Supplementary Methods, Table and Figures

Title: What controlled the occurrence of more than 116,000 human-mapped landslides triggered by Cyclone Gabrielle, New Zealand?

Journal: Landslides

DOI: 10.1007/s10346-025-02591-y

This document provides supplementary information relating to the data and methods used to analyses it, along with more details relating to specific results from the analyses.

S1.0 Additional methods

S1.1 Landslide Analyses Methods

Identifying the main factors controlling landslide occurrence within the training dataset.

Step 1: frequency-ratio method (e.g., Kritikos et. al. 2015) was used to identify and describe any relationships between the selected factors (listed in Table 1) and the landslides mapped within the priority mapping grid cells of the Training AOI. Refer to S1.1.1 for more detail.

Step 2: Logistic LASSO (L1) regression was used to identify the main factors that could be used to explain landslide occurrence, whilst reducing overfitting. This step identified the minimum number and therefore most important factors that could be used to explain landslide occurrence.

We performed a 10-fold cross-validation to identify the optimal lambda value (hyper parameter) using minimum average error. The initial run was on the entire data to get the lambda sequence. The cross-validation procedure was used to split the input data into 10 folds and the algorithm was run with each fold omitted. The average errors were calculated over each fold and their standard deviations computed. The errors were calculated using the mean squared error, which uses the squared residuals to calculate the mean cross-validated error. The logistic LASSO regression coefficients for the chosen factors were not used for hindcasting as the coefficients produced by the LASSO model are biased[1]

Step 3: The 'best-fitting' factors derived from Step 2 were further tested to explore the relationship between landside occurrence and the given factor(s). To do this, each of the continuous factors was standardised and used in logistic regression model training to explore the relative importance of each factor. The categorical factor of Land Cover Categories (LCC) was not standardised. Refer to S1.1.2 for more detail.

'Boot strapping' was used on the standardised dataset to estimate the variance of the coefficients for each factor and their relative 'weights' in explaining landslide occurrence. 100 boot strap samples were chosen, with each sample containing a randomly selected 75% of the training data, with 25% excluded for testing.

To ensure the models weren't over correlated, the correlation matrix of estimates was calculated for each pair of factors included in the logistic regression modelling. For this research any 'pairs' of factors with correlation values -0.5< X >0.5 were removed.

[1] https://crunchingthedata.com/when-to-use-lasso/

Once the best-fitting factors had been derived and checked, the same boot strapping approach was used to derive the logistic regression coefficients that could be used for hindcasting. This time model fitting was done using the actual values for the input factors, not the standardised values.

1.1.1 Frequency-ratio

The frequency ratio is the relative frequency of grid cells that are affected by landslides (Y = 1) and their associated variables, compared with the relative frequency of cells that are not landslides (Y = 0) and their associated variables. All grid cells within the Training AOI were used to calculate the failure frequencies for each variable tested, using the following equation:

$$Frequency\ ratio\ = \frac{N_{(Li)} \div N_{Ci}}{N_{(L)} \div N_{(A)}} \quad \text{Equation. 1}$$

where $N_{(Li)}$ is the number of grid cells that are classified as being a landslide in the variable *i*, $N_{(Ci)}$ is the total number of grid cells in the variable *i*, $N_{(L)}$ is the total number of grid cells in the dataset that are classified as being a landslide, and $N_{(A)}$ is the total number of grid cells in the dataset.

1.1.2 Logistic regression

Logistic regression (LR) is routinely used to forecast landslide occurrence (e.g. von Ruette et al. 2011; Budimir et al. 2015; Lombardo and Mai 2018) by calculating landslide probability from a series of input 'predictor' factors. The method is used to estimate the coefficients for predicting the probability (P_{LS}) that $Y = 1$, given the values of one or more predictor factors. Here, the condition $Y = 1$ corresponds to the occurrence of a landslide and $Y = 0$ corresponds to no landslide, within a sample grid cell on the ground. The regression coefficients were estimated using a maximum likelihood criterion. Each model was fitted using the equation:

$$P_{LS} = \frac{1}{1+exp^{-z}} = \frac{exp^{z}}{1+exp^{z}} \quad \text{Equation. 2}$$

where the probability of a landslide occurrence (P_{LS}) is within the range 0 to 1 on a sigmoid curve; Z is a linear fitting function for the set of landslide-related factors tested (Table 1), expressed in the form of:

$$Z = b_0 + b_1X_1 + b_2X_2 + \ldots b_nX_n \quad \text{Equation. 3}$$

and where b_0, b_n and X_n represent the intercept of the model, the partial regression coefficients 'parameter estimates', and the conditioning factors, respectively. The parameter estimates were derived by statistically comparing the mapped landslides with the features listed in Table 1, using the Statistica software (TIBCO Software Inc. 2020). For investigating the factors that best describe the landslide source-area locations, we used an unbalanced dataset, given that the occurrence of a landslide ($Y = 1$) is much rarer than not ($Y = 0$). Therefore, our dataset is inherently biased toward no landslides ($Y = 0$) and, as a result, are naturally unbalanced. The regression coefficients were estimated in Statistica by maximising the likelihood function. The maximisation of the likelihood is achieved by an iterative method called Fisher scoring. The algorithm completes when the convergence criterion is satisfied or when the maximum number of iterations has been reached. Convergence is obtained when the difference between the log-likelihood function from one iteration to the next is less than 0.0000001.

For a factor to be included in the model, it must have a logical and statistically significant influence on P_{LS}. Although LASSO regression was carried out to determine the minimum number of statistically important factors required by each mode, we also used a significance level (p-value) of $p < 0.05$ (using the Wald statistic) as the threshold to check whether the given factor should be included in the

model. Multiple groupings of factors were iteratively tested these are listed in Table 2. To ensure that the factors included do not exhibit multicollinearity, we used the estimated correlation matrix of estimates between the factors. Note that we use the correlation matrix of estimates rather than the covariance matrix of estimates. This is because the range of the values within each factor are relatively large – and vary considerably between each factor – which leads to very small covariance values that are difficult to interpret. The correlation values range between -1 and 1, thus making it easier to interpret the correlation relationship between any two variables. Values of >0.5 were avoided by removing the given variable from the model. The final factor coefficients models represent those variables that produced the best fit while meeting the significance level and multicollinearity criteria.

Logistic Lasso Regression

The Lasso Regression module in Statistica uses a penalized MLE method to fit the linear and logistic regression (LR) models. The regularization is performed over a grid of lambda values, the regularization parameter (the L1 penalty). The Lasso Regression module uses a cyclical coordinate descent to minimize the objective function, the likelihood contribution plus the penalty over a grid of lambda values successively[2].

The main advantage of a LASSO regression model is that it can set the coefficients for factors it does not consider interesting/important to zero. This means that the model decides which factors should and should not be included. Another advantage of a LASSO regression is that the L1 penalty that is added to the model helps to prevent the model from overfitting. One of the main disadvantages of LASSO regression is that the coefficients that are produced by a LASSO model are biased. The L1 penalty that is added to the model artificially shrinks the coefficients closer to zero, or in some cases, all the way to zero. That means that the coefficients from a LASSO model do not represent the true magnitude of the relationship between the factors and the outcome (landslide probability) but rather a shrunken version of that magnitude.

Once the main factors contributing to the landslide occurrence in each GeolCode were identified from the LASSO models, the chosen factors were then included in separate logistic regression models adopting the 'bootstrapping' method to select 100 samples from the training data within each GeolCode (see below for details). The estimated coefficients for each factor included in the models (shown in Table 3) were then used to hindcast landslide occurrence across the main area affected by Cyclone Gabrielle in the North Island of New Zealand (North Island AOI), using the Rain24hr_NIWA (Table 1) as the input along with the other statistically significant factors relevant to each GeolCode-specific LR model.

Bootstrapping

Validation and variation of the regression-model results was done using bootstrapping (Efron and Tibshirani, 1994), which allowed the variability of the parameter values for each factor to be estimated within each GeolCode group. Bootstrapping is a way to assess the robustness of the coefficients estimated for each factor derived from model fitting. For example, a robust estimate of the value for a given factor should have a narrow variation between the minimum and maximum and standard deviation either side of the mean. Conversely, a poor estimate would be one with a wider variation. In addition, a robust model should show little difference between the areas under

[2] Jerome Friedman, Trevor Hastie and Rob Tibshirani. (2008). Regularization Paths for Generalized Linear Models via Coordinate Descent- https://www.jstatsoft.org/article/view/v033i01. Journal of Statistical Software, Vol.33(1), 1-22 Feb 2010.

the ROC curves calculated for each of the 100 bootstrap models. To this end, we repeatedly drew random samples from the respective landslide datasets, where each sample represented 75% of the total GeolCode-specific dataset within the Training AOI. A different sample was drawn 100 times, logistic regression models were fitted to the sampled dataset and the resultant statistics were calculated – minimum, maximum, mean and standard deviation of each variable parameter along with the specificity, sensitivity and area under the receiver operating characteristic curve (ROC:AUC).

Standardisation

To test the importance of each factor used in the model, we converted each one to their standard score using the formula:

$$z = \frac{x-\mu}{\sigma} \qquad \text{Equation. 4}$$

where z is the standard score, x is the data value of the factor, μ is the factors mean and σ is the factors standard deviation. This conversion normalises the distribution of the factor, which in turn allows the model coefficients of each factor to be interpreted as a measure of its importance relative to other factors in the model (Lombardo and Mai 2018). As we standardised each factor prior to use within the regressions, all resulting coefficient values are in the same unit-less scale and can be directly compared. Bootstrapping was applied to the standardised values of each statistically important continuous factor (listed in Table 3), to allow the variance in the given factors coefficient to be quantified and compared to the other factors.

S2.0 Estimating the total number of landslides triggered by the event.

Step 1: The selected factors and their logistic regression model coefficients (from Table 3) – at the optimal grid-cell resolution of 16 by 16 m, which was fixed for hindcasting – were used to hindcast landslide probability across the North Island AOI.

The spatial extent of the hindcast region was the North Island AOI (Supplementary Figure S1). For GeolCode 5 (basement rocks), the 'Greywacke' rock type subunit polygons were removed from the North Island AOI for hindcasting, as very few landslides (about 13) within the Training AOI occurred in these materials[3]. All other factors required as inputs for the best-fitting logistic regression models (adopting the coefficients for these factors listed in Table 3), were created for the entire North Island AOI as per those that were created for the Training AOI (Table 1).

Hindcast landslide probabilities were then calculated for each sample grid cell within the North Island AOI to create the 'the hindcast landslide probability grid'. Areas of water were removed from the analyses. Only GeolCodes 1 to 3 and 5 were used for hindcasting, these account for 85 % of the North Island AOI, excluding areas classified as water.

Step 2: For the North Island AOI, the cumulative sum of the modelled hindcast probabilities, were summed separately for each GeolCode within each of the priority mapping grid cells (2.5 by 5 km^2), to estimate the number of hindcast landslide classified grid cells within each priority mapping grid cell. This was done in two ways:

a) using the full range of hindcast probabilities, and

[3] These materials are mainly greywacke and are in the Wellington region, they are different in both their geological structure and strength, when compared to those included in GeolCode 5 (within the main area affected by the rain from Cyclone Gabrielle), which do contain mapped landslides.

b) by excluding those grid cells with insignificantly low hindcast probability values (i.e., flat ground), lower than the threshold of 0.00167. This threshold was calculated by plotting the hindcast landslide probability distribution (for all GeolCodes together) and selecting only those grid cells with probability values within the upper quartile of the probability distribution – the lower probability value of the upper quartile was termed the 'threshold'. All probability values lower than the threshold were assumed to be 0 as these relate to areas of predominantly flat ground and were excluded from the computation of aggregate statistics[4].

Step 3: The summed probabilities were then converted into the number of landslides per priority mapping grid cell, by multiplying them by each GeolCode specific factor of difference derived from the Training AOI (Table 3). The ratios were calculated using the Training AOI data for each GeolCode separately – as the sizes of landslides vary per GeolCode – by dividing the number of mapped landslides with the number of landslide-classified grid cells.

Step 4: To calculate landslide density, the total number of hindcast landslides estimated within each priority mapping grid cell was then divided by the area of that grid cell for both the Training AOI – using the mapped landslides – and the North Island AOI – using the hindcast model results (Step 3). For the Training AOI, the landslide source area sizes (m^2) were also used to estimate the landslide areal density. The area of each grid cell was calculated by summing the total area covered by each of the GeolCodes, used in the hindcast modelling, excluding areas classified as water. For the North Island AOI, this is the area of the model extent within each priority grid cell, which only includes the GeolCodes used in the model meaning that it excludes GeolCode 4 and other excluded units and water. For the Training AOI this is the area of the priority grid cell with water and obscured areas excluded.

Step 5: To calculate the total area of landslides in the Northern AOI – source-areas only – the hindcast number of landslides were proportioned into the size classes using the mapped landslide source area frequency-to-area relationships from the Training AOI (Table 3).

2.1 LiDAR versus LINZ DEM-derived factors

For the North Island AOI, the same 16 by 16 m sample grid cell resolution was used as per the Training AOI but adopting the 8 m LINZ DEM resampled to 16 by 16 m, for those derivative factors listed in Table 1, including for those sample grid cells within the Training AOI.

A comparison of CumH-R relationships between those generated from factors derived from the LiDAR 1 by 1 m DEM, versus those derived from the LINZ 8 by 8 m DEM (both resampled to 16 x 16 m) was carried out within the Training AOI. This was done by including the factors (listed in Table 3) derived from the different DEMs in LR model training. The results showed that when using the values for those factors included in the LR model training derived from the LiDAR DEM, the number of estimated landslides – sum of the estimated probabilities – per GeolCode grouping are on average 1.2 times higher, compared to those results based on using the 8 by 8 m LINZ DEM. The results adopting the LINZ DEM under forecast the total number of grid cells estimated to be landslide sources, versus the actual number of landslide (source) classified grid cells, by a mean factor of 1.2.

[4] https://earthquake.usgs.gov/data/ground-failure/background.php#pref-ls-model

S3.0 Supplementary Tables

Table S1. Percentage of landslide classified sample grid cells (using the mapped landslide sources only, refer to Table 1 for details) versus non-landslide classified grid cells, at the optimal 16 by 16 m grid cell resolution, per GeolCode and Land Cover Category (categorical factors), within the priority mapping grid cells included in the Training AOI.

Categorical Factor	Percentage coverage of all grid cells within the given factor group, within the Training AOI	Percentage of landslide source-area classified grid cells within the given factor group
GeolCode 1	36.30%	0.17%
GeolCode 2	50.10%	1.39%
GeolCode 3	11.40%	0.38%
GeolCode 4	0.30%	0.04%
GeolCode 5	1.80%	0.20%
Land Cover Category 1	20.20%	0.82%
Land Cover Category 2	3.70%	0.22%
Land Cover Category 3	75.80%	0.83%

Table S2. Number of mapped landslides, the number of landslide-classified 'sample' grid cells (16 by 16 m), and the total number of sample grid cells in the priority mapping grids within the Training AOI, along with the proportions (%) of landslides of a given source-area bin size. The hindcast number of landslide source-area classified sample grid cells within the Training AOI are also given, which were derived from the best fitting regression models (Table 3), and their factors of difference between the hindcast versus mapped number (N) of landslide classified sample grid cells.

1. GeolCode factor	2. Total N grid cells in factor	3. N mapped landslides	4. N mapped Landslide classified grid cells	5. Factor of difference (= 3/4)	6. Proportion (%) of landslides in each source area size bin (m^2)							7. Hindcast N Landslide classified grid cells (source areas)	8. Factor of difference: hindcast versus mapped N landslide classified grid cells (= 7/4)
					2	9	46	225	1,100	5,382	26,340		
1	5,590,258	8,877	9,680	0.92	2.26	19.81	52.35	22.69	2.63	0.22	0.03	11,367	1.17
2	7,728,344	102,174	107,107	0.95	0.92	13.75	54.20	28.55	2.49	0.09	0.01	106,524	0.99
3	1,750,942	4,937	6,580	0.75	1.31	20.06	51.29	22.66	3.87	0.69	0.12	11,723	1.78
5	280,711	493	550	0.90	1.42	16.02	50.71	28.40	3.25	0.20	0.00	393	0.71

S4.0 Supplementary Figures

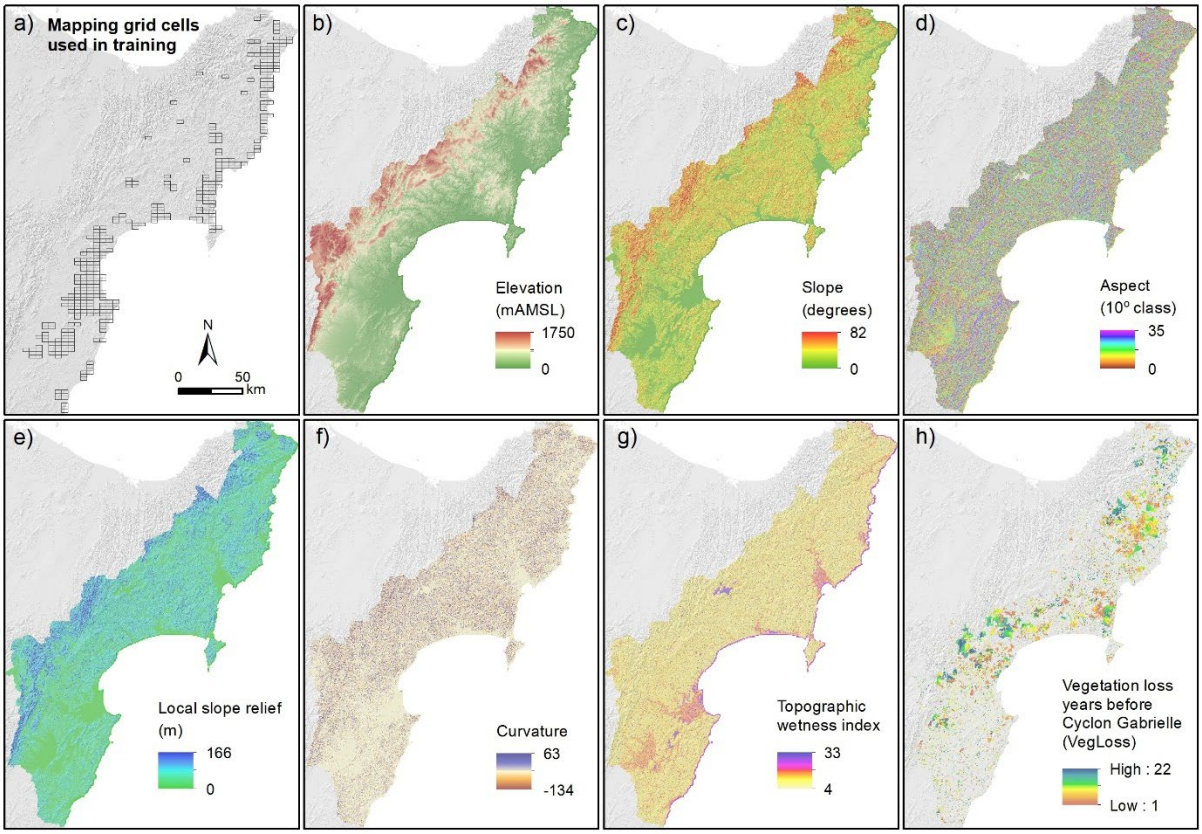

Fig. S1 Priority mapping grid cells used in this study – the Training AOI – and the spatial distribution and range of the continuous factors used in the regressions. a) The Training AOI, comprising the priority mapping grid cells containing the mapped landslides used in this study. b) Elevation. c) Slope. d) Aspect. e) Local slope relief (LSR). f) Curvature. g) Topographic wetness index. h) Vegetation loss years before Cyclone Gabrielle. Refer to Table 1 for details about hose these were generated.

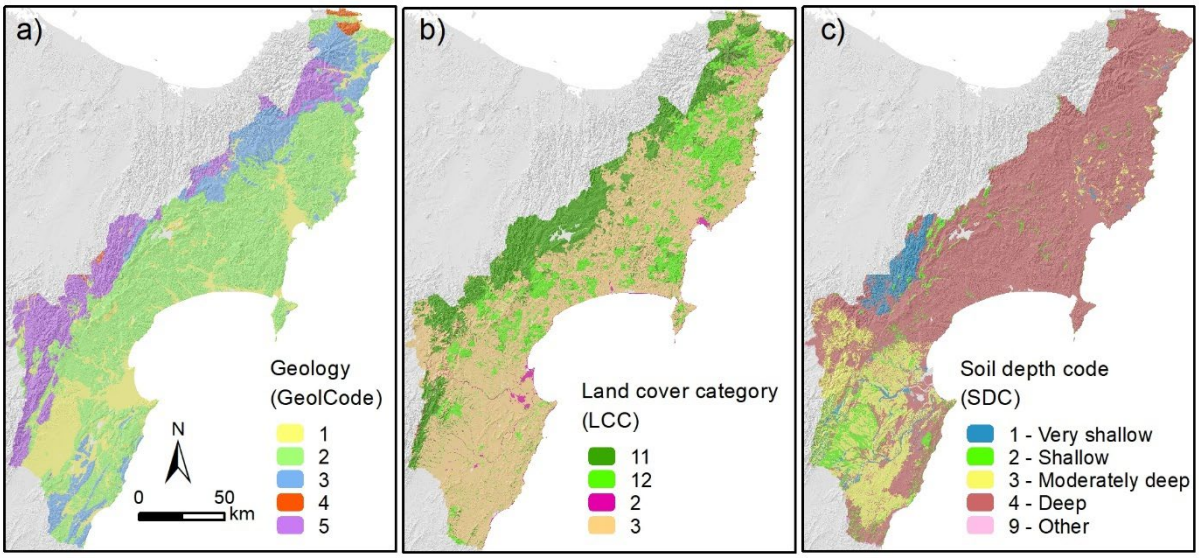

Fig. S2 Spatial distributions and ranges of the categorical factors used in the regressions. a) Geology (GeolCodes). b) Land cover categories (LCC). c) Soil depth codes (SDC). Refer to Table 1 for details about how these were generated.

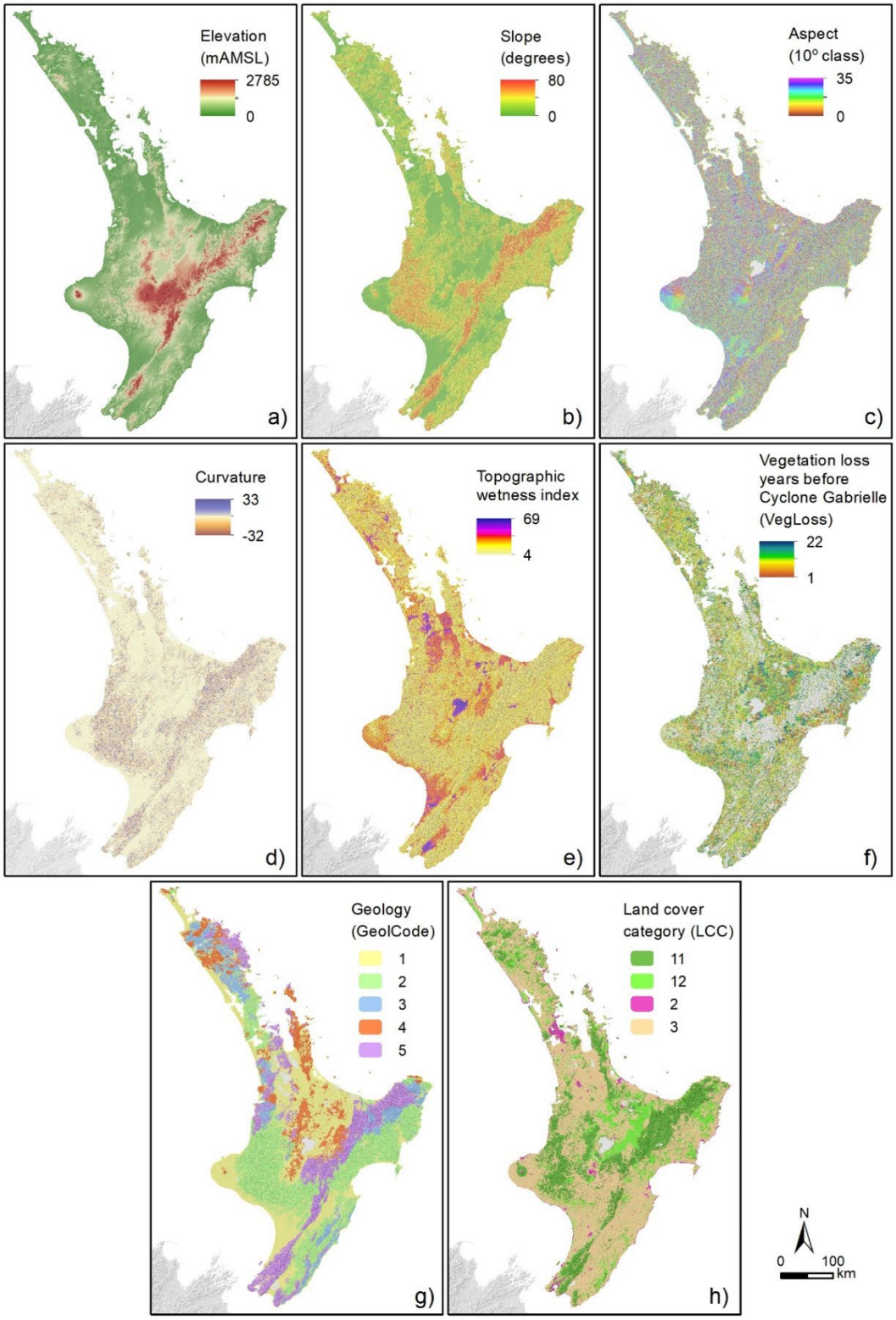

Fig. S3 North Island AOI and modelling parameters used in this study to hindcast landslide probabilities using the best fitting regression models listed in Table 3. a) Elevation. b) Slope. c) Aspect. d) Curvature. e) Topographic wetness index. f) Vegetation loss years before Cyclone Gabrielle (VegLoss). g) Geology (GeolCodes). h) Land Cover Categories (LCC). Refer to Table 1 for details.

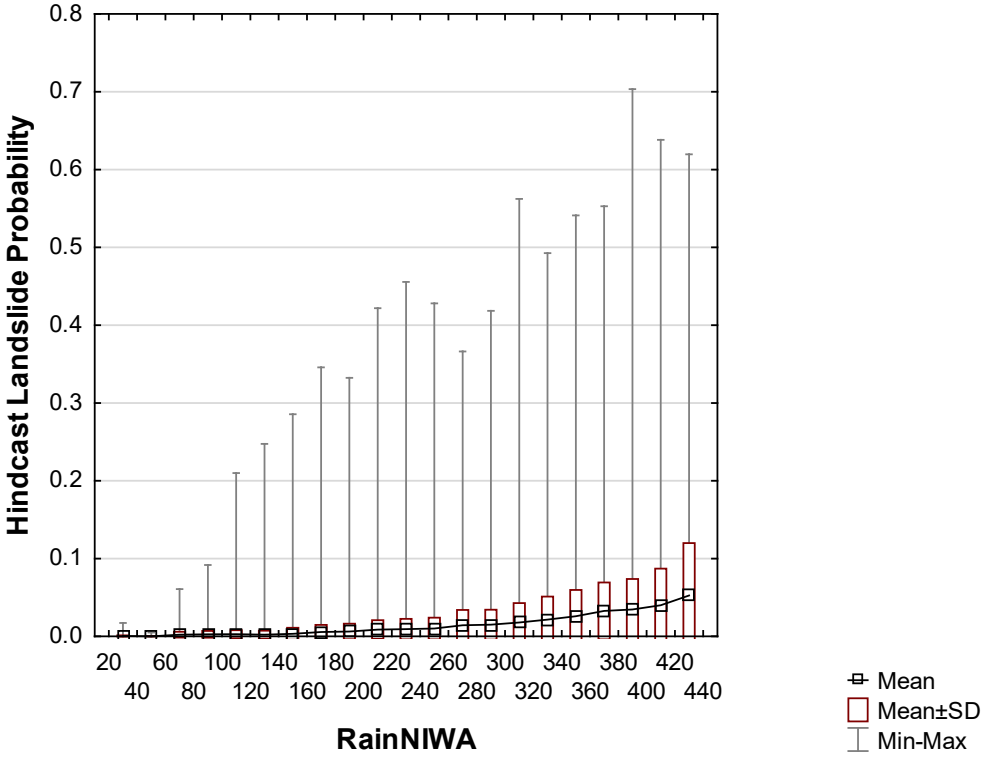

Fig. S4 Hindcast landslide probabilities (0 to 1) calculated using the best fitting regression 'models' (from Table 3) plotted against the 24-hour rain (RainNIWA), for those sample grid cells within the Training AOI – all GeolCodes included. The mean hindcast landslide probabilities are shown and are calculated for each rain bin, adopting 10 mm bins (from low to high rain). The boxes represent the standard deviation of the hindcast landslide probabilities within the given rain bin – noting that most are log-normally distributed – and the whiskers represent the min and max probabilities within each bin.