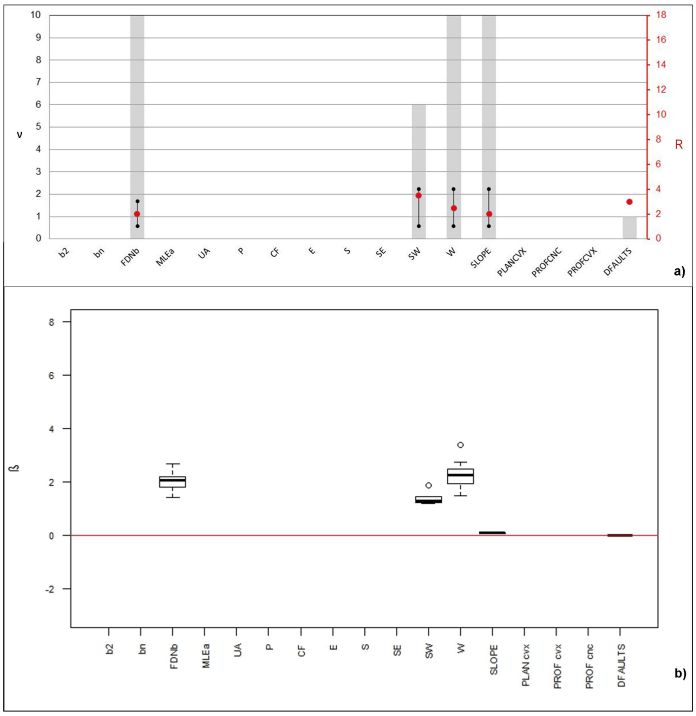

Fig. 9. Selected variables for the 2007 suite of models: a) ranking and frequency; b) β values.

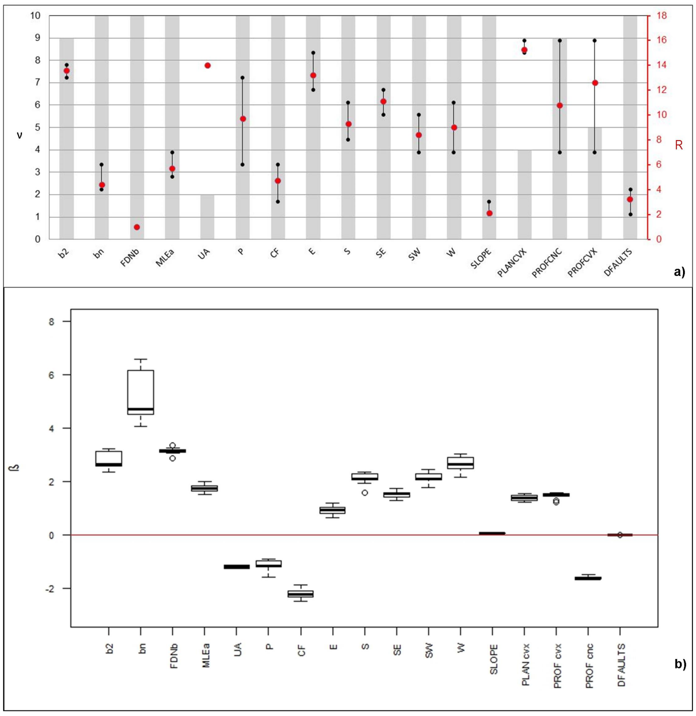

Fig. 10. Selected variables for the 2009 suite of models: a) ranking and frequency; b) β values. For need of representation, the coefficients of the topographic curvatures are reported as logβ values.

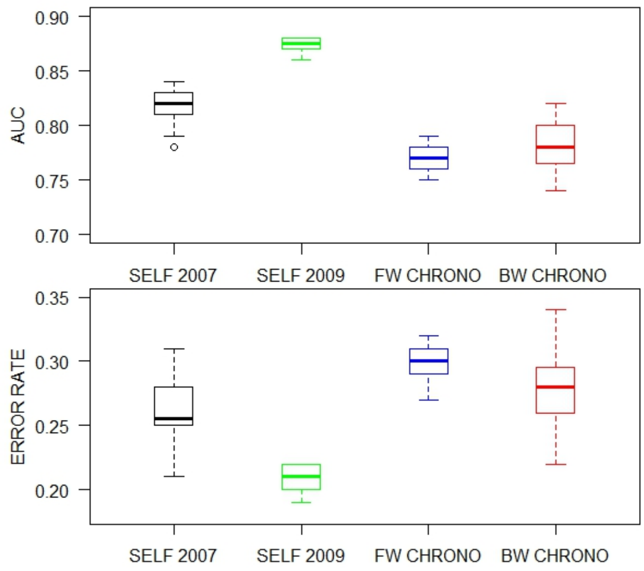

Fig. 11. Distribution of the AUC and error rate values calculated on the 10 replicates for 2007 and 2009 modelling and 100 models during the chrono-validation process.

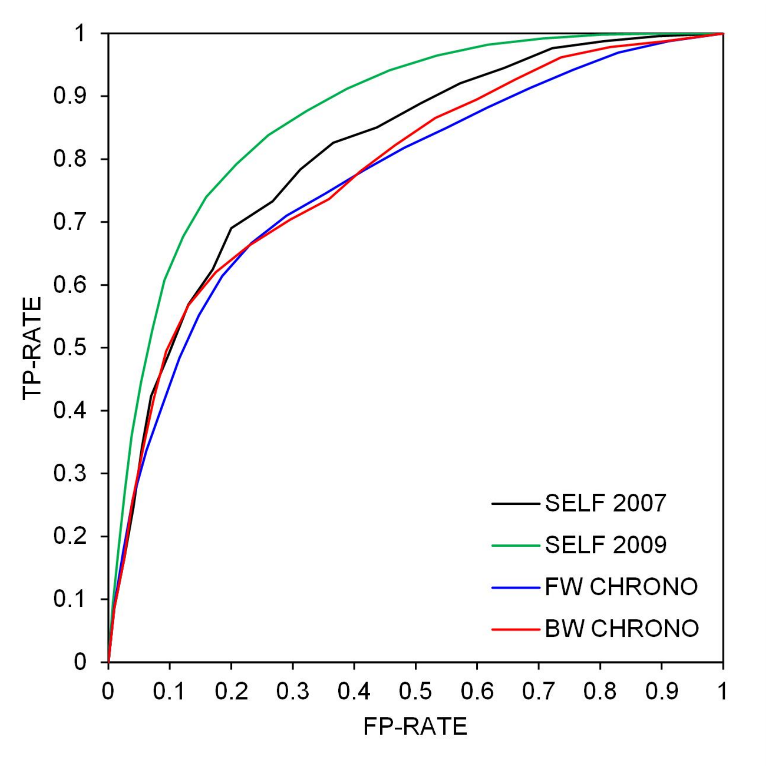

Fig. 12. Mean ROC curves.

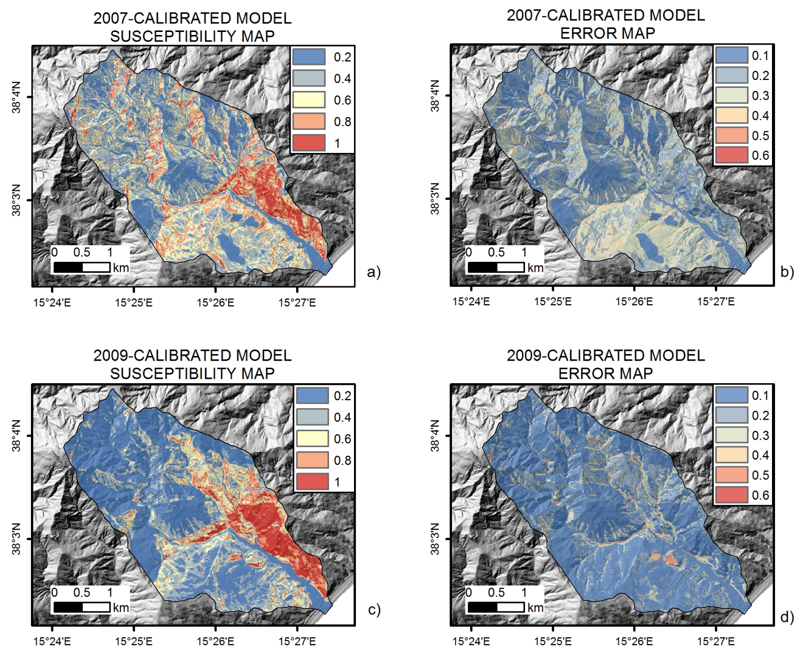

Fig. 13. Susceptibility and error maps for the 2007- and the 2009-calibrated models: a, c) mean susceptibility; b, d) error maps.

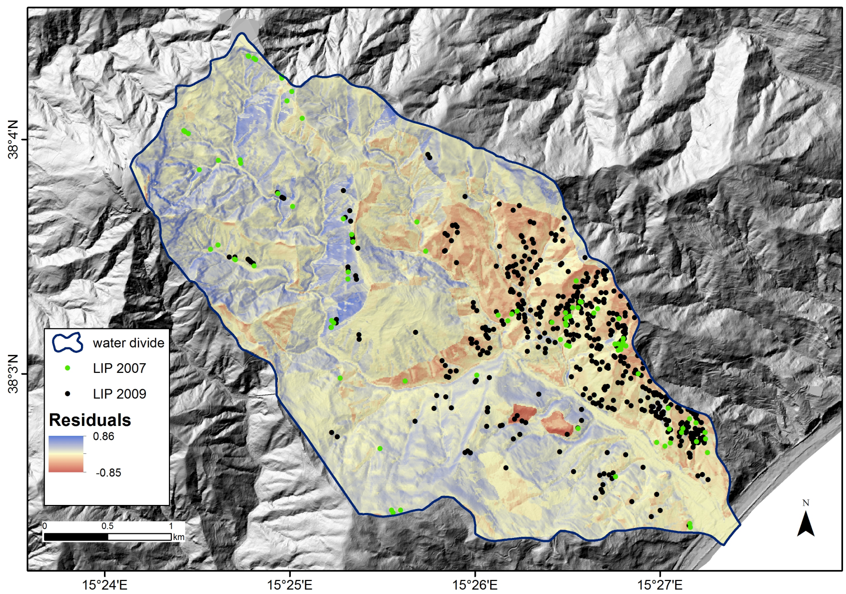

Fig. 14. Map of residuals calculated as percentage differences between the two (2007 and 2009) mean susceptibilities.

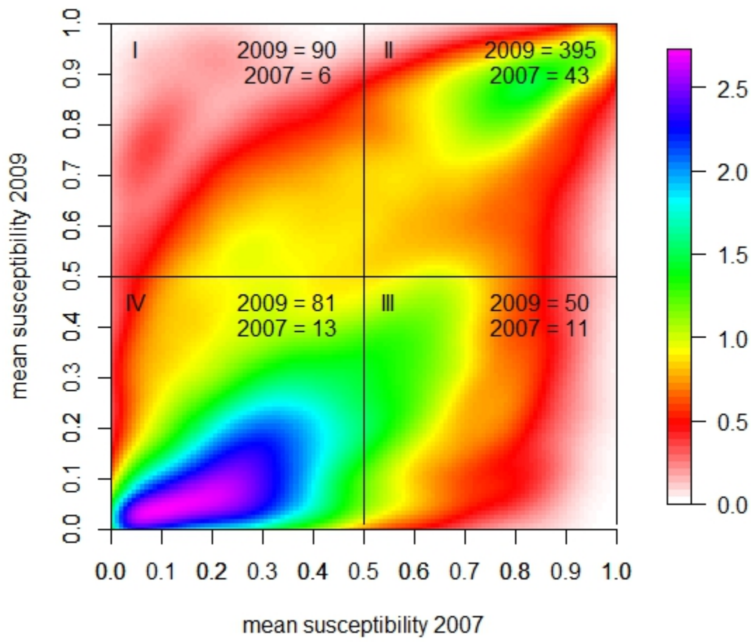

Fig. 15. Dispersion density plot calculated using 2d Binned Kernel Density algorithm (range for density calculation 0.045 xy). Positive cases for 0.5 cut-off values are reported for the two inventory events.