Landslides (2025) 22:3953–3972
DOI 10.1007/s10346-025-02591-y
Received: 1 April 2025
Accepted: 9 July 2025
Published online: 15 August 2025
© The Author(s) 2025

Chris Massey · Kerry Leith · Tom R Robinson · Biljana Lukovic · Sam McColl · Trevor Carey-Smith · Brenda Rosser · Liam Wotherspoon · Hugh Smith · Harley Betts · Rob Buxton · Janine Bidmead

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

**Abstract** This paper uses an extract of 116,655 landslide source areas from the >160,000 human-mapped landslides contained in the Version One (V1.0) landslide inventory triggered by Cyclone Gabrielle (11–17 February 2023). It describes the analysis of the human-mapped landslide distribution and the relationships between landslide occurrence, their susceptibility factors, and their hydrological forcing factors. The results show that rain is one of the main controlling factors on landslide probability, but the relationship is complex. Our results show the amount of 24-h rain becomes less important, relative to other landslide susceptibility factors, at high (>220 mm) to very high (>330 mm) 24-h rainfalls, suggesting that such extreme rain amounts may not drive the highest probability of landslide occurrence. Results also suggest that the 24-h rainfalls experienced by some areas during Cyclone Gabrielle were likely to have been sufficiently high enough to "overprint" antecedent soil water conditions. In these overprint areas, soil-water conditions prior to Cyclone Gabrielle rainfall may not have materially influenced the spatial landslide distributions. The results from this study suggest that during a storm, more intense rainfall may not directly translate to increasing probability of landslide occurrence; however, it may diminish the effect of pre-event soil water content.

**Keywords** Cyclone Gabrielle · GIS landslide analyses · Landslide impacts · Landslide susceptibility · Rainfall-induced landslides · Soil water content

## Introduction

Tropical Cyclone Gabrielle impacted the North Island of New Zealand between 11 and 17 February 2023, with the most intense rainfall occurring over 12 to 14 February. The maximum 24-h and cumulative 3-day rainfall recorded over New Zealand during the event were 500 and 750 mm, respectively, with heavy rainfall affecting an area >46,000 km², covering about 40% of the North Island (Fig. 1). Cyclone Gabrielle was estimated to be a 100-year ARI event (NIWA 2023). Rainfall recurrence intervals from the event ranged spatially but exceeded 250 years in places (NIWA 2023). This was the most significant rainfall event to affect New Zealand since Cyclone Bola in 1988 (NIWA 2023). Similarly significant cyclones have caused widespread impacts across New Zealand, occurring roughly every 30 to 40 years, including Cyclone Giselle (1968) and the "Cyclone of 1938" (Page et al. 1994). The event was well forecast, and the predicted intensity of the storm resulted in the declaration of a National State of Emergency, only New Zealand's third-ever declaration. Along with surface, coastal, and river flooding, Cyclone Gabrielle triggered hundreds of thousands of landslides over most of the North Island (Fig. 2), making it a multiple-occurrence regional landslide event (MORLE) (Crozier, (2005)). Sadly, 11 people were killed by the cyclone, of which four deaths were directly attributed to landslides. Total direct economic losses (Wilson et al. 2023) from the event are estimated to be more than NZ$14 billion, with direct landslide losses estimated at ~NZ$1.5 billion (Wilson et al. 2023), making it New Zealand's most impactful natural hazard event since the 2010 to 2011 Canterbury Earthquake Sequence (MFAT 2023).

Recent estimates suggest that globally, landslides occur in about 17.1% of landmasses, and that about 8.2% of the global population live in landslide-prone areas where people, properties, and the environment are exposed to potential landslides. Rain is the leading cause of landslides globally (e.g., Nadim et al. 2006; Froud and Petley 2018; Jia et al. 2021) and, given the projected climate changes leading to increased frequency and intensity of storm events (Pörtner et al. 2022), MORLEs and their impacts on people and infrastructure are expected to increase in the coming decades. Rainfall-induced MORLEs are the most common mode of landsliding in New Zealand, occurring somewhere in the country two or three times a year, on average (Crozier 2017). Understanding the factors and conditions that drive MORLEs is therefore critical for our ability to assess and reduce future risk both in New Zealand and globally. However, whether New Zealand MORLEs are similar to those that occur elsewhere in the world, or whether such events are unique to New Zealand, given its high/intense rainfall, steep topography, weak bedrock materials, shallow soils, and widespread deforestation, remains contentious.

This paper uses a sample of 116,655 landslide source areas from the >160,000 human-mapped landslides contained in the Version One (V1.0) landslide inventory for this event (Leith et al. 2025). This sample includes only those landslides that occurred in the Tairāwhiti-Gisborne and Hawke's Bay regions (Fig. 1, the "used for training" priority grid cells). These were two of the main regions affected by landslides triggered by Cyclone Gabrielle; both have similar geological and physiographical settings, and both had pre- and post-event high-resolution aerial photography available for mapping. The paper describes an analysis of the human-mapped landslide distribution and the statistical relationships between landslide occurrence, their susceptibility factors, and their hydrological forcing factors, in this case rain and proxies for soil-water content. The interaction of rain and vegetation (e.g., Gómez et al. 2024) on landslide occurrence is further investigated. The statistical relationships are then used to estimate the total number and extent

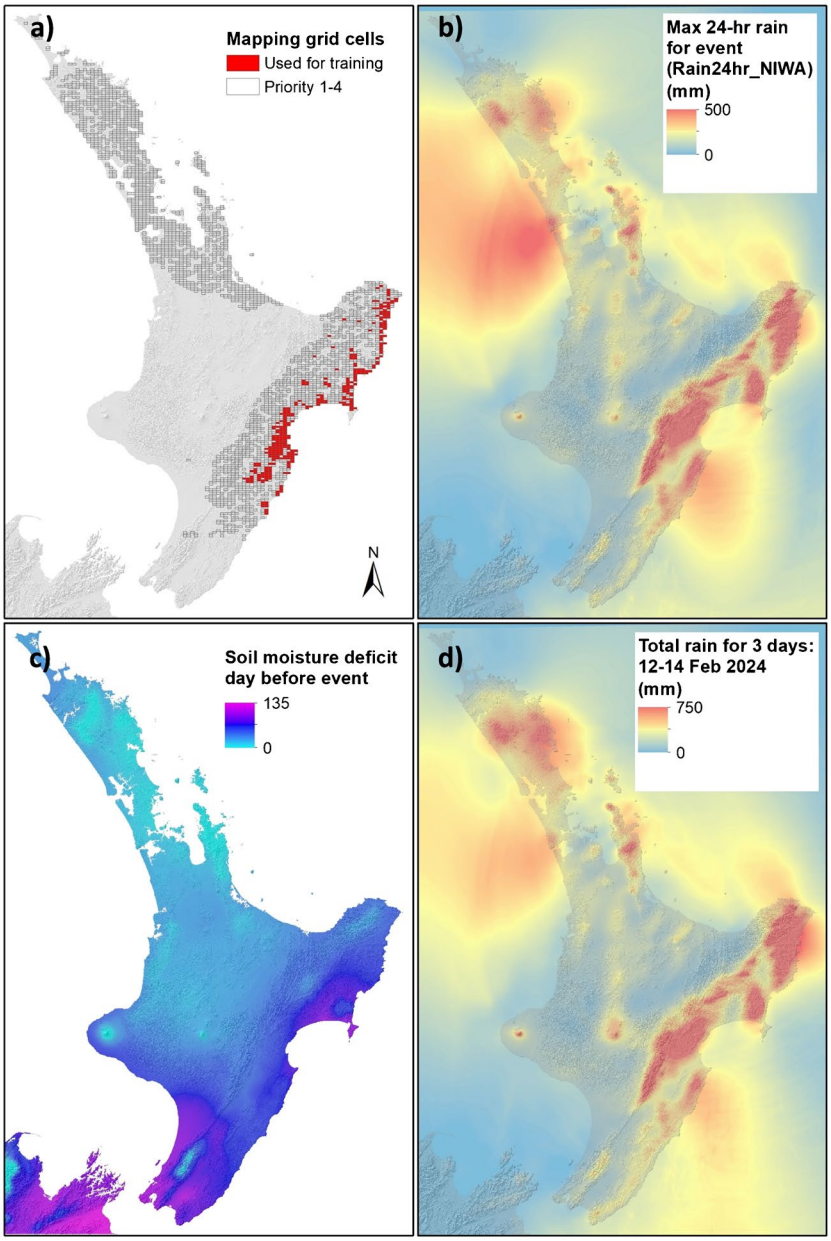



**Fig. 1** **a** Priority mapping grid cells classified as being priority 1–4 by Leith et al. (2025). Those shown as red are the mapped landslides that were used in this study—the Training AOI. **b** Maximum 24-h rainfall for Cyclone Gabrielle, derived from NIWA's Augmented Virtual Climate Station Network (VCSN) model at 500 m resolution. **c** Soil moisture deficit derived from NIWA's VCSN model at 5 km resolution. **d** Total rain (mm) for Cyclone Gabrielle between 12 and 14 February 2024, derived from NIWA's Augmented VCSN model

of landslides triggered by Cyclone Gabrielle in similar geological materials and physiographical settings across the entire North Island AOI. This provides an indication of the total scale in terms of landslide density of the Cyclone Gabrielle MORLE.

## Methods

### Rainfall

The rain data used for investigating and hindcasting landslide occurrence was the maximum 24-h rain amounts in mm derived from an augmented version of NIWA's Virtual Climate Station Network (VCSN) model created for this event at 500 m resolution. This dataset was created following the methodology in Tait et al. 2006), where a spatial spline model is fitted to daily rain gauge observations using a priori information on the spatial distribution of rainfall as an additional independent spline variable. However, the standard VCSN product is known to have biases in high elevation regions (Tait et al. 2012) and sometimes performs poorly for extreme precipitation events (Stone et al. 2022).

For this analysis, the independent variable used in the standard VCSN model (the mean annual rainfall surface) was replaced with an estimate of each day's 24-h rainfall throughout the event derived



**Fig. 2** Selected photographs from field and aerial reconnaissance of landslides triggered by Cyclone Gabrielle in the Tairāwhiti-Gisborne and Hawke's Bay Regions

from the NZ Convective Scale Model (NZCSM). NZCSM is a mesoscale numerical weather prediction system operated at NIWA with a horizontal grid length of 1.5 km and provides forecasts out to 48 h four times per day. This model is a local configuration of the UK Met Office Unified Model (Brown et al. 2012). NZCSM's high resolution means it can model convective rainfall processes explicitly and can adequately resolve complex terrain and its impact on precipitation.

For situations where an event's meteorology was reasonably well predicted by the weather forecasting model, using the predicted rainfall field allows the spline interpolation to incorporate gauge observations with lower error and reduced smoothing. For this event, rain gauge observations from the NZ National Climate Database (http://cliflo.niwa.co.nz) were augmented with regional council operated gauges in the Hawke's Bay and Tairāwhiti-Gisborne Regions, which greatly increased the observation density in these areas, particularly at higher elevations. In addition, hourly rain gauge observations were used to disaggregate the interpolated daily rainfall fields to provide an hourly gridded rainfall product throughout the event. From this product, the maximum 24-h rainfall accumulation was calculated for each grid cell across the region.

## Landslide mapping

Landslide sources mapped within the Tairāwhiti-Gisborne and Hawke's Bay regions were extracted from V1.0 of the landslide inventory (Leith et al. 2025). Of the >160,000 landslide sources contained within the V1.0 inventory, 116,655 were within the study area ("Used for training" cells in Fig. 1). Mapping for the V1.0 inventory was undertaken using 5 × 2.5 km "priority mapping grid cells" that

covered the affected area of the North Island—a total of 12,219 cells. Grid cells were assigned a priority level (1–4) based on population density, the presence of critical infrastructure, scientific interest, and to capture a representative spread of the landslides triggered by the event, including those with few and no mapped landslides. The extract from the V1.0 inventory includes only landslides mapped within a representative number (337) of these grid cells, which comprise the "Training" area of interest (AOI) ("Used for training" cells in Fig. 1). Mapping was carried out at a nominal scale of 1:1000 using high-resolution 0.2–0.3 m ground resolution, pre- and post-event orthorectified aerial imagery and LiDAR-derived shade models. For more details relating to the mapping methods refer to Leith et al. (2025). Field validation of the mapped landslides was carried out by the authors within the Training AOI immediately after Cyclone Gabrielle (Leith et al. 2025), and during field visits made to the affected areas since the event (Fig. 2).

## Landslide analyses

The objectives of the analyses of the landslides extracted from the V1.0 landslide inventory were to (1) identify the main "regional" factors controlling landslide occurrence using the mapped landslide distribution and (2) estimate the total magnitude—in terms of landslide density and number—of landslides triggered in similar materials and physiographical settings, across the North Island. A summary of the methodologies used to explore these objectives is provided here, and a more detailed description of the methods is contained in the Supplementary Section.

### Landslide-initiating controlling factors

The landslide inventory within the Training AOI was used to explore the relationships between the occurrence of a landslide and the factors that may control its occurrence. These factors represent those that are typically used in similar studies (e.g., Buxton 2025), and they were chosen to capture/represent "proxies" for the observations made by the authors in the field. These controlling factors were broadly grouped into (1) landslide susceptibility factors that capture the strength of the hillslope materials at a regional scale and the static shear stresses at the slope scale and (2) landslide forcing factors representing the intensity of the event-specific rain, land cover, soil water content, and their proxies. Table 1 describes the factors used in the analysis, their codes, description of how they were generated, their resolutions, and units. Supplementary Figures S1 and S2 show the spatial distribution of these factors within the Tairāwhiti-Gisborne and Hawke's Bay regions and the Training AOI.

The analysis was based on the frequency-ratio (e.g., Kritikos et al. 2015) and logistic regression (e.g., Lombardo and Mei 2018) methods. The frequency-ratio method was used to identify and describe any relationships between the selected factors, relative to their spatial proportion, and the landslides mapped within the Training AOI. Logistic LASSO (L1) regression was used to then select the main factors that could be combined to explain landslide occurrence—adopting the metric of landslide probability ranging from 0 to 1 (0 to 100%)—whilst reducing overfitting. This step selected the minimum number and therefore most important factors that could be used to explain landslide occurrence. Once the main factors were identified, more detailed analysis of the influence of rain and vegetation on landslide occurrence was carried out using a "bootstrap" approach (Efron and Tibshirani 1994) adopting standardized values for the factors used in the regressions. These are described in detail in the Supplementary Section 1.0.

Analyses of the landslides within the Training AOI used "sample" grid cells at multiple resolutions, including 3 by 3 m, 8 by 8 m, and 16 by 16 m. The training utilized different sample grid cell resolutions for the input factors (listed in Table 1) to see if there would be any statistical gain in using higher versus lower resolution datasets. Training within the AOI adopted a LiDAR-derived 1 m DEM, which was reprocessed to the given resolutions and used to generate the derivative factors listed in Table 1.

### Regional landslide hindcasting

The selected landslide susceptibility factors and their regression "model" coefficients, from the best-fitting logistic regression (LR) analyses and optimal "sample" grid-cell resolution, were used to hindcast landslide susceptibility across the North Island AOI. The landslide susceptibility metric calculated by the LR models—and used to estimate the number of landslides likely to have been triggered by the event—was the probability (0 to 1) of a grid cell being a landslide source area. This hindcast was used to estimate the total number of landslide-classified source-area grid cells which were translated into the estimated number of landslide sources triggered by the event, in similar geological materials and physiographical settings (Supplementary Figure S3). To do so, the susceptibility factor datasets were created for the entire North Island AOI, as per those that were created for the Training AOI (see Supplementary Section 2.0). Hindcast landslide probabilities were then calculated for each sample grid cell within the North Island AOI to create the "hindcast landslide probability grid." The North Island AOI used for hindcasting includes all the priority mapping grid cells (rated as priority 1–4) as well as those other grid cells (rated as priority 0) covering the entire North Island of New Zealand. Within these grid cells, only those sample grid cells comprised predominantly of geological units that were not represented within the Training AOI were excluded from the analyses. For hindcasting, the LINZ 8 by 8 m DEM was used, as no contiguous LiDAR coverage was available for the North Island AOI. The impact of using the 8 by 8 m LINZ DEM in the Training AOI versus the LiDAR-derived DEM on model training was explored for those sample grid cells within the Training AOI adopting a 16 by 16 m grid cell resolution. This was done by generating the DEM-derived factors listed in Table 1 from the 8 by 8 m LINZ DEM, which were then used in the LR analyses. These results were compared to those from the LR models adopting the same factors derived from the resampled LiDAR DEM. The results from these analyses are presented in Supplementary Section 2.1.

Using the hindcast landslide probability grid, the total number of sample grid cells hindcast to contain, or represent part of, a landslide source area was estimated for those sample grid cells within the Training AOI. This was done by cumulatively summing the hindcast landslide probability values—sorted from low to high probability—of all sample grid cells. It was calculated initially for all sample grid cells, and for those sample grid cells within each separate geology grouping, adopting the "GeolCode" groups described in Table 1. These were then compared to the corresponding record of mapped landslides. These

**Table 1** Data (factors) used in the analysis, their codes, description of how they were generated, their resolutions and units

| Input Factor/Data/Code | Original data source | Details, processing and resampling | Type/ Units<br>Resolution(s).<br>Those in bold were used in final models |
|---|---|---|---|
| Landslides | Landslide source area points initiated by Cyclone Gabrielle (Leith et al. 2025) | Landslide source area points within the priority mapping grid cells (Training area of interest [AOI]) were extracted from the database as at 26/10/2023. Landslide source-area points were converted to circles with an area equivalent to the captured source area attribute (e.g., fridge, house sizes etc). These were then used together with the source-area points to create a 'sample' grid (raster) of landslide versus non-landslide grid cells, adopting grid cell resolutions of 8 by 8 m, 16 by 16 m and 32 by 32 m. A given sample grid cell was classified as being a landslide (= 1) if the centroid of that grid cell fell within a source-area circle or if the source-area point fell within the grid cell – This is to make sure that all landslide sources are captured including the very small ones. Sample grid cells that were classified as containing a landslide source (tagged with 1), could include many landslides if e.g., several small landslides fall within one sample grid cell. The number of landslides within a sample grid cell has not been captured, such grid cells are all tagged with 1.<br>Only the mapping priority grids cells (and related landslides) that had been completed were selected. Mapping priority grid cells that were mostly covered with obscured areas (due to cloud cover in the imagery) were removed, and those priority mapping grid cells containing no landslides were kept – if we were confident that mapping of the grid cell had been completed and no landslides identified. | 1 = landslide,<br>0 = no landslide<br>Original landslides were points, which were converted to grids as described.<br>Grids of multiple resolutions were used for sensitivity analyses, these were: 8, 16 and 32 m grid cells. |
| Rainfall (Rain24hr_NIWA) | Grid of 24-hr rain amounts in mm derived from NIWA's Augmented Virtual Climate Station Network (VCSN[a]) model at 500 m resolution. | The rain represents the max 24-hr rain in the period from 9 am on 12/02/2023 to 9 am on 15/02/2023. Note the storm lasted several days and so the max 24-hr rain may occur on different days at different locations. The rain values were attributed to the sample grid cell by taking the original rain value at its centroid (original 500 m resolution) using bilinear interpolation. | Continuous, mm/day<br>8, **16** and 32 m (initial data was at 1.5 km resolution). |
| Soil moisture deficit (SMD) | Grid of soil moisture (soil moisture indices – SMI) for the day before the storm, taken from NIWA's Virtual Climate Station Network (VCSN) model at 5 km resolution. | The SMD values were attributed to the sample grid cell by taking the Soil Moisture Indices (SMI) value at its centroid using bilinear interpolation. SMI is calculated in mm from 9 am to 9 am (24-hr period), from rainfall and evapotranspiration. The base value is -150mm ("permanent wilting point") based on "soil store capacity". A value of "0" indicates the soil is at "field capacity" (amount of water held in the soil after the excess has drained away). A value greater than "0" indicates runoff. All positive values were set to 0. All negative values were changed to their absolute value to create the soil moisture deficit grid (SMD).<br>Note on the SMI calculation model used to create the SMD grid used in this work: The SMI model has one parameter, the soil store capacity (Smax). This model assumes that rainfall infiltrates into a soil store from which moisture is depleted by evapotranspiration loss. The model keeps track of the soil store for each timestep. When the soil store is full, it is assumed to be saturated, and the excess rainfall becomes surface runoff. Evapotranspiration is assumed to continue at its potential rate until the soil water store drops to half capacity. Evapotranspiration then decreases linearly to zero when the store is empty. For the simulations here, an Smax value of 150 mm is used for all the modelling grids.<br>https://niwa.co.nz/nz-drought-indicator-products-and-information/drought-indicator-maps/soil-moisture-deficit-smd | Continuous, mm<br>3, 8 and 16 m (initial data was at 5 km resolution) |

**Table 1** (continued)

| Input Factor/Data/Code | Original data source | Details, processing and resampling | Type/ Units<br>Resolution(s).<br>Those in bold were used in final models |
|---|---|---|---|
| Geology (GeolCodes) | The geology classification is based on 'dominant unit' determined for each geological unit within the NZ QMAP 2018 dataset. Mapping resolution of 1:250,000. | The QMAP geological units were mapped as polygons. The geological units covering the North Island AOI (including the Training AOI) were classified into seven groups based on the QMAP 'dominant unit' field, and a number code (1 to 7) was assigned to each unit and stored in the GeolCode field. This number code was used to generate a grid out of the polygons using the MAXIMUM_COMBINED_ AREA option in ArcGIS. A total of 5 GeolCodes (1-5) cover the priority mapping grid cells within the Training AOI, and an additional two GeolCodes (6 and 7) exist in the hindcast North Island AOI, which were ignored as it was not possible to train models for these GeolCodes, given that they don't exist in the Training AOI.<br>GeolCode1: Quaternary sands, silts and gravels. These materials typically form terrace deposits on the top of the steep coastal cliffs, as well as on inland slopes adjacent to the main rivers of the area. Many of these terraces have been incised by rivers.<br>GeolCode2: Neogene limestones, sandstones and siltstones, which are typically weak. These represent the largest (in land area) geology group within both the priority mapping grid cells and the wider area affected represented by the 'hindcast' area.<br>GeolCode3: Upper Cretaceous to Paleogene rocks, including limestones, sandstones, siltstones and minor volcanic rocks. These are typically weak (like the Neogene limestones and sandstones) and easily erodible, and they can contain thin clay seams, which are volcanic in origin.<br>GeolCode4: All intrusive and extrusive igneous rocks, ranging in age but mainly Jurassic. Includes andesites, basalts, rhyolites, ignimbrites, dacites, granites and granodiorites. These represent the smallest (in land area) geology group within the area affected.<br>GeolCode5: Lower Cretaceous Torlesse (Pahau terrane) 'basement' rocks are predominantly sandstones and argillite.<br>GeolCode6: Quaternary lava. Predominantly basalt lavas.<br>GeolCode7: Neogene debris, including landslide debris, alluvium etc. | Categorical<br>3, 8 and **16** m. (initial data was at 1:250,000 scale resolution) |
| Land cover categories (LCC) | LCDB v5.0 was obtained from the LRIS portal[b]. LCDB v5.0 was mapped in the summer of 2018/2019. Mapping resolution of 1:50,000. | The data were grouped based on the land cover description, a code was assigned to each group and the polygons were turned into integer grids using the MAXIMUM_ COMBINED_AREA option in ArcGIS.<br>For this work the land cover grid LC2018 was used as it represents the land cover types at the time closest to when the landslides occurred. To simplify the training, the LC grid codes were initially combined into three categories, with one subsequently split into two, see below.<br>LCC1 (Forest): Group 1 includes: indigenous forest, deciduous hardwoods, exotic forest and harvested forest. This was further split into:<br>• LCC11 indigenous forest and<br>• LCC12 exotic forest comprising deciduous hardwoods, Pines, and harvested forest.<br>LCC2 (Bare ground and Urban): Group 2 includes: bare ground, urban, urban open (parks and open space), transport and landslides.<br>LCC3 (Scrub and grassland): Group 3 includes: alpine vegetation, scrub, grassland and horticulture.<br>Water and swamp land cover areas were excluded. | Categorical<br>3, 8 and **16** m |

**Table 1** (continued)

| Input Factor/Data/Code | Original data source | Details, processing and resampling | Type/ Units<br>Resolution(s).<br>Those in bold were used in final models |
|---|---|---|---|
| Soil Depth Code (SDC) | Soil depth classes taken from SMap. S-map is a digital soil map for Aotearoa New Zealand | Soil depth classes taken from SMap[c] polygons where available, and any gaps were infilled with the Fundamental Soils Layer (FSL) data[d]. Unfortunately, SMAP and FSL data are not available for all the area affected by Cyclone Gabrielle and so are included in the training data to investigate any potential linkages between soil depth and landsliding in each of the GeolCodes. They were not used for forecasting.<br>Class 1 – Very shallow: 0–0.44 m<br>Class 2 – Shallow: 0.45–0.59 m<br>Class 3 – Moderately deep: 0.6–0.89 m<br>Class 4 – Deep: > 0.9 m | Categorical<br>Mapped as polygons at 1:50,000 and resampled to 3, 8 and **16** m |
| Digital elevation model (DEM) | LiDAR derived DEMs provided by Land Information New Zealand (LINZ). 1-m resolution. | The original 1 m DEMs for Tairāwhiti and Hawke's Bay supplied by LINZ were mosaiced together and resampled to multiple resolutions (3, 8 and 16 m) to test the impact of varying the DEM resolution on derived topographical variables used to investigate landslide occurrence. The Elevation, Slope, Aspect, Local Slope Relief, Curvature and Topographic Wetness Indices were calculated from these DEMs as described below. | Continuous, m<br>3, 8 and **16** m |
| Elevation | Grid cell elevation resampled from the 1 m LiDAR DEM. | Elevation calculated in ArcGIS from the LINZ 1 m LiDAR DEM, resampled to 3, 8 and 16 m resolutions. This variable is a proxy used to represent different climatic and vegetation conditions and erosion mechanisms, which could dominate at different elevations. | Continuous, m AMSL<br>3, 8 and **16** m |
| Slope | Grid cell slope angle calculated from the resampled 1 m LiDAR DEM. | Hillslope gradient calculated in ArcGIS at 3, 8 and 16 m resolutions using resampled LiDAR DEMs. This variable is a proxy for the static shear stresses in the slope. | Continuous, degrees<br>3, 8 and **16** m |
| Aspect | Grid cell aspect calculated from the 1 m LiDAR DEM. | Hillslope aspect calculated in ArcGIS at 3, 8 and 16 m resolutions using the resampled 1 m LiDAR DEM. Used in the Frequency-ratio analyses. The values were grouped into 10 degrees classes. | Coded value representing a 10-degree aspect class, with values 0–35.<br>3, 8 and **16** m |
| AspectSin | Sine of the aspect calculated from the 1 m LiDAR DEM. | Sine of the hillslope aspect calculated in ArcGIS at 16 m resolution using the resampled 1 m LiDAR DEM. Used in the logistic regression models. | Continuous, values between –1 (West) and 1 (East).<br>16 m |
| AspectCos | Cosine of the aspect calculated from the 1 m LiDAR DEM. | Cosine of the hillslope aspect calculated in ArcGIS at 16 m resolution using the resampled 1 m LiDAR DEM. Used in the logistic regression models. | Continuous, values between –1 (South) and 1 (North).<br>16 m |
| Local Slope Relief (LSR) | Local height of a slope within a defined search radius, calculated from the LINZ 1 m DEM. | Local slope relief generated at 3, 8 and 16 m resolutions using resampled LiDAR DEMs, using focal statistics in ArcGIS. It is calculated as the difference in elevation between a given grid cell and the lowest in elevation grid cell within an 30 m radius.<br>LSR represents the local height of the sample grid cell. This variable represents a proxy for slopes that could generate larger landslides due to their 'larger-scale steepness' (larger than just a sample grid-cell size), where larger values of LSR represent the steeper and higher slopes of the region. | Continuous, m<br>3, 8 and **16** m |
| Curvature | Grid cell curvature calculated from the LINZ 1 m DEM | Standard Curvature calculated in ArcGIS at 3, 8 and 16 m resolutions using resampled LiDAR DEMs. For each cell, the curvature value is calculated as a second derivative of the surface fitted through that cell and its eight surrounding neighbours. A negative value indicates that the surface is upwardly convex at that cell. A positive profile indicates that the surface is upwardly concave at that cell. A value of 0 indicates that the surface is flat. This variable is a proxy for slope 'hollows' that represents topographic depressions, where surface water might accumulate. | Continuous One hundredth (1/100) of az-unit<br>3, 8 and **16** m |

**Table 1** (continued)

| Input Factor/Data/Code | Original data source | Details, processing and resampling | Type/ Units<br>Resolution(s).<br>Those in bold were used in final models |
|---|---|---|---|
| Topographic Wetness Index (TWI) | Grid cell value calculated from the LINZ 1 m DEM | The TWI quantifies the topographic control on hydrological processes. The TWI is a terrain attribute that estimates where water will accumulate in an area based on elevation differences. It's a function of slope and the upstream contributing area. It was calculated using the resampled DEMs and several ArcGIS geoprocessing tools[e].<br>Low TWI values indicate 'dryer' regions where water is less likely to accumulate, compared to higher values indicating 'wetter' regions. | Continuous, −10 to 10<br>3, 8 and **16** m |
| Vegetation lost years before Cyclone Gabrielle (VegLoss) | Taken from the University of Maryland (USA) Global Forest change between 2000 and 2022. | The years before 2023 since the vegetation (forest) was lost. These are the results from a time-series analysis of Landsat images in characterizing global forest extent and change from 2000 through 2022 (assumed to be Dec 2022)[f].<br>The original resolution of the dataset is 30 m. Trees are defined as vegetation taller than 5 m in height and are expressed as a percentage per output grid cell as '2000 Percent Tree Cover'. 'Forest Cover Loss' is defined as a 'stand-replacement disturbance', or a change from a forest to non-forest state, during the period 2000–2022. 'Forest Cover Gain' is defined as the inverse of loss, or a non-forest to forest change entirely within the period 2000–2022. 'Forest Loss Year' is a disaggregation of total 'Forest Loss' to annual time scales.<br>Each cell in the original grid represents the year in which the forest (vegetation) was lost or gained. The derived grid shows the number of years before 2023 when this change happened. All the NULL values – meaning no forest loss within the time period – were changed to 30 years, as this is the average time between planting and harvesting of most commercial forest (pine) in NZ. This was done for all areas of forest including indigenous and exotic forestry types – noting that it mainly relates to exotic forest. | Continuous. The N years before 2023 that vegetation was lost in a grid cell<br>30 m |

*Note*: The resolutions shown in bold were those used for hindcasting in the North Island area of interest (AOI) once the models had been trained on data from the Training AOI. The other resolutions were used for sensitivity analyses

[a]https://niwa.co.nz/climate/our-services/virtual-climate-stations

[b]https://lris.scinfo.org.nz/data/

[c]https://smap.landcareresearch.co.nz/maps-and-tools/app

[d]https://lris.scinfo.org.nz/data/

[e]https://mapscaping.com/topographic-wetness-index-in-arcgis-pro/

[f]Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science 342 (15 November): 850-53. Data available on-line from: https://glad.earthengine.app/view/global-forest-change. https://www.science.org/doi/10.1126/science.1244693

cumulative hindcast versus record (CumH-R) relationships are a way to show whether the hindcast models are under- or over-predicting the number of grid cells that are hindcast as being a possible landslide source. They therefore show the "consistency" of the modelled hindcast probabilities against the actual grid cells classified as being landslides from the mapped landslide source areas (e.g., Parker et al. 2015; Massey et al. 2018; 2020; 2022). The corresponding recorded landslide values (i.e., 1 = landslide, or 0 = no landslide based on the mapped landslides) were also summed on the sorted dataset to calculate the cumulative sum of grid cells classified as being a landslide source. A well-performing model should plot as a "straight" line and close to the line of equality in cumulative hindcast versus cumulative record space. A poorly performing model will not be a straight line and will either over- or under-hindcast the cumulative number of landslide cells.

The best fitting LR models were used to hindcast the landslide probability and the total number and density (N landslides/km$^2$) of landslides triggered by Cyclone Gabrielle across the North Island AOI. This was done by using the max 24-hourly augmented VCSN rainfall for the event (Table 1, and Supplementary Section S2). The number of hindcast landslides in each GeolCode was estimated by multiplying the summed probabilities by the "factor of difference" between the number of mapped landslides and landslide-classified grid cells, per GeolCode, derived from the mapped landslides within the Training AOI.

## Results

### Landslide inventory

The landslides extracted from the V1.0 landslide inventory used in this study were predominantly landslides that initiated as slides in soil or debris; they were small (<46 m$^2$) in area and were predominantly first-time failures (Leith et al. 2025). Such landslides were similar to those mapped outside of the Training AOI (Leith et al. 2025). The largest percentage of landslide-classified grid cells within the Training AOI occurred in GeolCode 2, which also covered the largest proportion of the sample grid cells within the Training AOI (Supplementary Table S1). Supplementary Table S2 shows the proportions of landslide versus non-landslide sample grid cells and their source area sizes, per GeolCode category, within the priority mapping grid cells included in the Training AOI. These results show that the ratio of mapped landslide source areas to landslide-classified grid cells within the Training AOI varies per GeolCode as a function of source area size.

The landslide source number and area densities, calculated using the human-mapped landslide sources inside the priority mapping grid cells within the Training AOI, are log normally distributed, ranging from 0 to 327 landslides/km$^2$, with a mean of 28 and a standard deviation of 46 landslides/km$^2$, and a skew of 2.7. The aerial density ranges from 0 to 3.8%, with a mean of 0.4% and a standard deviation of 0.6% with a skew of 2.6. The priority mapping grid cells with the highest densities are those mainly comprising GeolCode 2, with scrub and grassland cover (LCC3).

### Field observations

Field observations carried out by the authors, in the Tairāwhiti-Gisborne and Hawke's Bay regions post-Cyclone Gabrielle, confirmed that most of the landslides—observed in the field—were relatively shallow, less than a meter deep, and most were a combination of slides and flows. Within their source areas, the rupture surfaces were coincident with permeability boundaries, mainly between soil/completely weathered in situ rock overlying highly weathered rock—adopting NZGS (2005) descriptions. Observations of the soils suggested that most occurred in residual soils, although localized clusters of landslides were observed to have occurred in soils derived from volcanic ash and related deposits. Most landslides occurred on slopes where the land cover was grassland and or had been recently (within the past few years) deforested.

### Frequency-ratio

Graphs showing the results of the frequency-ratio analyses are shown in Fig. 3. A summary of the results, adopting the optimal 16 by 16 m grid cell resolution within the Training AOI, indicates that landslide-classified sample grid cells occurred proportionally on slopes that had the following characteristics (with the largest categories underlined):

- Geological groupings (GeolCodes; Table 1): 7.8% = Geolcode 1 (Quaternary soils), *86.4% = GeolCode 2 (Neogene sedimentary rocks)*, 5.3% = GeolCode 3 (Upper Cretaceous to Paleogene sedimentary rocks), 0.05% = GeolCode 4 (Intrusive igneous rocks) and 0.45% = GeolCode 5 (Lower Cretaceous Torlesse sedimentary rocks).
- Land cover categories (LCC; Table 1): 20.7% = LCC 1 (forest), 1.0% = LCC 2 (bare ground and urban) and *78.3 % = LCC 3 (scrub and grassland)*.
- Soil depth classes (SMap; Table 1): 68% = *deep (>0.9 m) soils*, 29% = moderately deep (0.6–0.9 m) soils, and 3% = shallow and very shallow (<0.6 m) soils.
- Forest loss (Global Forest change; Table 1): 51% of landslides in forested/recently forested areas had lost tree cover 3–5 years before Cyclone Gabrielle.
- 24-h rainfall, soil moisture (SMD), and topographic wetness indices (TWI) (Table 1): Of the forcing factors tested, landslides preferentially occurred in areas of high 24-h rainfall where the hindcast soil moisture deficits were between 45 and 75 mm, indicating partially saturated conditions, and in areas of higher TWI, indicative of "wetter" regions.

These results show that landslides preferentially occurred in soils derived from Neogene sedimentary rocks (GeolCode 2), which are at/near surface across most of the Training AOI. Most preferentially occurred in scrub and grassland (LCC 3). Of those that occurred in forest (LCC 1), areas that had lost tree cover in the preceding 3 to5 years before Cyclone Gabrielle saw the most preferential rates of landsliding.

### Logistic regression

For the sample grid cells within the priority mapping grid cells included in the Training AOI, initial model testing and results (Models 1 to 3, Table 2) found no statistical gain in model performance—ROC:AUC and pseudo R$^2$—when adopting the 3 or

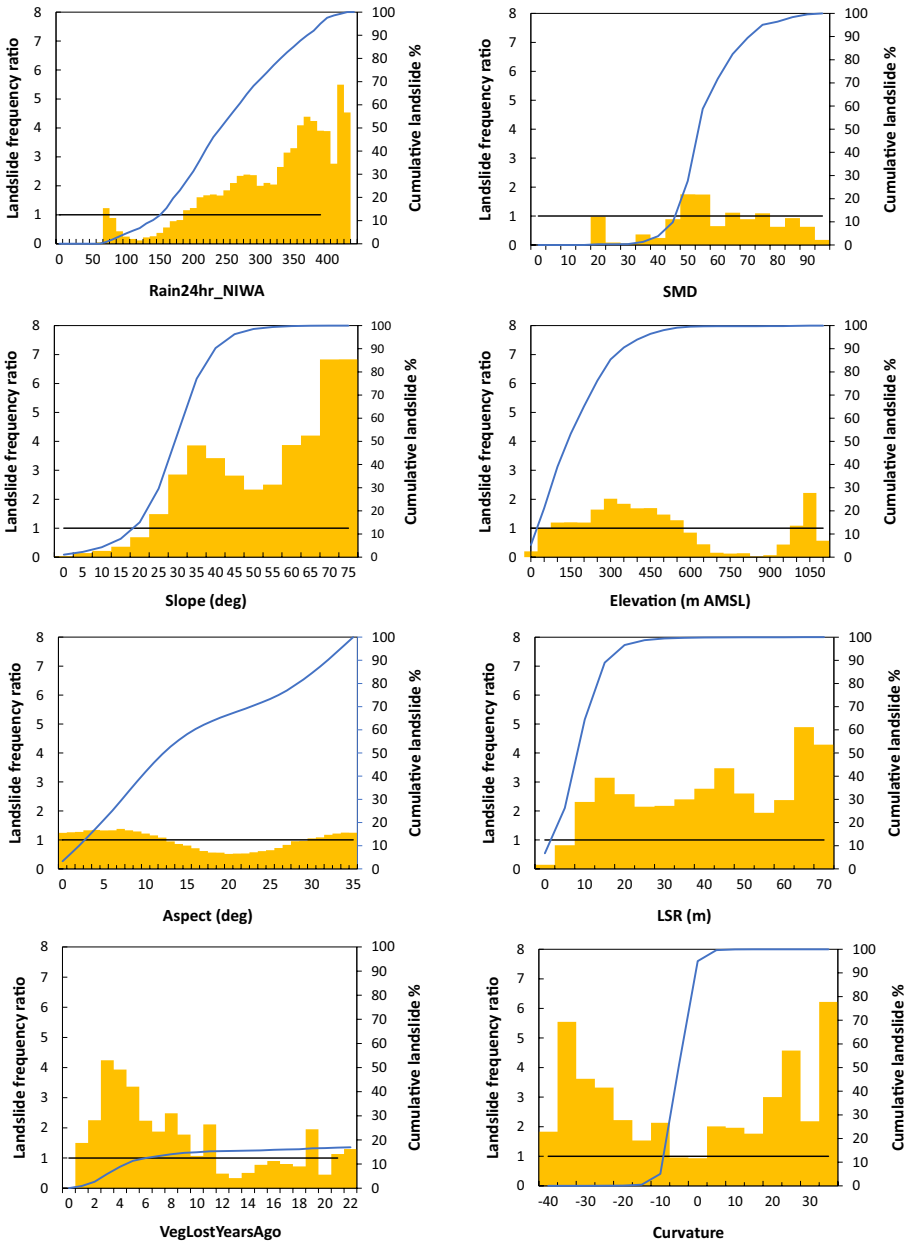



**Fig. 3** Landslide (sources only) frequency ratios (histograms) for selected landslide "forcing" and "susceptibility" factors. The dark blue lines show the cumulative % of the grid cells classified as being a landslide. The black line represents the frequency ratio = 1.0; frequency ratios >1.0 are thought to represent statistically important values for the given factor. The total number of grid cells classified as landslide = 123,943; and non-landslide = 15,295,850, within the Training AOI. Note the "VegLostYearsAgo" (VegLoss) factor plot is only for those grid cells with Forest land cover (LCC 1)

8 m resolution LiDAR DEM and derivative factors of elevation, slope, aspect, LSR, curvature, and TWI, relative to the 16 m LiDAR DEM. In fact, the performance decreased when going from 8 to 3 m, possibly because the finer resolution data does not represent the size of most of the mapped landslide source areas, or the positional accuracy of the mapping. For the results presented from here on, unless stated differently, the sample grid cell resolution used for training and testing within the Training AOI was 16 by 16 m, adopting the LiDAR DEM for those derivative factors listed in Table 1.

Initial regression "models" were run using all factors combined to gauge the relative importance of each one in the regressions (Models 1 to 3). Based on these results, different factors were then grouped using the categorical factors of land cover types (LCC, Models 4 to 6, Table 2) and geological materials (GeolCodes, Models 7, 9, 11, 13, and 14, Table 2), and regressions carried out on the factors contained within each grouping. Results showed that the performance of the models adopting grouped factors performed better than those containing the categorial factors of GeolCode and LCC together. The GeolCode-grouped

**Table 2** Summary of the logistic regression hindcast "models" investigated, and the continuous and categorical factors included in each one, along with their performance adopting the pseudo $R^2$, and ROC:AUC values

| Model | Combining factor | Description | Continuous factors | | | | | | | | | Categorical factors | | | | | | | | | | | | Pseudo $R^2$ | Mean ROC:AUC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 24 hr Rain | SMD | TWI | VegLost | Elevation | Slope | Aspect | LSR | Curve | GeolCode 1 | GeolCode 2 | GeolCode 3 | GeolCode 4 | GeolCode 5 | LCC 1 | LCC 2 | LCC3 | LCC 11 | LCC 12 | SDC 1 | SDC 2 | SDC 3 | SDC 4 | | |
| 1 | None | All factors used together | O | O | O | O | O | O | O | Y | O | O | O | O | O | O | O | O | O* | | | | | | | 0.189 | 0.879 |
| 2 | None | All factors used together, no GeolCodes | O | O | O | O | O | O | O | Y | O | | | | | | O | O | O* | | | O | O | O | O* | 0.193 | 0.877 |
| 3 | None | All factors used together, no SDC | O | O | O | O | O | O | O | Y | O | O | O | O | O | O* | O | O | O* | | | | | | | 0.189 | 0.878 |
| 4 | LCC 1 | GeolCodes incl. as categories | O | O | O | O | O | O | O | Y | O | O | O | O | X | O* | | | | | | | | | | 0.201 | 0.871 |
| 5 | LCC 2 | GeolCodes incl. as categories | O | O | O | X | O | O | O | X | X | O | O | O | X | O* | | | | | | | | | | 0.280 | 0.952 |
| 6 | LCC 3 | GeolCodes incl. as categories | O | O | O | O | O | O | O | Y | O | O | O | O | O | O* | | | | | | | | | | 0.189 | 0.879 |
| 7 | GeolCode 1 | LCC incl. as categories | O | X | O | O | O | O | O | Y | O | | | | | | | X | O* | O | O | O | O | X | O* | 0.276 | 0.937 |
| 8 | GeolCode 1 | LASSO RESULTS | O | | | O | | O | | | | | | | | | | O | O* | O | O | | | | | 0.272 | 0.940 |
| 9 | GeolCode 2 | LCC incl. as categories | O | O | O | O | O | O | O | X | O | | | | | | | O | O* | O | O | O | O | O | O* | 0.145 | 0.825 |
| 10 | GeolCode 2 | LASSO RESULTS | O | O | | O | | O | O | | O | | | | | | | O | O* | O | O | | | | | 0.132 | 0.829 |
| 11 | GeolCode 3 | LCC incl. as categories | O | O | O | O | O | O | O | Y | X | | | | | | | O | O* | O | O | X | X | O | O* | 0.120 | 0.818 |
| 12 | GeolCode 3 | LASSO RESULTS | O | O | O | O | O | O | | | | | | | | | | O | O* | O | O | | | | | 0.119 | 0.816 |
| 13 | GeolCode 4 | Not enough landslides to train a model | | | | | | | | | | | | | | | | | | | | | | | | | |
| 14 | GeolCode 5 | LCC incl. as categories | O | O | O | X | O | O | O | X | X | | | | | | | X | O* | X | X | O | X | O | O* | 0.155 | 0.879 |
| 15 | GeolCode 5 | LASSO RESULTS | O | O | O | | O | O | O | | | | | | | | | | O* | | O | | | | | 0.160 | 0.876 |
| 16 | GeolCode 2 | Rain <220 mm (Lower 2 Quartiles) | O | O | X | O | O | O | O | Y | O | | | | | | X | X | X | X | X | | | | | 0.061 | 0.751 |
| 17 | GeolCode 2 | Rain >220 mm (Upper 2 Quartiles) | O | O | O | O | O | O | O | X | O | | | | | | O | O | O | X | X | | | | | 0.120 | 0.800 |

**Notes**

| | |
|---|---|
| O | P<0.001 |
| X | P>0.001 |
| Y | Co-correlation >0.5 |
| * | Reference level used for the categorical factors |
| | Preferred Models |

models performed slightly better overall, thus justifying the development of models for each GeolCode categorical grouping. However, it is noted that there were not enough landslides in GeolCode 4 within the Training AOI to train a GeolCode 4 model. These materials only cover 0.3% of the Training AOI and <10% of the North Island AOI, so the impact on model training and hindcasting is relatively small.

The GeolCode-grouped factors were then used in the LASSO-penalization regressions (Models 8, 10, 12 and 15, Table 2) to optimize the regressions and select the most important factors contributing to landslide probability. Subsequent logistic regression models were developed that excluded the unimportant factors as identified by the LASSO penalization. LASSO results from all GeolCode-specific models showed that slope, rain, and land cover (LCC) were the factors with the highest importance, whilst the relative importance of the other factors varied between the GeolCode groupings. Only those factors with values given in Table 3 are those shown to be statistically significant by the LASSO penalization. The coefficients for the relevant factors shown in Table 3 are derived from the best fitting LR models (in Table 2) based on the mean of the bootstrap samples—not the LASSO regression coefficients—but adopting only those factors shown to be statistically significant by the LASSO models. The importance of geology—as a grouping factor— combined with slope and landcover is not surprising given the review by Buxton (2025), which showed many previously published models support similar relationships. For example, in New Zealand, Smith et al. (2023) also found geology, land cover, and slope had the largest influence on landslide susceptibility, ahead of intra-event rainfall intensities and pre-event rainfall accumulations. What is interesting about the results presented here—and different from most previous studies—is the relative importance of the Max 24-h rain factor in all the LASSO models; it is the 2nd most important continuous factor after slope.

The relative importance of each of the (statistically significant) factors tested, per GeolCode grouping, is shown in Fig. 4. For context, the relative number of landslide sources (and landslide-classified grid cells) within each GeolCode category group, within the Training AOI, is given in Supplementary Tables S1 and S2. The results shown in Fig. 4 are based on using the standardized values for each factor in the regressions, so the relative weights of the coefficients in the models can be compared. The box and whisker plots show the range in the coefficients derived from 100 bootstrap samples of data—split 75% for training and 25% for testing—randomly selected from each GeolCode group within the Training AOI.

Of all the continuous susceptibility factors tested, the plots show that slope has the most explanatory power on landslide probability. Slope-angle distributions do vary between the different GeolCodes and have previously been used as proxies for material shear strength (e.g., Massey et al. 2018; 2020). Landcover is also an important factor for all GeolCodes. In general, ground with forest cover (LCC 1) tends to have the lowest probability of landslides occurring within it, compared to ground that is bare and urban (LCC 2) or scrub and grassland (LCC 3) cover, noting there are very few urban areas in the Training AOI. For all GeolCode-specific models, the results show that ground with scrub and grassland (LCC 3) cover is overall more susceptible to landslides. When forest (LCC 1) is split into indigenous (LCC 11) and exotic (LCC 12) forest, ground with indigenous forest cover is less susceptible than ground with exotic forest cover. The VegLoss factor is another important explanatory factor for landslides in GeolCodes 1–3, but not 5, indicating slopes that more recently lost vegetation (assumed to be exotic forest in the data set used) correlate with higher landslide probability, indicating a path dependency. GeolCode 5 mainly comprises stronger basement rocks and is either bare (rock exposed at/near surface) or covered by indigenous forest, which tends not to change quickly over time and reduces the susceptibility

of a slope to landslides. Elevation is another important explanatory factor in GeolCodes 3 and 5, where lower elevations show higher levels of landslide probability, which appears to be related to soil thickness—as shown by elevation distributions of the different soil depth codes. Larsen et al. (2023) found that soils on lower elevation slopes tend to be generally thicker than those at higher elevations, due to increased chemical weathering at lower elevations and the accumulation of sediment from physical weathering processes, compared to increased physical weathering at higher elevations. However, thicker soils can develop under indigenous forest at higher elevations. Rainfall also correlates with elevation, but only in locations where increased elevation tends to lead to increasing rain. Such relationships are already accounted for in the VCSN rainfall model used here.

Of all the factors tested, for all GeolCodes, rainfall is one of the most important influencing the estimated landslide probabilities, and it is the most important of the three forcing factors tested. The results show that hindcast landslide probabilities increase with increasing 24-h rain, with the largest increases in probability occurring with rainfalls ≥320 mm (Fig. 5 and Supplementary Figure S4). For GeolCodes 2, 3, and 5, soil moisture deficit (SMD) is also important, and for GeolCodes 3 and 5, the topographic wetness index (TWI) is also important. Both SMD and TWI factors are typically proxies for how wet the ground might have been before (SMD) and during (TWI) Cyclone Gabrielle. For GeolCodes 2, 3, and 5, SMD is correlated with landslide probability, indicating increasing landslide probabilities at increasing SMD levels—i.e., landslides occurred preferentially in materials that were partially/less saturated at the time of the rain. For GeolCodes 1 and 2, TWI has limited explanatory power. For GeolCodes 3 and 5, increasing landslide probabilities occur with increasing TWI values, indicating that landslides occurred in wetter (higher TWI) regions during the storm. For GeolCode 1 (soil), rain is the only statistically significant forcing factor, suggesting SMD and TWI are not important for this GeolCode, which typically includes deeper

**Table 3** Results from the best fitting logistic regression (LR) models (from Table 2), fitted to each GeolCode, where GeolCode 1 (Model 8) comprises Quaternary sands, silts and gravels; GeolCode 2 (Model 10) comprises Neogene limestones, sandstones and siltstones; GeolCode 3 (Model 12) comprises Upper Cretaceous to Paleogene rocks, including limestones, sandstones, siltstones and minor volcanic rocks; and GeolCode 5 (Model 15) comprises Lower Cretaceous Torlesse (Pahau terrane) 'basement' rocks

| | GeolCode1 | Std. Error | GeolCode2 | Std. Error | GeolCode3 | Std. Error | GeolCode5 | Std. Error |
|---|---|---|---|---|---|---|---|---|
| Intercept | −8.754 | 0.053 | −8.405 | 0.035 | −13.08 | 0.17 | −17.84 | 0.60 |
| Elevation | N/A | N/A | N/A | N/A | −0.00198 | 0.00010 | −0.00254 | 0.00042 |
| Slope | 0.09643 | 0.00062 | 0.07406 | 0.00027 | 0.1070 | 0.0013 | 0.0880 | 0.0048 |
| AspectSin | N/A | N/A | 0.3729 | 0.0046 | N/A | N/A | 0.458 | 0.066 |
| AspectCos | N/A | N/A | 0.3235 | 0.0046 | N/A | N/A | 0.460 | 0.067 |
| LSR | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| Curvature | N/A | N/A | -0.01251 | 0.00093 | N/A | N/A | N/A | N/A |
| RainNIWA | 0.00972 | 0.00012 | 0.007919 | 0.000040 | 0.00882 | 0.00026 | 0.0122 | 0.0012 |
| SMD | N/A | N/A | 0.00871 | 0.00033 | 0.0563 | 0.0015 | 0.122 | 0.011 |
| TWI | N/A | N/A | N/A | N/A | 0.1318 | 0.0080 | 0.106 | 0.027 |
| VegLostYearsAgo | −0.0488 | 0.0015 | −0.04342 | 0.00044 | −0.0213 | 0.0025 | N/A | N/A |
| Bare ground and Urban (LCC2) | 0.108 | 0.051 | 0.173 | 0.048 | 1.299 | 0.036 | N/A | N/A |
| Indigenous forest (LCC11) | −1.118 | 0.048 | −0.805 | 0.021 | −0.968 | 0.046 | N/A | N/A |
| Deciduous hardwoods, exotic forest, and harvested forest (LCC12) | 0.230 | 0.032 | −0.370 | 0.018 | −0.509 | 0.040 | −0.53 | 0.20 |
| Scrub and grassland (LCC3*), used as the reference level | 1.0 | | 1.0 | | 1.0 | | 1.0 | |
| Likelihood ratio $X^2$ | −52,053 | | −490,522 | | −38,257 | | −3,368 | |
| *p*-values (max) | 0.033 | | 0.000 | | 0.000 | | 0.009 | |
| RUC:AUC (mean all bootstrap models | 0.94 | | 0.83 | | 0.82 | | 0.88 | |
| Pseudo $R^2$ | 0.27 | | 0.13 | | 0.12 | | 0.16 | |

*Note:* The factors with 'N/A' were not statistically significant based on the LASSO regressions, and thus not included in the LR hindcast models

*LCC3 was chosen as the reference level as it covers the most area within the Training AOI and contains the highest percentage of mapped landslide sources and landslide classified sample grid cells

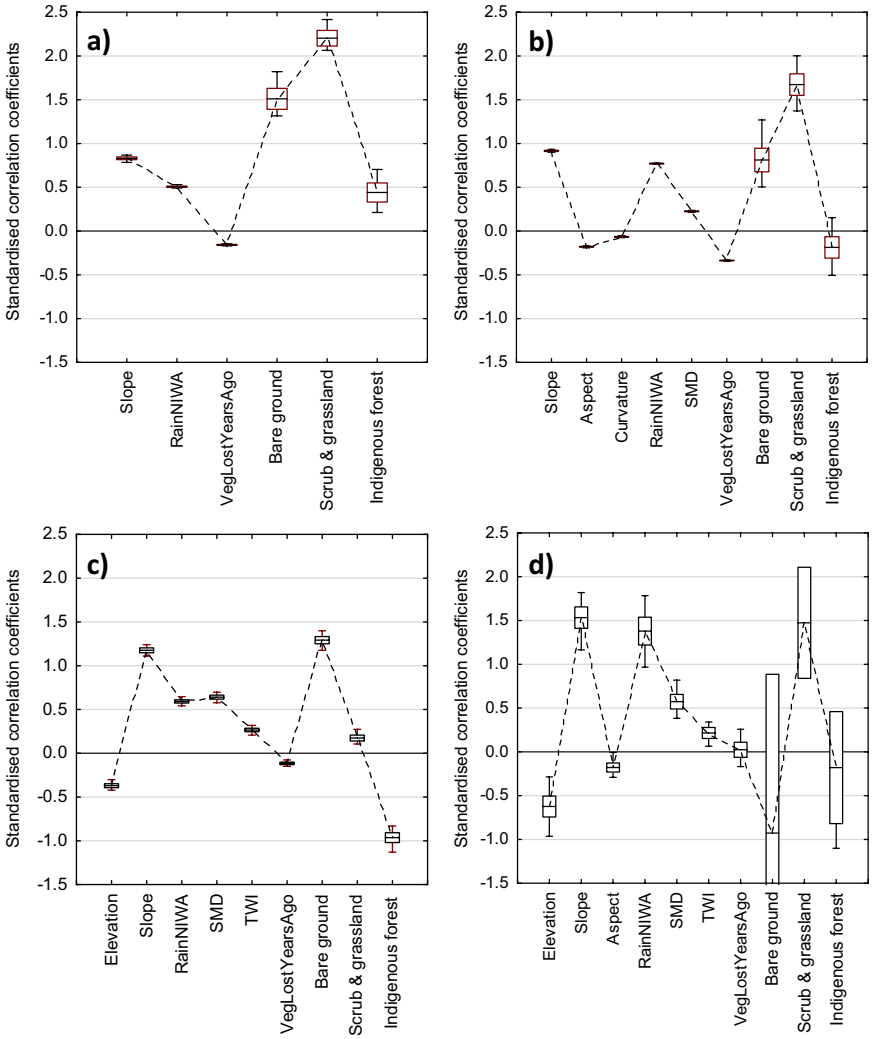



**Fig. 4** Bootstrap model results for those statistically significant factors, derived from the LASSO regression, per GeolCode grouping. The values for each of the continuous factors used in these regressions were standardized to allow the relative importance of each factor to be compared. Those factors with the largest ranges (boxes and whiskers) and closest to zero reflect those with lower statistical significance. The vegetation lost years before Cyclone Gabrielle VegLostYearsAgo (Veg Loss) and Land Cover (LCC) categories were not standardized. The ranges (boxes and whiskers) were derived using "Bootstrapping," where 100 bootstrap samples were taken, with each bootstrap sample comprising 75% of the total dataset. Regression models were fitted to each bootstrap sample and the results were combined to calculate the means (dashed lines), standard deviations (the boxes), and the non-outlier variances (the whiskers) for each factor. For the categorical factor of land cover, the exotic forest land cover code was used as the reference in the regressions: **a** GeolCode 1, Quaternary sands; silts and gravels; **b** GeolCode 2, Neogene limestones, sandstones, and siltstones; **c** GeolCode 3, Upper Cretaceous to Paleogene rocks, including limestones, sandstones, siltstones, and minor volcanic rocks; and (**d**) GeolCode 5, Lower Cretaceous Torlesse (Pahau terrane) "basement" rocks

soils (derived from alluvium and colluvium) and those not derived from the in situ weathering of rock.

Overall, the results from each GeolCode-specific model (Fig. 5) show that GeolCode 1 (soils derived from Quaternary sands, silts and gravels) and 2 (soils derived from Neogene sedimentary rocks) show higher levels of landslide probability, when compared to those estimated for soils derived in GeolCodes 3 (upper Cretaceous to Paleogene rocks) and 5 (Pahau terrane sedimentary rocks). GeolCode 5 has the lowest landslide probabilities; this is thought to be because the soils are very shallow or not present with rock exposed at/near surface, and GeolCode 5 tends to be covered by indigenous forest. These comparisons were made by calculating the landslide probabilities using the best fitting regression "models" (from Table 3) for various 24-h rainfall amounts (0 to 500 mm as per the range during the event) and adopting a value of 35° for the slope factor used in all regressions, and zero values for all other factors. A slope of 35° was chosen as this represents the peak landslide source area frequency for landslides within all GeolCodes combined, and it is the most important factor used in the regressions.

### Modelled hindcast versus recorded landslide-classified grid cells

The coefficients listed in Table 3 are used to calculate the GeolCode-specific landslide probabilities, which were summed from low to high hindcast probability to allow the hindcast number of

landslide-classified grid cells to be compared to the cumulative number of corresponding recorded landslide-classified grid cells (CumH-R). Overall, these relationships show a general linear trend (Fig. 6), indicating the hindcasts for GeolCodes 2, 3, and 5 have no apparent biases in any part of their range. The CumH-R plot for GeolCode 1 (Fig. 6) is biased—slightly non-linear—in the upper part of the range, where the hindcast number of landslide-classified grid cells is increasing more rapidly than the number of recorded landslide-classified grid cells.

Overall, the cumulative sum of the modelled hindcast number of landslide-classified grid cells matches the cumulative sum of the recorded landslide-classified grid cells relatively well (Table 3). The factors of difference—hindcast versus recorded number of landslide-classified grid cells—range between 0.7 and 1.8, with a mean of 1.2 for the GeolCode-specific models (Supplementary Table S2). For GeolCode 2, comprising 86% of the recorded landslide-classified grid cells, the factors of difference between the hindcast versus recorded number of landslide-classified grid cells are 1.0, indicating a good relationship.

The results of the best-fitting LR models (Table 3) were converted to the summed probabilities, per GeolCode, and multiplied by the "factor of difference" between the number of landslide-classified grid cells and mapped landslides (Supplementary Table S2), to estimate the total number of landslides within each GeolCode group. The results give an estimate of between 806,000 (±50,000) landslide-classified grid cells, assuming no probability "cut-off" threshold, and 706,000 (±40,000), adopting the upper quartile probability cut-off threshold of 0.00167, meaning that any probability value below this is not summed. Refer to Supplementary Section S2., Step 2 for more description of the methods used. The range is estimated based on the standard error of each GeolCode-specific regression. Using the landslide area-frequency distribution derived from the mapped landslides within the Training AOI (Supplementary Table S2), the ≈800,000 estimated number of landslides across the North Island AOI approximates to an estimated area of landsliding (sources only) of about 104 km² assuming no probability cut-off threshold, and 91 km² adopting the upper quartile probabilities only.

### Uncertainties

These landslide estimates are likely to be at the lower end of the actual number of landslides triggered, as they do not account for landslides that might have occurred in the GeolCodes not included in the Training AOI—representing 15% of the North Island AOI. They are also likely to be underestimated by a mean factor of 1.2 based on using the 8 by 8 m LINZ DEM to derive the factors adopted in the North Island AOI hindcast. The uncertainty range is also likely to be underestimated, given the LR models are trained and tested on the same dataset, and as a result, they are likely to be understated in those grid cells outside the Training AOI. These estimates of the affected area of landsliding also do not consider the area affected by landslide debris. Hancox and Wright (2005b) noted that the landslide "footprint" (source and runout areas) had an average ratio of runout length to footprint length of between 2.5 and 3.1.

There are several other factors that contribute towards the uncertainties on the results presented in this study. For example, this study primarily captures small (<46 m²), shallow slides. Although larger, deeper-seated, or slow-moving landslides are included in the inventory, they might be underrepresented as they tend to be harder to identify and map. The method adopted to hindcast landslide probabilities and numbers assumes similar landslide responses across regions with similar rainfall and geology. Such assumptions can overlook local hydrological or mechanical variabilities, which are not captured in the "proxy" factor datasets. For example, soil moisture and topographic wetness index (SMD and TWI) used in the training and hindcasting are estimated from regional-scale models, and local-scale hydrological controls on landsliding are not considered. Whilst the SMD models adopt limited in situ measurements, the density of such measurements across the region is poor. In addition, better model performance may be achieved by varying the type of statistical model used to explore non-linearity between landslide probability and the input factors used, which are not captured by the logistic regression method.

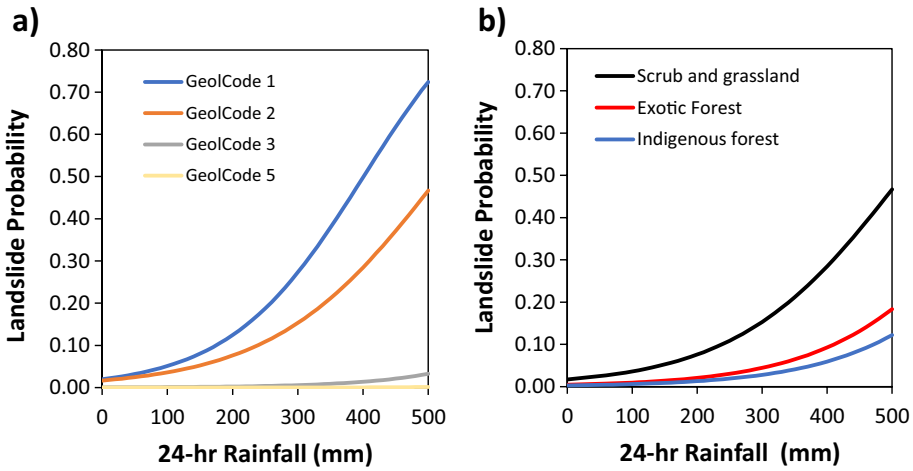



**Fig. 5** Landslide probabilities calculated using the best fitting regression "models" (from Table 3) for various 24-h rainfall amounts and adopting a value of 35° for the slope factor used in all regressions, and zero values for all other factors used in each regression model. **a** Landslide probabilities calculated for each GeolCode. **b** Landslide probabilities calculated for GeolCode 2, per Land Cover Category

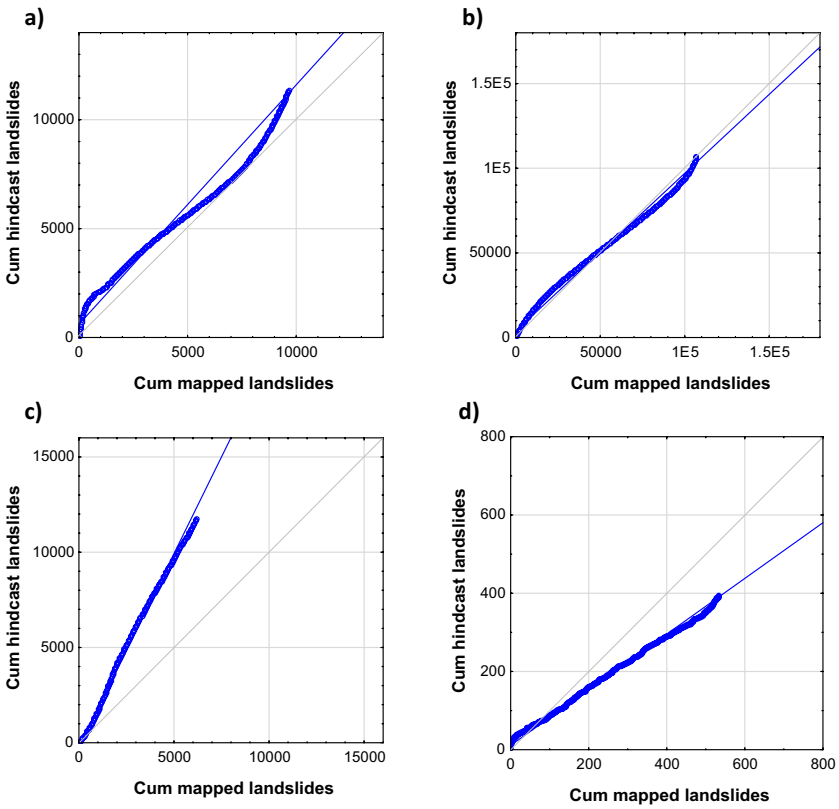



**Fig. 6** CumF-R plots per GeolCode. The cumulative hindcast versus mapped (recorded) number of landslide grid cells (16 × 16 m) from the best fitting models (Table 3), within the Training AOI. **a** Quaternary sands, silts, and gravels (GeolCode 1). **b** Neogene limestones, sandstones, and siltstones (GeolCode 2). **c** Upper Cretaceous to Paleogene rocks, including limestones, sandstones, siltstones, and minor volcanic rocks (GeolCode 3). **d** Lower Cretaceous Torlesse (Pahau terrane) "basement" rocks (GeolCode 5). The trends are linear fits to the given data. Note: the data were created by calculating the landslide probabilities from the best fitting models for all grid cells within the Training AOI. The data were then ordered from low to high hindcast probability, and the probabilities cumulatively summed from low to high. The corresponding landslide source area-classified grid cells (1 or 0) were also cumulatively summed

## Discussion

### Landslide densities

At the regional scale, landslide impact may be represented simply by landslide point or areal density—the cumulative area of landslide footprints, i.e., source and runout, represented as a percentage of the study area [e.g., Hancox and Wright (2005a; b); Crozier (2017)]. The total number of landslides estimated to have occurred within the North Island AOI is in the order of 800,000 (Fig. 7).

The landslide source number and area densities calculated using the human-mapped landslides, inside the priority mapping grid cells within the Training AOI, are log-normally distributed, ranging from 0 to 327 landslides/km$^2$, with a mean of 28 and a standard deviation of 46 landslides/km$^2$, and skew of 2.7. The aerial density ranges from 0 to 3.8%, with a mean of 0.4% and a standard deviation of 0.6% with a skew of 2.6. These landslide source-number densities are high to extreme when compared to those from overseas studies that occurred under comparable rainfall amounts. For these overseas events, landslide number densities range from 10 to 28 landslides/km$^2$ (e.g., Ko 2003; Turner et al. 2010), and between 0.01 and 1.0% for aerial density (e.g., Marc et al. 2018). These densities represent mean reported values. However, the Cyclone Gabrielle landslide densities are comparable to those caused by previous MORLEs in New Zealand, associated with similar 24-h and total event rainfalls. Examples include the following: multiple events in the Wairarapa, e.g., Crozier et al. (1980) and Basher et al. (2018); Cyclone Bola in 1988, e.g., Marden and Rowan (1993) and Page et al. (1994; 1999); the Manawatu storm in 2004, e.g., Hancox and Wright (2005a,b) and Crozier (2017); and multiple storm events recorded in similar materials that occurred within the Training AOI, e.g., Page et al. (1999) and Reid and Page (2002). What causes the New Zealand landslide densities to be so high? Multiple authors (e.g., Reid and Page 2002; Marden and Rowan 1993; Marden 2004; Philips et al. 2024; and Gilmore et al. 2023) suggest that it is a combination of the removal of indigenous forest and changing land use (pasture), coupled with shallow soils, and frequently occurring high-intensity rain events.

### Exploring the influence of rainfall, soil moisture, and groundwater

Crozier (1999) and Glade (2000) demonstrated that in many rainfall events the antecedent water status within the slope prior to the

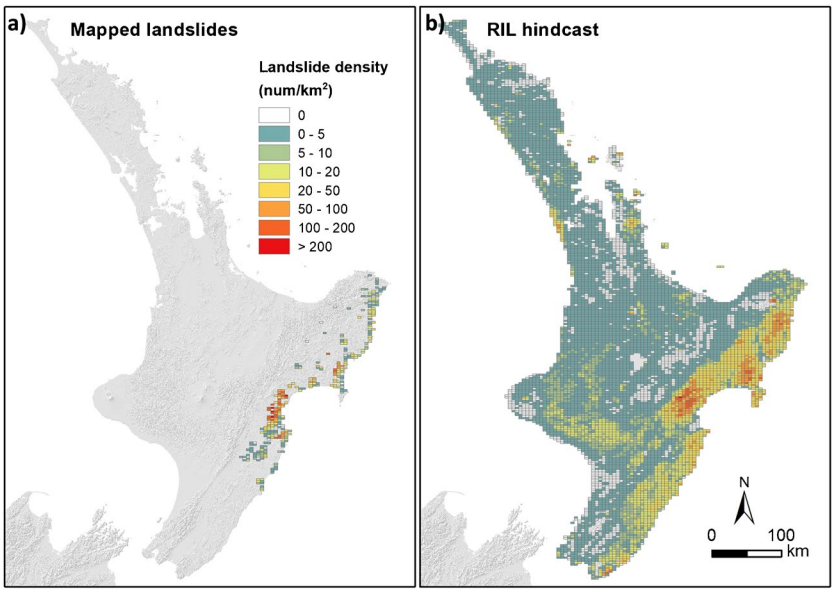



**Fig. 7** Maps showing the landslide densities calculated for the priority mapping grid cells in the Training AOI and North Island AOI. The units are the number of mapped landslides per km$^2$ within the given priority mapping grid cell (num. LS/km$^2$). **a** Landslide densities calculated using the mapped landslides within those priority grid cells that were mapped. **b** The estimated landslide densities calculated using the hindcast landslide probabilities for the different GeolCodes and Land Cover categories using the best fitting regression models (Table 3)

event will dictate the amount of storm rainfall (event water) needed to reach the water content (mass of water divided by mass of soil in a soil sample, expressed as a percentage) and pore-water pressures to initiate landslides. They suggested the amount of rainfall required to initiate landsliding will be reduced if there has been a wet period before the event (Crozier 2017). As well as the antecedent rainfall and evapotranspiration, physical factors that promote the concentration or drainage of "slope water" during a storm are also important; these include elements of surface and subsurface topography as well as substrate conditions such as permeability pathways and contrasts and the void ratio of the soil (volume of voids/volume of solids in a soil. The relationships between rainfall intensity, surface and subsurface flow paths, and ultimately soil water content and pore-water pressure are complex and rarely measured in situ, thus being represented by regional proxies. For example, the frequency-ratio analyses results show that 91% of landslide-classified sample grid cells occurred in soils where the SMD (a proxy for soil water content) was around 45–75 mm—minimum is 0 mm (wettest) and maximum is 150 mm (driest)—suggesting the ground is more susceptible to landslides when it is partially saturated (neither at its wettest or driest). This assumes the estimated SMD the day before Cyclone Gabrielle, which is derived from the standard NIWA VCSN soil moisture model (Kandel et al. 2005), is representative of actual conditions. There might be a SMD threshold where, if the soil is too dry, it is less susceptible to landslides, leading to more runoff than infiltration. The opposite is also true; if SMD is low (wettest)—indicating high antecedent rainfalls—then the slope might be less susceptible to failure because it may have failed previously, rendering it less susceptible in the subsequent rain event.

Mondini et al. (2023) suggest it is possible to adequately anticipate the location and timing of rainfall-induced landslides over large areas using rainfall values alone (without considering antecedent conditions). In our study, the LR model results show that rainfall is the dominant landslide forcing factor and that SMD and TWI are only marginally important. To further analyze the importance of rainfall, regressions were carried out on the GeolCode 2 dataset. GeolCode 2 was chosen as it contains 86.4% of the mapped landslide-classified grid cells within the Training AOI. Sample grid cells were grouped into rainfall bands adopting quartiles where quartile 1 = 25 to 110 mm/24 h, quartile 2 = >110 to 220 mm/24 h, quartile 3 = >220 to 330 mm/24 h, and quartile 4 = >330 to 436 mm/24 h. The rainfall ranges for each quartile show that the priority mapping grid cells, which comprise the Training AOI, represent the full range of rainfall experienced by the region during Cyclone Gabrielle and not just those areas that experienced extreme rain (Fig. 1). Regressions were carried out on the grouped sample grid cell data, comprising the 1st and 2nd rainfall quartiles, the 3rd and 4th quartiles, and the 4th quartile on its own. Models are fitted using those factors listed in Table 3 as being statistically important, adopting standardized values for these factors. Bootstrapping (Supplementary Section S1.1.2) was used to explore the variance of each factor's standardized coefficient (Fig. 8). All regression results show that slope is the main dominant factor and rainfall the second most dominant. Rainfall is almost as dominant as slope in the regressions carried out on the 1st and 2nd quartiles of rainfall data. However, the importance of rainfall reduces in the regressions carried out on the combined 3rd and 4th quartile and reduces again in the regressions carried out on the 4th quartile data only. The importance of elevation, aspect, LSR, curvature, and TWI on landslide susceptibility does not vary much between the regressions. However, SMD becomes less important at higher rainfalls and VegLostYearsAgo becomes more important.

The main results from these regressions show that 24-h rainfall becomes less important as a landslide forcing factor, at high (>220

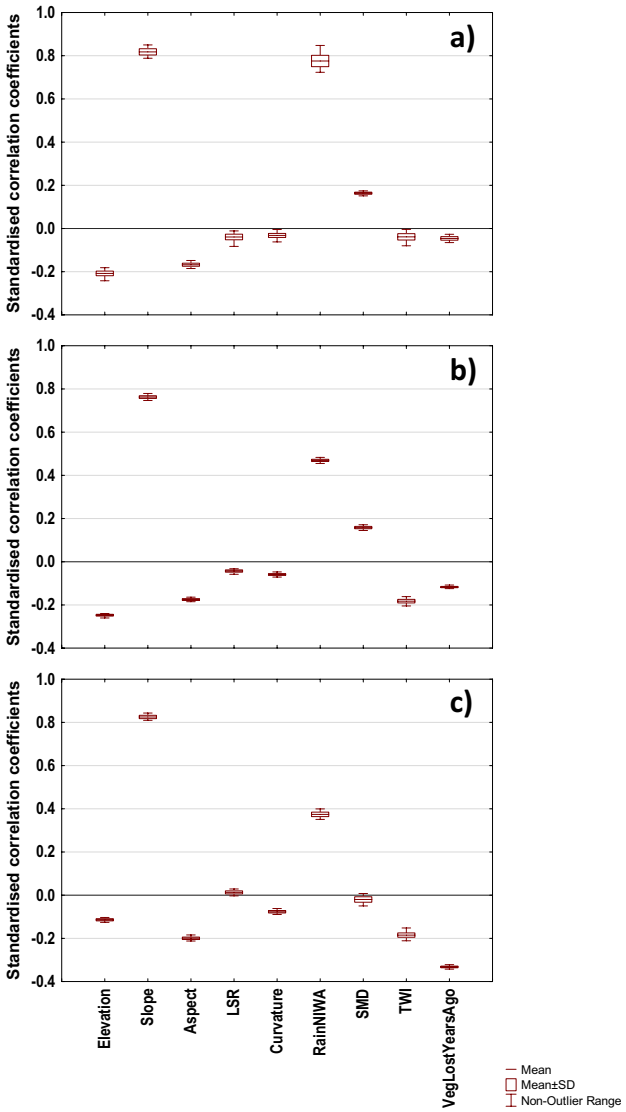



**Fig. 8** GeolCode 2 bootstrap model results for all continuous factors used in the regressions, adopting different rainfall ranges. To do this, the sample grid cells (and their corresponding values for the different continuous factors) were grouped into rainfall bands adopting quartiles where quartile 1 = 25 to 110 mm/24 h, quartile 2 = >110 to 220 mm/24 h, quartile 3 = >220 to 330 mm/24 h, and quartile 4 = >330 to 436 mm/24 h. **a** Regression results adopting the 1 st and 2nd quartiles of data. **b** Regression results adopting the 3rd and 4th quartiles of data. **c** Regression results adopting the 4th quartile of data. Note: The values for each of the continuous factors used in these regressions were standardized to allow the relative importance of each factor to be compared. Those factors with the largest ranges (boxes and whiskers) and closest to zero reflect those with low statistical significance. The ranges (boxes and whiskers) were derived using "Bootstrapping," where 100 bootstrap samples were taken, with each bootstrap sample comprising 75% of the total dataset. Regression models were fitted to each bootstrap sample, and the results were combined to calculate the means, standard deviations (the boxes), and the non-outlier variances (the whiskers) for each factor

mm) to very high (>330 mm) rain amounts. These results suggest that such extreme rain amounts may not drive the highest probability of landslide occurrence. This might be because such high rainfalls are likely to be associated with intensities that exceed the infiltration capacity of the ground. Alternatively, the ground may already be fully saturated (and already failed where susceptible), such that additional rainfall does not increase the landslide rate. This is supported in part by the SMD data. Whilst SMD is marginally important in the regressions carried out on the combined 1st and 2nd and combined 3rd and 4th quartiles of data, it becomes irrelevant in the 4th quartile regressions, suggesting saturated conditions are reached regardless of antecedent conditions. These results suggest that the 24-h rainfalls experienced by some areas during cyclone Gabrielle were likely to have been sufficiently high enough to "overprint" antecedent soil water conditions. In these areas, soil water contents within the slopes before Cyclone Gabrielle rainfall occurred may not have been important for generating landslides, supporting the findings of Mondini et al. (2023). However, at lower rainfall values, the antecedent soil water contents remain important, consistent with the antecedent rainfall hypothesis (e.g., of Crozier 1999 and Glade 2000). These findings support the cell model for MORLEs proposed by Crozier (2017) which argues that the importance of susceptibility factors differs between different zones of rainfall intensity.

### Exploring the influence of land cover on landslide occurrence

Overall, the results show that slopes with tree cover are less susceptible to landslides, compared to slopes that are bare or have scrub and grassland cover (Fig. 5). Such observations have been well documented previously (e.g., Spiekermann et al. 2022). However, it should be noted that in this case, the effect of vegetation might be exaggerated because of poor ground visibility in forested areas—leading to landslide mapping bias. However, Leith et al. (2025) show that it is unlikely there are enough unobserved landslides in forested areas to change these results.

Bergin et al. (1995) found that on the East Coast of New Zealand, landslide occurrence in fully stocked stands of reverting Kanuka and Manuka scrub (LCC3) was 65% lower than those in pasture, and 90% lower in 20-year-old indigenous forest. Marden (2004) found that following the Manawatu storm in 2004, landsliding under forest was 90% less than that under pasture, and 80% less than that under scrub. The results from this study show similar findings, where slopes that had scrub and grassland cover had significantly higher landslide probabilities compared to those with forest cover (Fig. 5b), with mean factors of difference of 5.4 times higher for indigenous forest, and 3.4 times higher for exotic forest, where a factor of 10 would be one order of magnitude higher.

Gilmore et al. (2023) state that the understanding of how vegetation contributes to slope stability and erosion control is well advanced. A closed canopy forest cover contributes to an increase in slope stability and a significant reduction in erosion by reducing the ability of rainfall to cause slope failure through the processes of interception and evapotranspiration, whilst the roots provide mechanical reinforcement and are how trees extract moisture from the soil to reduce soil water contents. These processes become most effective after branches of individual trees touch (canopy closure) and lateral roots of adjacent trees overlap (full-root occupancy). The rate at which canopy closure and root occupancy occur is largely determined by plant spacing, growth rate, and silviculture (tree tending) regimes. Marden and Rowan (1993) and Gilmore et al.

(2023) showed that during Cyclone Bola, trees helped to reduce the frequency/density of shallow landslides. They found that soils under a closed canopy forest are less prone to rainfall-induced landslides than similar soils under pasture, or if the canopy is more open, such as in young stands of pines and scattered regenerating scrub.

Our results (Fig. 3) also show that areas of ground that lost trees a few years before the storm occured (VegLostYearsAgo) have an increased susceptibility to landslides, and that at higher rainfalls, the importance of this factor in the landslide susceptibility regressions increases (Fig. 8). This finding is similar to the findings of others. For example, during Cyclone Bola in New Zealand (1988), areas under closed canopy indigenous forest (LCC 11) and exotic plantations (LCC 12) older than 8 years were 16 times less susceptible to landsliding than areas in pasture (scrub and grassland—LCC 3) and under exotic pines (LCC 12) less than 6 years old, and four times less susceptible than closed canopy regenerating scrub (age unknown) and exotic pines 6–8 years old. Results from this study (Fig. 5) show that landslides preferentially occurred on slopes that had lost forest cover 3–5 years before Cyclone Gabrielle. This finding is consistent with the "window of vulnerability" (Sidle 2005) for forestry rotations, relating to the time it takes for the root system to decay post harvesting and the time for new growth to compensate for this. For example, in a study from New Zealand, including a site in Tairāwhiti-Gisborne, Phillips et al. (2024) measured that the maximum rates of landsliding occurred an average of 2–3 years following harvesting.

## Conclusions

The total estimated number of hindcast landslides initiated by Cyclone Gabrielle across the North Island area of interest (AOI) was in the order of >800,000. This number of landslides approximates to an area of landsliding (sources only) of about 91–104 km$^2$. These estimates are likely to be at the lower end of the actual number triggered and area affected. In terms of the number of landslides triggered, this suggests Cyclone Gabrielle was one of the most extreme landslide-triggering events recorded globally, far exceeding even the highest estimates of the number of landslides triggered by the 2008 Wenchuan earthquake.

Of the 116,655 human-mapped landslides extracted from the V1.0 landslide inventory for Cyclone Gabrielle within the Tairāwhiti-Gisborne and Hawke's Bay (the Training AOI), most were small slides and flows in shallow soils covered by scrub and grassland—pasture. The density of mapped landslides in the Training AOI ranged from a mean of 28 to a max of 327 per km$^2$. Such densities are similar to past New Zealand multiple-occurrence landslide events MORLEs, but much larger than internationally recorded events, again highlighting the extreme nature of Cyclone Gabrielle.

The results show that rainfall is one of the main controlling factors on landslide occurrence/density, but the relationship is complex. This study shows that 24-h rainfall becomes less important, as a landslide forcing factor, at high (>220 mm) to very high (>330 mm) rainfalls. Results suggest that the 24-h rainfall experienced by some areas during Cyclone Gabrielle were—in some areas—likely to have been sufficiently high enough to "overprint" antecedent soil water conditions. This suggests that in these areas, soil water content within the slopes before Cyclone Gabrielle rainfall occurred may not have been important for generating landslides in these areas.

The results from the multivariate analyses show that geology and land cover were important high-level factors influencing landslide occurrence. Results from all GeolCode-specific regression models showed that slope, rainfall, and land cover were the factors with the highest importance, whilst the relative importance of the other factors varied between the GeolCode groupings. After slope, the 24-h maximum rainfall over the duration of the storm was the second most important factor on landslide susceptibility. Land cover was also an important factor for all GeolCodes.

For all GeolCode-specific models, the results show that ground with scrub and grassland cover is overall more susceptible to landslides. When forest is split into indigenous and exotic forest, ground with indigenous forest cover is less susceptible than ground with exotic forest cover. The vegetation lost years before Cyclone Gabrielle factor is another important explanatory factor for landslides in GeolCodes 1–3, but not 5, indicating slopes that more recently lost vegetation (assumed to be forest in the data set used) correlate with higher landslide probability, indicating a path dependency. This study further supports the findings from previous studies, which show the benefits of exotic and indigenous forest on reducing the occurrence of shallow rainfall-induced landslides.

## Acknowledgements

The authors would like to acknowledge the team of landslide mappers and spatial database experts who contributed towards the mapping of V1.0 landslide inventory. These people are, in no particular order, Dougal Townsend, David Barrell, Alec Zoeller, Alfredo Jaramillo, Amelia Lin, Andrea Wolter, Andrew Boyes, Anna Knoltsch, Ariane Pola, David Heron, Hanna Lyford, Harley Betts, Hui Zhou, Imogen Daysh, Jason Farr, Joe Pelmard, Joshua Wight, Julia Harvey, Julie Lee, Katie Jones, Kuba Rozmus, Laura Gnesko, Livio Dreyer, Luke Easterbrook-Clarke, Matt Hill, Niraj Bal Tamang, Phil Glassey, Rachel Lawson, Salman Ashraf, Saskia de Vilder, and Zane Bruce. We would also like to acknowledge Dr Rob Buxton and Dr Bipin Peethambaran (GNS Science) and the three anonymous reviewers, for providing their constructive comments on previous versions of the manuscript.

## Funding

This study was funded by New Zealand's Ministry of Business, Innovation and Employment, Award Number: C05X170 and C05X2304.

## Data Availability

The data used in the analyses presented in this paper can be retrieved from https://osf.io/8sbrj. Leith K, McColl S, Robinson T, Massey C, Smith H, Wotherspoon L, Carey-Smith T, Ashraf S, Barrell D, Betts H, et al. 2025. 2023 Cyclone Gabrielle Landslide Data. https://doi.org/10.17605/OSF.IO/QBMD9.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

## References

Basher L, Betts H, Lynn I, Marden M, McNeill S, Page M, Rosser B (2018) A preliminary assessment of the impact of landslide, earthflow, and gully erosion on soil carbon stocks in New Zealand. Geomorphology 307:93–106

Bergin DO, Kimberley MO, Marden M (1995) How soon does regenerating scrub control erosion? N Z J For 38(2):38–40

Brown A, Milton S, Cullen M, Golding B, Mitchell J, Shelly A (2012) Unified modelling and prediction of weather and climate: a 25-year journey. Bull. Amer. Meteor. Soc. 93:1865–1877. https://doi.org/10.1175/BAMS-D-12-00018.1

Buxton R (2025) Machine learning landslide susceptibility models – a review. GNS Science Report 2025/36. Lower Hutt (NZ): GNS Science. https://doi.org/10.21420/BDCH-EY76

Crozier MJ, Eyles RJ, Marx SL, McConchie JA, Owen RC (1980) Distributions of landslips in the Wairarapa hill country. New Zealand Journal of Geology and Geophysics 23:575–586

Crozier MJ (1999) Prediction of rainfall-triggered landslides: a test of the antecedent water status model. ESPL 24(9). https://doi.org/10.1002/(SICI)1096-9837(199908)

Crozier MJ (2005) Multiple-occurrence regional landslide events in New Zealand: Hazard management issues. Landslides 2:247–256. https://doi.org/10.1007/s10346-005-0019-7

Crozier MJ (2017) A proposed cell model for multiple-occurrence regional landslide events: implications for landslide susceptibility mapping. Geomorphology 295:480-488. ISSN 0169-555X. https://doi.org/10.1016/j.geomorph.2017.07.032. https://www.sciencedirect.com/science/article/pii/S0169555X17301332. Accessed 21 Jun 2025

Efron B, Tibshirani RJ (1994) An introduction to the bootstrap. 1st Edition. Chapman and Hall/CRC. https://doi.org/10.1201/9780429246593

Froude MJ (2018) Petley DN (2018) Global fatal landslide occurrence from 2004 to 2016. Nat. Hazards Earth Syst. Sci. 18:2161–2181

Gilmore B, Marden M, Visser R (2023) Expert Report on land use issues in Tairāwhiti and forestry management practices in response. Appendix 3. MILU-Combined Citizen Space Submission. https://environment.govt.nz/assets/what-government-is-ng/Land/MILU-Combined-Citizen-Space-Submission-Attachments-Part-1.pdf. Accessed 21 Jun 2025

Glade T (2000) Modelling landslide-trigering rainfalls in different regions of New Zealand - the soil water status model. Zeitschrift fur Geomorphologie, Supplementband 122:63–84

Gómez D, García EF (2023) Aristizábal E (2023) Spatial and temporal landslide distributions using global and open landslide databases. Nat Hazards 117:25–55. https://doi.org/10.1007/s11069-023-05848-8

Hancox, GT, Wright K (2005a) Landslides caused by the February 2004 rainstorm and floods in southern North Island New Zealand. In: Institute of Geological and Nuclear Sciences Science Report 2005/10, p 32

Hancox, GT, Wright K (2005b) Analysis of terrain and landsliding caused by the 15–17 February 2004 rainstorm in the Wanganui-Manawatu hill country, southern North Island, New Zealand. Lower Hutt: Institute of Geological & Nuclear Sciences Limited. Institute of Geological & Nuclear Sciences science report 2005/11, p 64

Jia G et al (2021) A global landslide non-susceptibility map. Geomorphology 389:107804

Ko FWY (2003) Correlation between rainfall and natural hazard terrain landslide occurrence in Hong Kong. GEO Report No. 168. Originally produced in December 2003 as GEO Special Project Report No. SPR 7/2003. Geotechnical Engineering Office, Hong Kong, p 74. https://www.scribd.com/document/470986947/er168links

Kritikos T, Robinson TR, Davies TRH (2015) Regional coseismic landslide hazard assessment without historical landslide inventories: a new approach. J Geophys Res Earth Surf 120(4). https://doi.org/10.1002/2014jf003224

Reid LM, Page MJ (2002) Magnitude and frequency of landsliding in a large New Zealand catchment. Geomorphology 49(1-2):71–88

Larsen I, Eger A, Almond PC, Thaler EA, Rhodes JM, Prasicek G (2023). The influence of erosion and vegetation on soil production and chemical weathering rates in the Southern Alps, New Zealand. Earth and Planetary Science Letters. 608:118036. ISSN 0012-821X. https://doi.org/10.1016/j.epsl.2023.118036 .https://www.sciencedirect.com/science/article/pii/S0012821X23000493. Accessed 21 Jun 2025

Kandel D, Tait A, Woods R (2005) A soil moisture model for New Zealand. Journal of Hydrology (New Zealand) 44(2):123–135

Leith K, McColl S, Robinson T, Massey CI, Townsend DB, Rosser BJ, Smith H, Barrell DJA, Zoeller A, Lukovic B et al (2025) Cyclone Gabrielle (12-16 February 2023): landslide inventory for North Island New Zealand, version 1.0. Lower Hutt (NZ): GNS Science. p 72 (GNS Science report; 2023/28). https://doi.org/10.21420/MMFP-S330

Lombardo L, Mai PM (2018) Presenting logistic regression-based landslide susceptibility results. Engineering Geology. 244:14–24. https://doi.org/10.1016/j.enggeo.2018.07.019

Marc O, Stumpf A, Malet J-P, Gosset M, Uchida T, Chiang S-H (2018) Initial insights from a global database of rainfall-induced landslide inventories: the weak influence of slope and strong influence of total storm rainfall, Earth Surf. Dynam. 6:903–922. https://doi.org/10.5194/esurf-6-903-2018

Marden M (2004) Future-proofing erosion-prone hill country against soil degradation and loss during large storm events: have past lessons been heeded? New Zealand Journal of Forestry 49:11–16

Marden M, Rowan D (1993) Protective value of vegetation on tertiary terrain before and during Cyclone Bola, East Coast, North Island, New Zealand. NZJ For Sci 23(3):255-263

Massey CI, Townsend DB, Rathje E, Allstadt KE et al (2018) Landslides triggered by the 14 November 2016 Mw 7.8 Kaikoura earthquake, New Zealand. Bull Seismol Soc Am 108(3B):1630-1648. https://doi.org/10.1785/0120170305

Massey, CI, Townsend DB, Lukovic B, Morgenstern R, Jones KE, Rosser BJ, de Vilder SJ (2020) Landslides triggered by the 14 November 2016 Mw 7.8 Kaikoura earthquake : an update. Landslides 17(10):2401-2408. https://doi.org/10.1007/s10346-020-01439-x

Massey CI, Wolter A, Huso R, Lukovic B, Brideau M-A (2022) Coseismic landslide susceptibility and triggering analyses. In: Towhata I, Wang G, Xu Q, Massey CI (eds), Coseismic landslides: phenomena, long-term effects and mitigation. Singapore: Springer. Springer natural hazards. 633-679. https://doi.org/10.1007/978-981-19-6597-5_18

MFAT (2023) Rapid assessment of land damage - Cyclone Gabrielle. Contract report for Ministry of Environment, LC4292 MFAT: https://www.mfat.govt.nz/en/trade/mfat-market-reports/cyclone-gabrielles-impact-on-the-new-zealand-economy-and-exports-march-2023. Accessed 21 Jun 2025

Mondini AC, Guzzetti F, Melillo M (2023) Deep learning forecast of rainfall-induced shallow landslides. Nat Commun 14:2466. https://doi.org/10.1038/s41467-023-38135-y

Nadim F, Kjekstad O, Peduzzi P, Herold C, Jaedicke C (2006) Global landslide and avalanche hotspots. Landslides 3:159–173. https://doi.org/10.1007/s10346-006-0036-1

NIWA (2023) https://niwa.co.nz/sites/niwa.co.nz/files/Climate_Summary_February_2023_NIWA-web.pdf. Accessed 21 Jun 2025

Page MJ, Trustrum NA, Dymond JR (1994) Sediment budget to assess the geomorphic effect of a cyclonic storm. New Zealand. Geomorphology 9(3):169–188

Page MJ, Reid LM, Lynn IH (1999) Sediment production from Cyclone Bola landslides, Waipaoa catchment. Journal of Hydrology, New Zealand. 38(2):289–308

Parker RN, Hancox GT, Petley DN, Massey CI, Densmore AL, Rosser NJ (2015) Spatial distributions of earthquake-induced landslides and hillslope preconditioning in the northwest South Island, New Zealand. Earth Surf Dyn 3(4):501–525. https://doi.org/10.5194/esurf-3-501-2015

Pörtner H-O, DC, Roberts H, Adams I et al (2022) Impacts, adaptation and vulnerability. In: Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegría A, Craig M, Langsdorf S, Löschke S, Möller V, Okem A, Rama B (eds) Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge and New York, USA, pp 37–118. https://doi.org/10.1017/9781009325844.002

Reid LM, Page MJ (2003) Magnitude and frequency of landsliding in a large New Zealand catchment. Geomorphology. 49(1–2):71–88

Phillips C, Betts H, Smith HG, Tsyplenkov A (2024) Exploring the post-harvest 'window of vulnerability' to landslides in New Zealand steepland plantation forests. Ecol Eng 107300:107300. https://doi.org/10.1016/j.ecoleng.2024.107300

Sidle RC (2005) Influence of forest harvesting activities on debris avalanches and flows. In: Debris-flow Hazards and Related Phenomena. Springer Praxis Books. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-27129-5_16

Smith HG, Neverman AJ, Betts H, Spiekermann R (2023) The influence of spatial patterns in rainfall on shallow landslides. Geomorphology. 437:108795. ISSN 0169-555X. https://doi.org/10.1016/j.geomorph.2023.108795. https://www.sciencedirect.com/science/article/pii/S0169555X23002155. Accessed 21 Jun 2025

Spiekermann RI, Smith HG, McColl S, Burkitt L, Fuller IC (2022) Quantifying effectiveness of trees for landslide erosion control. Geomorphology 396:107993. ISSN 0169-555X. https://www.sciencedirect.com/science/article/pii/S0169555X21004013. Accessed 21 Jun 2025

Stone DA, Rosier SM, Bird L, Harrington LJ, Rana S, Stuart S, Dean SM (2022) The effect of experiment conditioning on estimates of human influence on extreme weather. Weather Climate Extremes 36:100427

Tait A, Henderson R, Turner R, Zheng X (2006) Thin plate smoothing spline interpolation of daily rainfall for New Zealand using a climatological rainfall surface. International Journal of Climatology 26(14):2097–2115

Tait A, Sturman J, Clark M (2012) An assessment of the accuracy of interpolated daily rainfall for New Zealand. J Hydrology (New Zealand) 51(1):25–44. https://search.informit.org/doi/10.3316/informit.628479729772838

Turner TR, Duke SD, Fransen BE, Reiter ML, Kroll AJ, Ward JW, Bach JL, Justice TE, Bilby RE (2010) Landslide densities associated with rainfall, stand age, and topography on forested landscapes, southwestern Washington, USA. For Ecol Manag 259(12):2233-2247. ISSN 0378-1127, https://doi.org/10.1016/j.foreco.2010.01.051

Wilson N, Broadbent A, Kerr J (2023) Cyclone Gabrielle by the numbers – a review at six months. Public Health Expert Brief. https://www.phcc.org.nz/briefing/cyclone-gabrielle-numbers-review-six-months. Accessed 21 Jun 2025

**Publisher's Nwote** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10346-025-02591-y.

**Chris Massey** (✉) · **Kerry Leith** · **Biljana Lukovic** · **Sam McColl** · **Brenda Rosser** · **Rob Buxton** · **Janine Bidmead**
GNS Science, 1 Fairway Drive, Avalon 5011, New Zealand
Email: c.massey@gns.cri.nz

**Tom R Robinson**
School of Earth and Environment, University of Canterbury, Christchurch 8140, New Zealand

**Trevor Carey-Smith**
NIWA, 301 Evans Bay Parade, Hataitai, Wellington 6021, New Zealand

**Liam Wotherspoon**
University of Auckland, Engineering Block 1, Bldg 401, 20 Symonds Street, Auckland 1010, New Zealand

**Hugh Smith** · **Harley Betts**
Manaaki Whenua - Landcare Research, Massey University, Riddet Road, Palmerston North, New Zealand