SUPPLEMENTARIES

Methods and Metrics

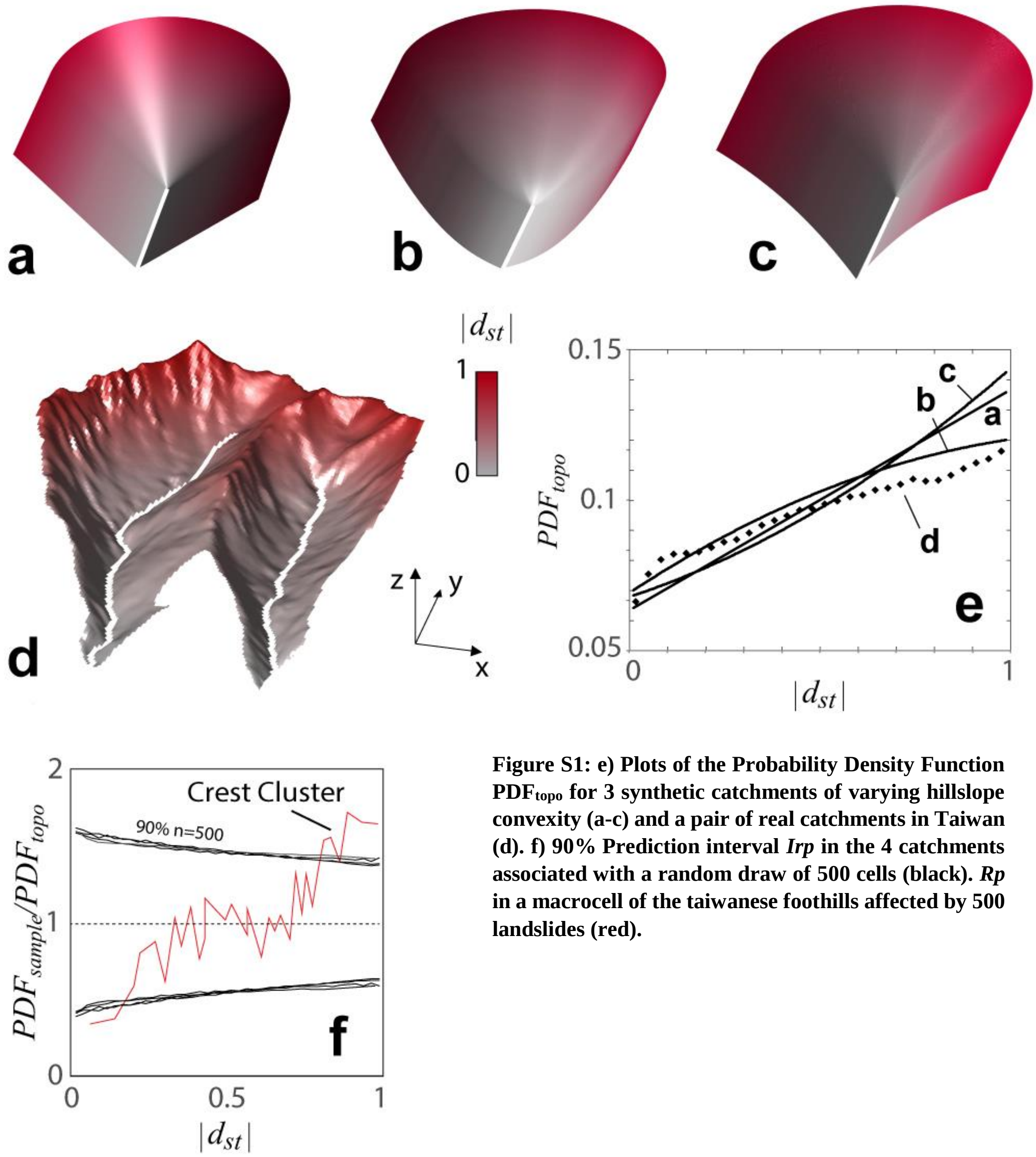

Figure S1: e) Plots of the Probability Density Function PDF_{topo} for 3 synthetic catchments of varying hillslope convexity (a-c) and a pair of real catchments in Taiwan (d). f) 90% Prediction interval *Irp* in the 4 catchments associated with a random draw of 500 cells (black). *Rp* in a macrocell of the taiwanese foothills affected by 500 landslides (red).

Every cell in the landscape is defined by its normalized distance to the river network $|d_{st}|$ expressed as:

$$|d_{st}| = \frac{d_{st}}{d_{st}+d_{tp}}, \qquad (1)$$

where d_{st} and d_{tp} are its flow distances to the nearest river and the nearest crest respectively. $|d_{st}|$ ranges from 0 for cells located in the floodplains to 1 for cells located on the crests. In any part of the landscape, one can compute the probability density function (*PDF*) of a cell being at a given $|d_{st}|$ over the interval (0,1). The boundaries are kept opened in order to keep the *PDF* unaffected by the potentially large variations of ridge/river densities introduced by the use of different *DEM* or ridge/river mapping methods. Figure S1 shows examples of this *PDF* in 3 synthetics catchments of straight, concave and convex hillslopes respectively. Whatever the hillslope curvature, the *PDF* is a monotonic function with no asymptotic behavior toward nods.

To study the spatial variation of the landslide location across the epicentral area, we divide the landscape into macrocells within wich PDF_{topo} and PDF_{ls} are computed. PDF_{topo} is built from the distribution of every points in the macrocell while PDF_{ls} only accounts for cells affected by landsliding. The *PDF* ratio R_p is then defined as:

$$R_p = \frac{PDF_{ls}}{PDF_{topo}} \qquad (2)$$

If the landscape is uniformly sampled by landsliding along $|d_{st}|$, $R_p=1$ over (0,1). In the right region of the curve, high values of R_p $(>>1)$ express a significant crest oversampling by landslides while low values (<<1) express undersampling. Inversely, a large R_p in the left region of the curve expresses hillslope toe clustering (Fig. S1f).

Statistics

Plots of R_p computed in macrocells with low number of landslides exhibit a high variability that is not statistically significant. To correct from this bias, we use the central limit theorem (*CLT*). As the area affected by landsliding is generally less than 10% of the landscape, the sampling of the topography by landslides can be approximated by a Bernoulli sampling. *CLT* then gives the 90% prediction intervals as

$$I_p = \left[p - 1.96\sqrt{\frac{p(1-p)}{n}}; p + 1.96\sqrt{\frac{p(1-p)}{n}}\right], \qquad (3)$$

for a given value $PDF_{ls}(|d_{st}|_i)=p$ and n numbers of independent random draws. The convergence towards a normal distribution of n draws centered on p also requires $n>30$, np and $n(1-p)>5$. By construction, the 90% prediction intervals on R_p is then defined as

$$I_{Rp} = \left[1 - 1.96\sqrt{\frac{(1-p)}{np}}; 1 + 1.96\sqrt{\frac{(1-p)}{np}}\right], \qquad (4)$$

Figure S1f shows the 90% interval on R_p in the 3 synthetic catchments mentioned above for 500 points randomly drawn in the *DEM*. Note than as the *PDF* are monotonically growing with $|d_{st}|$, the prediction interval is generally smaller in the right region of the plot, *i.e* near the crests.

As the Probability ratio R_p is built from the normalized ratio of area of given $|d_{st}|$, n should be the number of cells affected by landsliding in the macrocell. But this method introduces a bias: as a landslide is composed of several cells, for any cell i affected by landsliding of given $|d_{st}|_i$, its neighboring cell j has a higher probability of being at $|d_{st}|_j \approx |d_{st}|_i$. In this approach, the draws are not independent anymore and the Bernoulli sampling hypothesis is not met. We bypass this problem by defining n as the number of landslides included in the macrocell.

In the case of a random sampling along $|d_{st}|$, Rp should be confined within I_{rp}. As a consequence, we define "clusters" as regions of the plot where Rp is found to be beyond I_{rp}. Figure S1f shows an example of Rp computed in a macrocell affected by 500 landslides (red) and exhibiting crest oversampling.

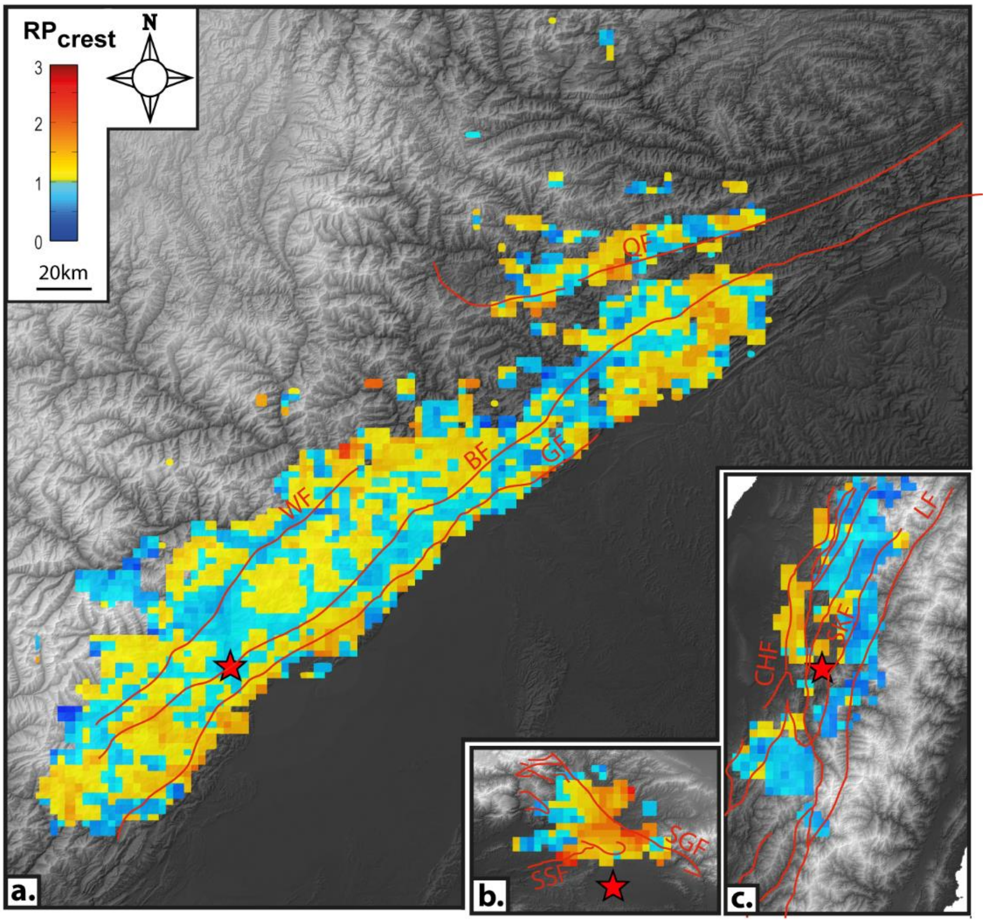

Figure S2: Rp_{crest} values in the a. Wenchuan, b. Northridge and c. Chi-Chi epicentral areas. The 3 maps are at the same scale. All statistically meaningful macrocells are represented (*i.e.* verifying the 3 conditions for normal law convergence). Main faults are represented by red lines and epicenters by red stars. WF: wenchuan fault; BF: Beichuan fault; GF Guanxian fault. b. SSF: Santa Susanna fault; SGF: San Gabriel fault .c. CHF Chelungpu thrust fault; SKF: Shuilikeng fault; LF: Lishan fault.

In every macrocells, we have defined $Rp_{crest} = \overline{Rp_{[0.75-1)}}$ and $Rp_{toe} = \overline{Rp_{(0-0.25]}}$ as the mean value of Rp over the upper and the lower quarter of the hillslope respectively. Crest-clustering is thus found in macrocells within wich $Rp_{crest} > \overline{IRp_{crest}}_{max}$ while toe-clustering corresponds to macrocells of $Rp_{toe} > \overline{IRp_{toe}}_{max}$. Macrocells of $\overline{IRp_{crest}}_{min} < Rp_{crest} < \overline{IRp_{crest}}_{max}$ are not represented on Fig. 2 and should be interpreted as regions of uniform sampling. Maps containing unsolved Rp_{crest} are represented on Fig. S2. The patterns identified as clusters on Fig. S2 are separated by zones of $R_p \sim 1$, i.e. of uniform sampling. Figure S3 shows that crest clustering is generally equivalent to toe undersampling and vice versa. As a consequence, regions of toe-clustering are represented by macrocells of $Rp_{crest} < \overline{IRp_{crest}}_{min}$ on Fig. 2 and S2.

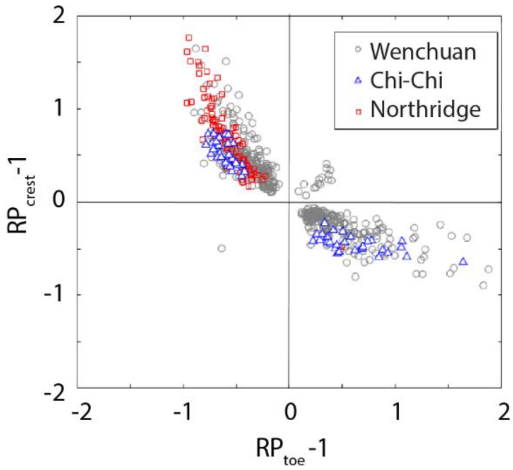

Figure. S3: Values of *RP_{crest}*-1 plotted along Values of *RP_{toe}*-1 in every macrocells of the 3 inventories. The inverse correlation shows the absence of double clustering.

Validity of the metrics

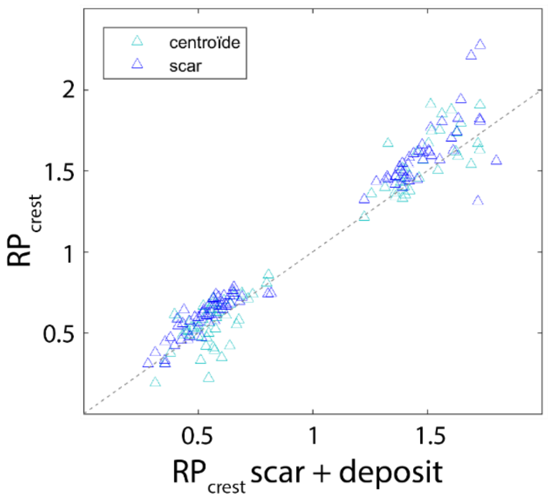

Figure S4: Comparison of the crest clustering values (*RP_{crest})* in Chi-Chi obtained using the total landslide surface with the one obtain using the landslide centroid (light blue triangles), and the landslide scar (dark blue triangles) considering Domej et al, 2017 observations. A quasi 1:1 relation is observed between the methods.

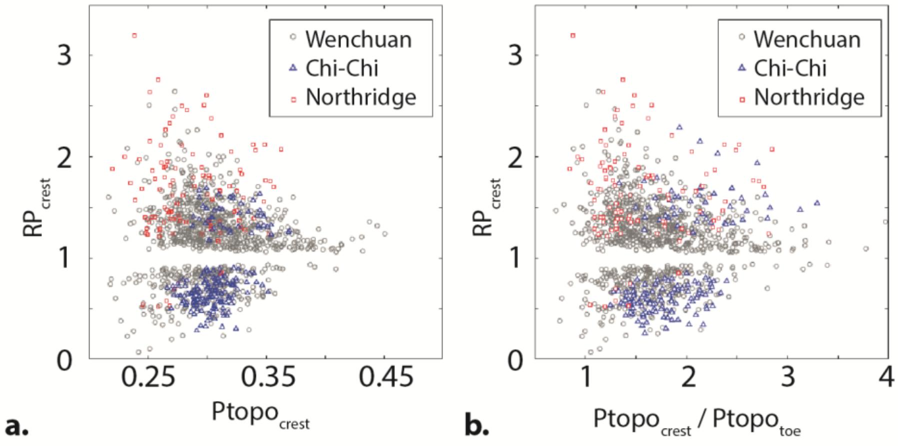

Figure S5: *RP_{crest}* plotted along a. *$Ptopo_{crest}$* and b. *$Ptopo_{crest}/Ptopo_{toe}$* for the three study area. The two plots show no correlation, insuring that the crest clusters are independent of the amount of landscape standing along crest/river in the landscape.

The method we introduce aims at defining the landslide position independently of the distribution of area with $|d_{st}|$ in the landscape. This condition is satisfied since Rp_{crest} is uncorrelated to $Ptopo_{crest} = \overline{PDF_{topo[0.75-1)}}$ and $Ptopo_{toe} = \overline{PDF_{topo(0-0.25]}}$ (Fig. S5a and S5b).

Dependence on the dataset

Three landslide databases are available for the Wenchuan earthquake. The Xu et al, 2014 has the higher number of landslides and cover the larger area. We notice that the number of landslides tends to converge to a maximul value for landslide densities above 10^{-2} in the Parker *et al.*, and Gorum *et al.* datasets, while it still increases in the Xu *et al.* one. This proves the presence of landslide amalgation in some places in the Parker *et al.*, and Gorum *et al.* datasets.

Figure S7 shows the *Rp_{crest}* maps obtained using these three databases. The one resulting from the *Xu et al*, 2014 covers a larger area than the two others.

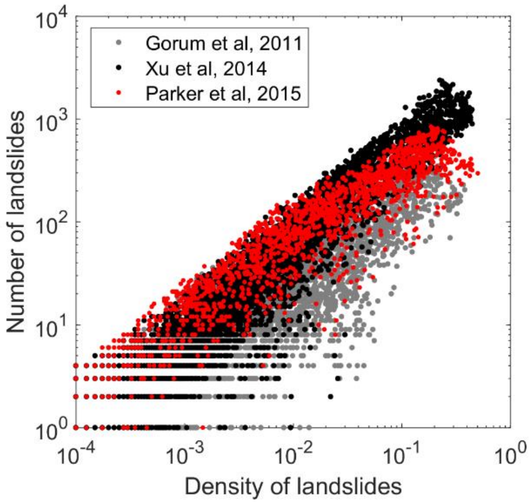

Figure S6: Number of individual landslides plotted with landslide density (ratio of surface covered) computed in each macrocell using the 3 landslide databases of the Wenchuan case: Gorum et al., 2011; Parker et al., 2015; Xu et al., 2014. The more precise is the catalog, the more small landslides there is. Amalgation and over mapping are observed in Parker et al., 2015 and Gorum et al., 2011 inventories.

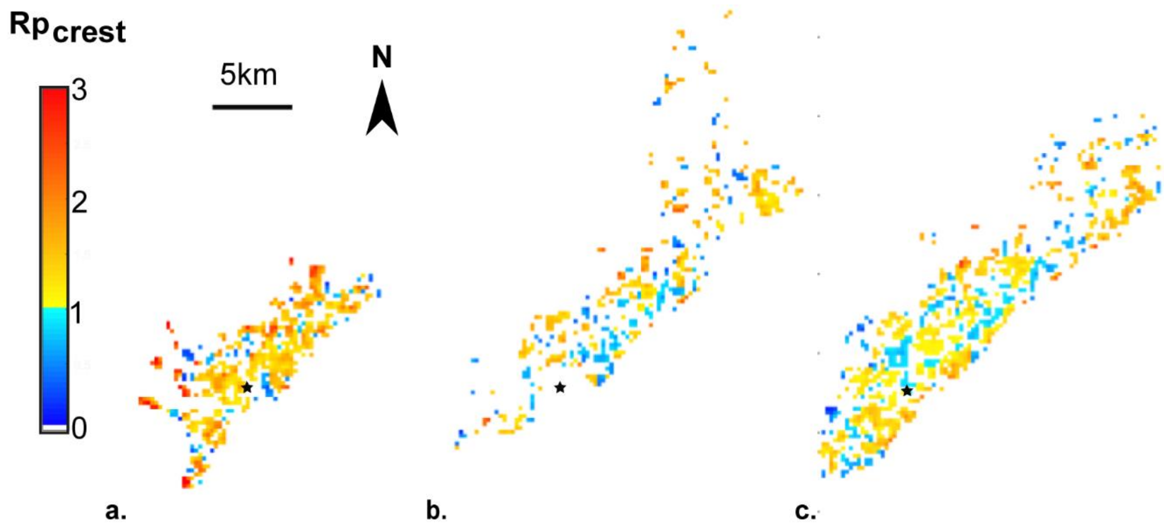

Figure S7: ***Rp_{crest}*** **maps obtained using the a. Parker et al. 2015, b. Gorum et al. 2011 and c. Xu et al. 2014 landslide databases. The Wenchuan earthquake epicenter is represented by the black stars. The Xu et al, 2014 one covers a larger area and has more statistically valid macrocells than the two others.**

Extraction of topographic features

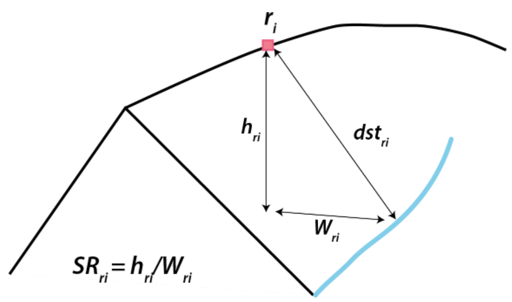

Figure S8: Extraction of topographic features on each ridge cell. *h_{ri}* ridge relief, distance to the stream *dst_{ri}* and half base width *W_{ri}* measured at a given ridge point *ri*. The shape ratio S_{ri} is defined as the ratio of *h_{ri}* to *W_{ri}*.

The ridge relief h_{ri}, is defined as:

$$h_{ri} = H_{ri} - \min(H_j) \,, \qquad (5)$$

where H_{ri} is the elevation of the r_i and H_j is the elevation of a point j distant of dst_{ri} from r_i. The half base width of a hill on a *ri*, W_{ri}, is calculated as:

$$W_{ri}^{\,2} = \frac{h_{ri}^{\,2}}{dst_{ri}^{\,2}} \,, \qquad (6)$$

The shape ratio at a given ridge point S_{ri} is defined as the ratio of ridge relief h_{ri} to the half-width W_{ri}.

Additional information on clustering controls

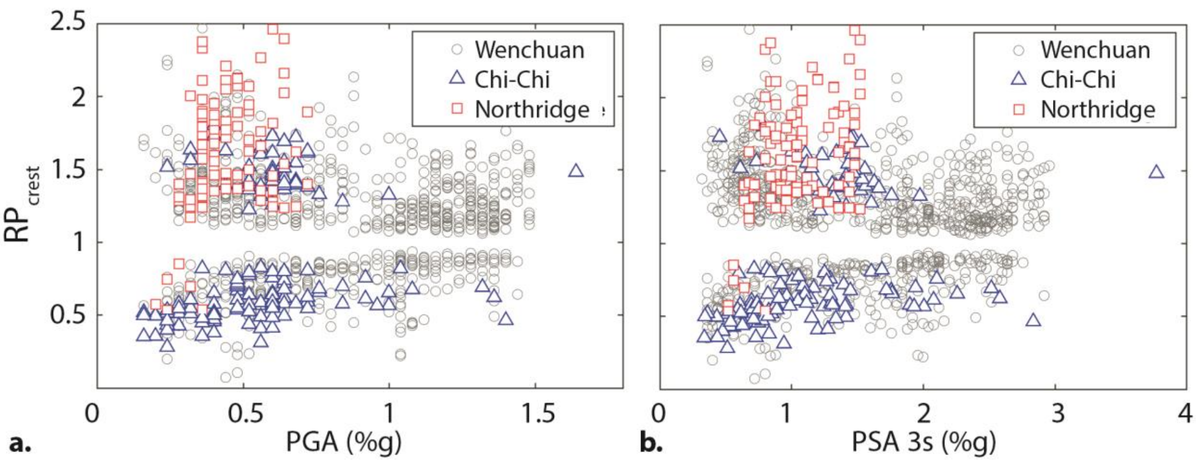

Figure S9: *Rp_{crest}* as a function of seismic features: a. Peak Ground Acceleration (PGA) (%g), b. Pseudo Spectral Acceleration at 3s (PSA 3s) (%g) for the Wenchuan, Northridge and Chi-Chi epicentral areas. Regional seismic parameters do not seem to explain landslide position along hillslope.

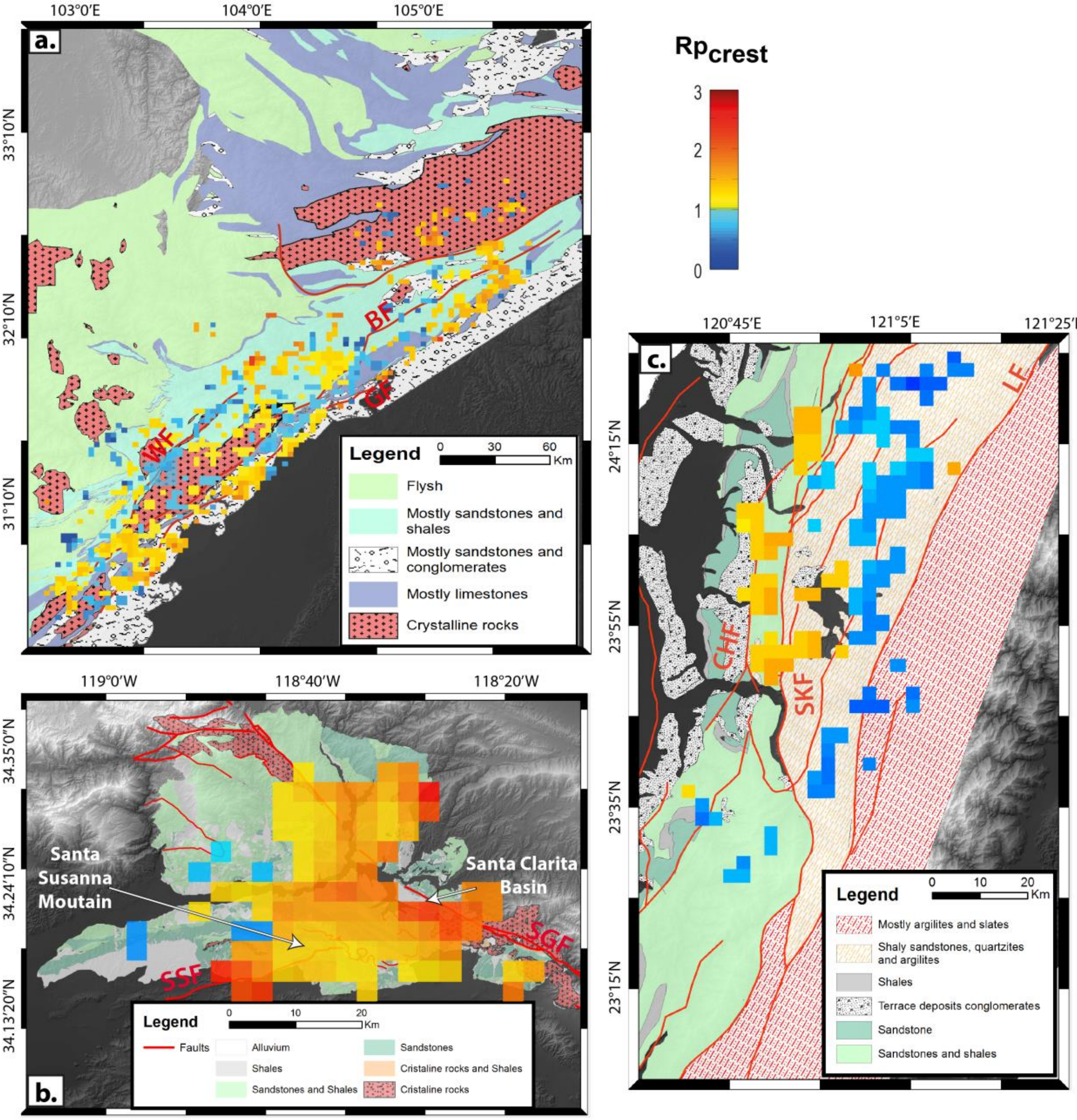

Figure S10: Lithological unit maps superimposed by Rp_{crest} maps in the a. Wenchuan, b. Northridge, c. Chi-Chi epicentral areas. The main faults are represented by red lines (a. WF: Wenchuan fault; BF: Beichuan fault; GF Guanxian fault. b. SSF: Santa Susanna fault; SGF: San Gabriel fault .c. CHF Chelungpu thrust fault; SKF: Shuilikeng fault; LF: Lishan fault). The upper slope clustering 90% maps are represented in transparency (Rp_{crest}>1: crest-clustering, Rp_{crest}<1: toe-clustering).

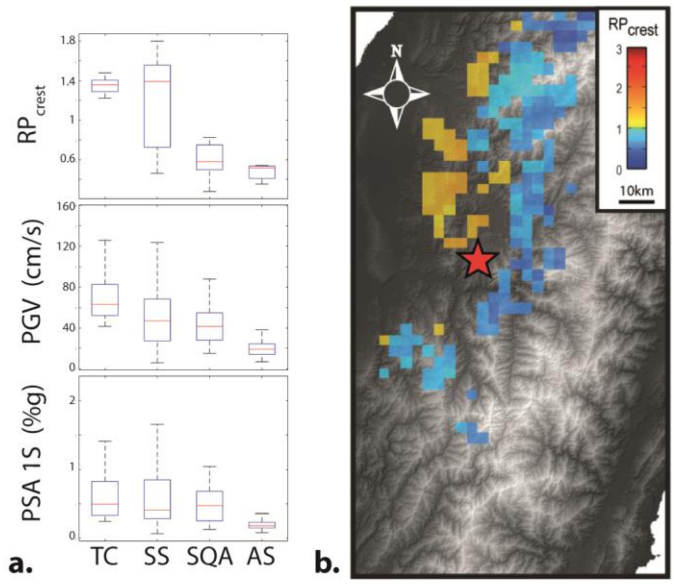

Figure S11: Rp_{crest}, *PGV* and *PSA* distributions plotted in a. lithological groups and b. Rp_{crest} map in the Chi-Chi epicentral area. TC: terrace and conglomerates SS: sandstones and shales SQA: shaly sandstones quartzite and argillites AS: argillites and slates. Both *PGV* and *PSA 1s* decrease with rock strength. Consequently, it is uneasy to dissociate ground motion control from lithological control on the reduction of Rp_{crest} toward the east.

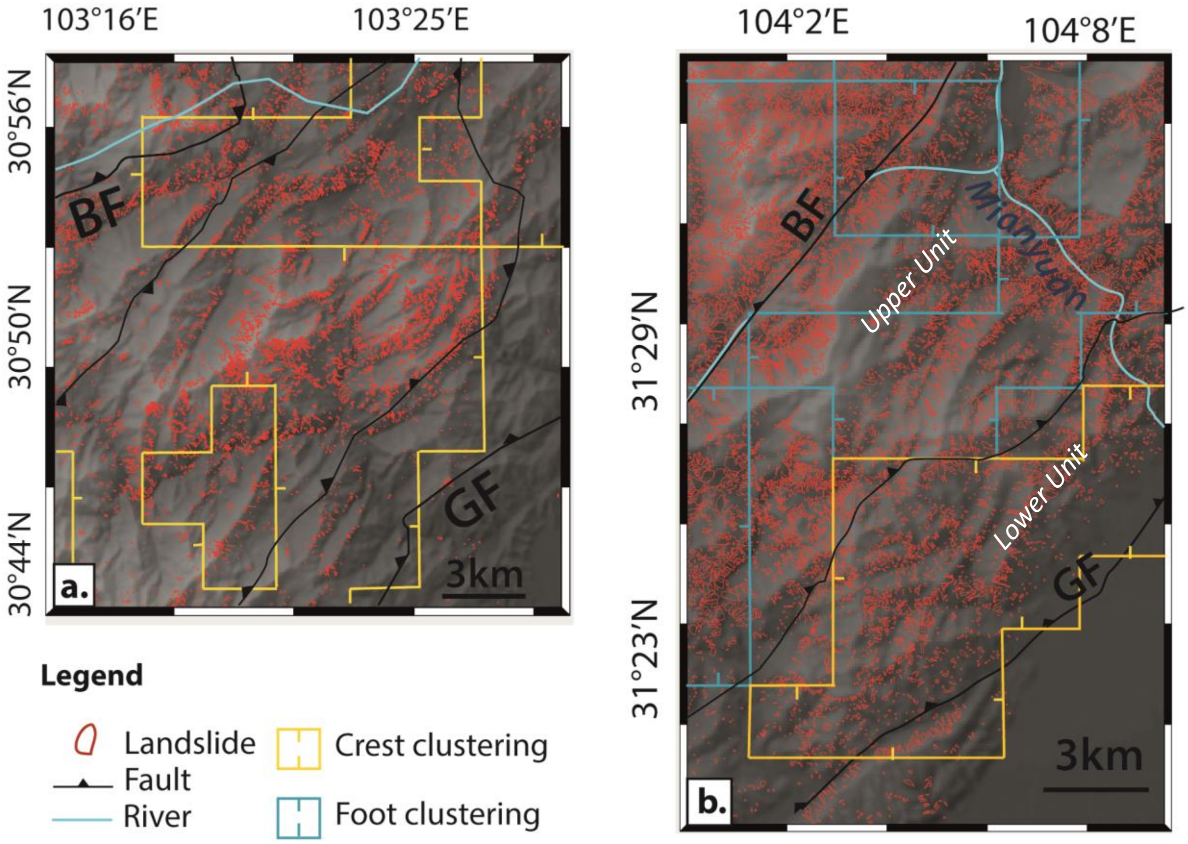

Figure S12: Snapshots of the landslide maps in a. the Sanjiang klippe and the b. foothills. The locations of a. and b. are reported in Fig. 8a. In the lower unit of the central zone and the Sanjiang klippe the landslides cluster around the

crests. In the upper unit of the central zone the landslides cluster downslope along the Beichuan Fault. BF: Beichuan fault; GF Guanxian Fault.

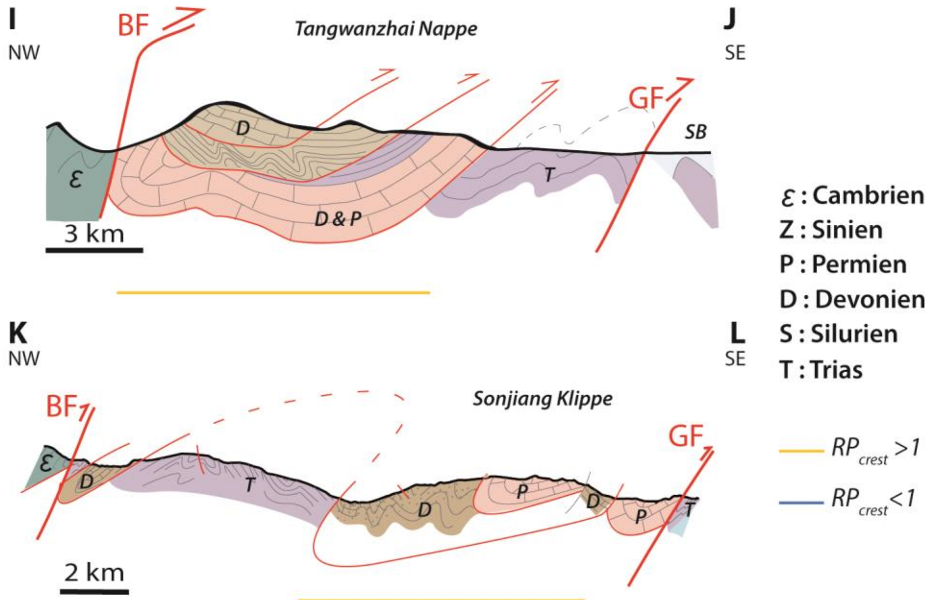

Figure S13: Cross sections of the South of the Tangwanzhai nappe (I-J) and of the Songjiang Klippe (K-L). Cross section locations are reported in Fig. 8a. BF: Beichuan fault; GF Guanxian Fault (after Robert 2011).

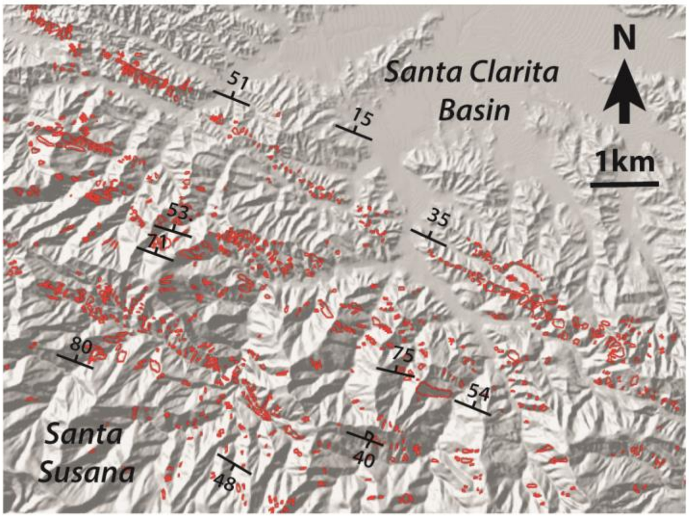

Figure S14: Northridge earthquake-induced landslides in the Northern flank of the Santa Susana Mountain. Landslides are represented by red polygons. Most of the landslides are concentrated along the ridge crests following the stratigraphic layers.

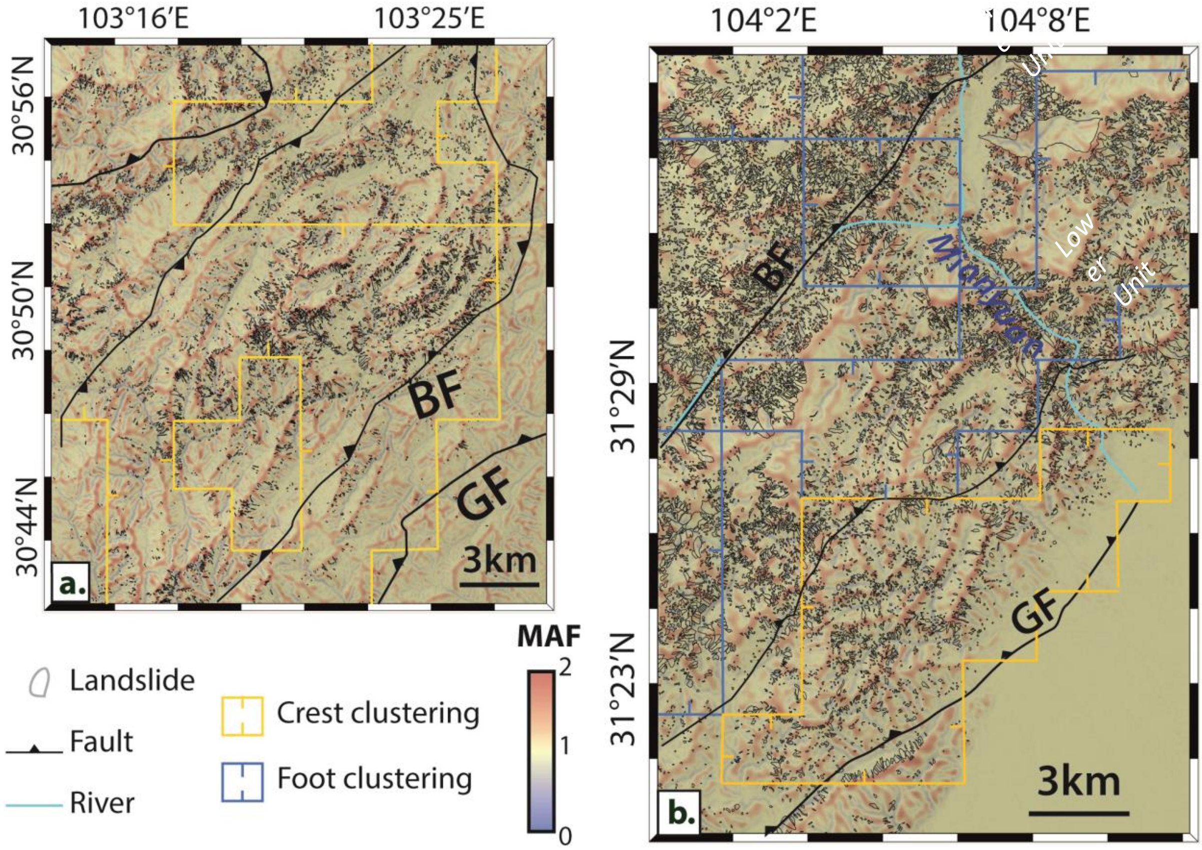

Figure S15: Position of the landslides along hillslope compared to the Maximum medium Amplification Frequency scale curvature (MAF) over S-wavelength 500-833m in Sichuan. The location of a. and b. is reported in Fig.8a. In the lower unit of the central zone (b.) and the Sanjiang klippe (a.) the landslides cluster at some place around the ridge crests where the predicted amplification is close to 1.8. BF: Beichuan fault; GF Guanxian Fault.