

## ***Interactive comment on “Spatial distributions of earthquake-induced landslides and hillslope preconditioning in northwest South Island, New Zealand” by R. N. Parker et al.***

**Anonymous Referee #2**

Received and published: 11 February 2015

General comments:

Parker and co-authors present a very interesting study of the factors controlling landslide failure during two earthquakes in New-Zealand that have occurred in the same area in 1929 and 1968. Using a logistic regression based on many potential factors driving failure, they demonstrate convincingly that for each earthquake seismic ground motion (or more precisely distance to the fault for the 1929 EQ) and hillslope gradient (with an influence of lithology) are the main predictors of failure probability. This first result is well supported by the data and analysis. By itself, it is worth being published in esurf as it adds new constraint on EQ triggered landslides. I have only few comments on this part which is very clearly written.

C5

The authors build on this first result by exploring whether the 1929 EQ could have pre-conditioned hillslope material to make the area closest to the 1929 fault line more susceptible to rupture (all other predictors being accounted for). This is a very interesting question which have hardly been addressed in the past. After some quick statistical analysis, the authors argue that the 1929 EQ has likely increased the failure probability during the 1968 EQ. However, I do not think the data analysis clearly support their conclusion. There might be a small effect, but this part of the paper lacks a more robust statistical treatment, and a more objective analysis of the results to be really convincing. The authors are trying too hard to see preconditioning in their data and are overlooking a careful discussion of the limitation of their approach. I'm afraid the uncertainties regarding the 1929 event (the landslide inventory is made 38 years after the 1st EQ, there is no pga data nor DEM prior to the event) are too large for the effect this EQ could have on the 1968 event to be computed precisely enough to. I hope I'm wrong as I'd really like preconditioning to be important, but for the moment the data and results do not support this notion. I have many detailed comments on this aspect amounting at a large revision of the MS.

Apart from this, the paper is well written, clearly organized easy to follow and the bibliography is adequate. On top of the major comment that I have regarding the demonstration of preconditioning, I have many minor comments that should not distract from the fact that this is a very good paper worthy of publication in esurf once properly revised.

Congratulations to the authors for a very interesting idea.

Detailed comments

P9L12: have you checked for any correlation between the size of the landslide and any other parameters studied in your logistic regression ?

P10L24: that ~ 18 % of the landslides occurred in over-steepened source areas (and the logistic regression results) shows that local slope is a critical element. Yet, you do not discuss your choice of the scale of measurement of the slope (why 90 m ?), nor the

C6

fact that you don't have a DEM prior to the 1929 and 1968 EQ.

P13L10-24 , Figure 9c: while I understand that the residuals could be correlated to PIs(1929) for high PIs(1929), the fact that there could be a trend at low PIs(1929) does not make sense to me. Should it not be uncorrelated here ? If the 1929 had little probability of failure, it should not reduce the probability of failure for 1968, but simply not change it.

P13L25: if I remember well the 10 m resolution is not a great DEM (e.g. staircase effects due to the contour lines of the 50000 maps it is derived from). Resampling it at 30 m is a good idea, however you should justify the choice of this spatial scale as it has many implications for the subsequent analysis (slope calculation, aspect etc...).

P14L23: for many landslide the scale at which the local hillslope gradient (SL) is measured is actually smaller than the landslide itself. Given that the DEM is post-1968, this means that the local hillslope gradient used in the logistic regression is posterior to landsliding and would typically be either smaller than before the EQ (in the center and at the base of the landslide) or much steeper (in the boundary of the landslide). Given that (SL) is a critical discriminant factor of the logistic regression, I think the authors should explore other values of the hillslope gradient (e.g., derived from unruptured area immediately close to the landslide), or estimated at a much larger scale. At least, they should discuss potential biases induced by using a DEM posterior to the EQs.

P15L24: I'm not familiar with McFadden's Pseudo  $R^2$ . Maybe giving a bit more details on how it is actually computed could help the audience to better understand its meaning.

IMPORTANT COMMENT: it is not clear how the observed PIs (which varies between 0 and 1) is estimated from the map of hillslope failures (binary values). This should be explained in detail.

P16L13: remove "this".

C7

P17L18: I don't understand the notion of size in the context of this sentence. Is it the magnitude, is it the size of the landslides ? Rephrase.

P18L10: "topographic amplification": this is an interpretation. You're just measuring the position with respect to the divide. Hillslope geometry (convex vs concave) could generate a dependency to NDS by the sole hillslope gradient dependency. This is why exploring the cross-correlation within your predictors could be interesting to reduce the number of meaningful ones.

P18L15: the following section is a bit weak as it does not objectively describe the results and try to push the idea that the variation in the residuals are determined by PIs 1929. There's also one thing not really clear in Figure 13: the amplitude and values of the residuals should theoretically be the same in each graph. We're looking at  $Y=[\text{Observed PIs}_{1968}-\text{Eq}(5)]$  as a function of various predictors and PIs<sub>1929</sub> and Distance from 1929 fault. I suppose that the variations in the amplitude of the residuals (from a range of 0.006 in fig. 13G to less than 0.00001 in fig. 13E) comes from the binning that varies with the chosen predictor. But this is far from being clear in the text or the legend, and suggests that there are actually very large variations in the residuals that the 'mean' is hiding (hence my request for displaying the standard deviation of each bin).

P18L19: this is not what I observe. In figure 13C there's a pattern (i.e. there's no random variation in the residuals) but you choose to ignore it, while you choose to see a trend in figure 13G and a correlation in fig13H !. You need to beef up the statistical analysis of these graphs.

P18L26: I don't see any correlation in Fig. 13h...quite the contrary.

P19L2: you could also mention that the pseudo- $R^2$  increases from 0.246 to 0.247 when refitting the model. However, I'm not sure I would qualify this as a significant effect.

Figure 10: a bit difficult to read and evaluate. At first order, it seems that pd3, pd6,

C8

SL,ES, ER are highly correlated. Any logistic regression would probably not been able to separate the effect of these 5 predictors. I'd suggest to explore cross-correlation and keep only the predictors that are truly independent of each other.

Figure 11: I'd suggest to invert the color scale for Pls such that it matches (to some extent) the observed map of hillslope failures. Comparing figures 11F and 11G, it seems that the residuals for 1968 are not much larger than the residuals of 1929 and are both patterned. This led me to 2 suggestions:

- the first one is to plot the residuals of 1929 against the predicted Pls(1929), to evaluate how much variability there's "naturally" in the 1929 model. The whole demonstration of preconditioning relies on the assumption that Pls1929 is a good predictor of observed failure probability, but what if it is not good enough because you don't have the pga for 1929 (which seems a critical element in 1968) or a DEM prior to the event ?
- Add for each graph of figure 13, the standard deviation for each quantile bins of the x variable. This would help in deciphering whether a trend in the average is statistically significant or not.

P20L1-3: you should first try to discuss the limitations of your approach, rather than first trying to push the idea that the 1929 EQ has influenced the 1968. I see at least three points to discuss:

- at this stage of the paper, it seems that the authors have forgotten that they are working on a very "strange" landslide inventory, taken 38 years after the first EQ over which the 1968 have superimposed new landslides. Clearly, the 1929 data is censored of small events (as the authors acknowledge themselves), but how these limitations could play a role in the model regression is never discussed. I suggest for instance to redo the calculation by censoring the 1929 with a very large minimum landslide size and see how it affects the subsequent calculations.
- given the small changes in probability you're looking at, how accurate and precise

C9

is Eq. (4) to predict the "impact" of the 1929 EQ on the 1968 EQ ? Given the lack of predicted pga for 1929, is it really realistic to assume that Pls1929 is good enough to detect a very small effect on the 1968 data ?

- what could possibility be the impact of building a logistic model using hillslope gradient measured after the EQ. and use it to predict failure probability on non-ruptured hillslopes ? That probably means that the sensitivity to hillslope gradient built into the 1929 logistic model is likely incorrect, or not as precise as it could be. I understand there's no easy way to sort for this effect, but it must be discussed given the importance of hillslope gradient.

P20L27 and end of discussion: to really make a case for considering preconditioning in landsliding susceptibility, you should give the reader an order of magnitude of how incorrect would be eq. (4) if it did not take into account previous effects : e.g. what would be the surface affected by landslides with or without preconditioning. As much as I like the idea, it seems your data are showing this is only a third or even fourth order parameter compared to pga, hillslope gradient and lithology (fig 12C). Maybe a more balanced and less arm-waving discussion would be better.

P21L8: no, you do not directly demonstrate the topographic amplification effects, you postulate it.

P21L9: as above you do not demonstrate that hillslope weathering is the predictor. It is just hillslope orientation (which in itself is a great result that you could emphasize better in the discussion!). You could also emphasize in the conclusion that many tested parameters do not appear critical for predicting EQ triggered landslides !

P21L24: point 1 of the conclusion clearly states that failure probability is explained at 90% by pga and hillslope gradient (with litho effect). How come that the current damage state of hillslopes represents a significant source of uncertainty ? Again this looks like unnecessary arm waving given the uncertainties in the data and subsequent analysis.

C10

