# Spatial Statistical Modeling of Rockfall Hazard in a Mountainous Road in Cantabria (Spain)

Daniel Jato-Espino, Alejandro Roldan-Valcarce ⓘ, Felipe Collazos-Arias, and Jorge Rodriguez-Hernandez

*Abstract*—**Rockfall events are one of the most frequent types of mass wasting in mountainous areas, causing service and traffic disruption, as well as infrastructure and human damage. Hence, having accurate tools to model these hazards becomes crucial to prevent fatalities, especially in a context of climate change whereby the effects of these phenomena might be exacerbated. Under this premise, this article concerned the development of a framework for assessing rockfall hazard in mountainous areas. First, a set of factors expected to favor rockfalls were processed and aggregated using spatial analysis tools, yielding a series of hazard maps with which to fit observed data through statistical modeling. The validation process was undertaken with the support of a database containing the number of rocks removed from a mountainous road section located in Cantabria, northern Spain. The results achieved, which demonstrated the accuracy of the proposed approach to reproduce rockfall hazard using frequency data, highlighted the primary role played by factors such as slope, runoff threshold, and precipitation to explain the occurrence of these events. The effects of climate change were considerably influenced by the fluctuations in the projections of precipitation, which limited the variations in the spatial distribution and magnitude of rockfall hazard.**

*Index Terms*—**Climate change, hazard mapping, rockfall, spatial analysis, statistical modeling.**

## I. INTRODUCTION

**T**OGETHER with flooding, mass wasting is the most frequent and fatal natural disaster, causing major economic and human losses worldwide [1], [2]. Haque *et al.* [3] undertook a comprehensive review of the 476 fatal mass movements occurring in Europe during the last 20 years (1995–2014), revealing a total of 1370 deaths and 784 injuries, with an annual average

Daniel Jato-Espino is with the Universidad Internacional de Valencia, 46002 Valencia, Spain (e-mail: djato@universidadviu.com).

Alejandro Roldan-Valcarce and Jorge Rodriguez-Hernandez are with the GITECO Research Group, Universidad de Cantabria, 39005 Santander, Spain (e-mail: roldanva@unican.es; jorge.rodriguez@unican.es).

Felipe Collazos-Arias is with the State Road Demarcation in Cantabria, Spanish Ministry of Transport, Mobility and Urban Agenda, 39010 Santander, Spain (e-mail: felipe.collazos@unican.es).

economic loss of 4–7 B€. Mountainous zones are particularly prone to these phenomena, especially in the form of rock slides or rockfalls, due to the sensitive morphologic characteristics of these areas [4], [5].

An increasing trend in the number of fatalities derived from these processes was identified from 2008 to 2014, which might be also related to the effects of climate change, since heavy rainfall was identified as the main factor triggering the occurrence of mass movements [3]. However, there is a variety of other topographic, geologic, and hydrologic factors contributing to generating slope instabilities [6]. In consequence, the management and processing of data to characterize these factors is crucial to properly evaluate and model the hazards represented by these phenomena [7].

As a result of the previous considerations, the assessment of mass wasting is an increasingly widely addressed topic in the literature [8]–[10]. Since their spatial distribution is argued to be crucial in prioritizing areas showing greater susceptibility to experience these phenomena [11], Geographic Information Systems (GIS) has been widely used to map these phenomena throughout the years.

Creating hazard maps is especially interesting for preventing rockfalls, which are particularly sudden mass movements [12] that require ex-ante prediction models. This circumstance has motivated the development of a number of studies devoted to the spatial modeling of rockfall hazards throughout the last years [13].

The most recent trends in this area focus on addressing rockfall geohazard through artificial intelligence. This includes artificial neural networks [14], ensemble models [15], learning classifiers [16], Gaussian mixture models [17], which are applied to determine rockfall probability based on a series of conditioning factors. Although these techniques are very powerful in computational terms, they are often used as black boxes and calibrated indiscriminately to fit observed data [18].

This lack of understanding of what happens "inside" the models notably hinders rationalizing the relationships between contributing factors and rockfall hazard. This situation is especially undesirable if it applies to high-stakes decisions as those concerning fatal hazards such as rockfalls. In line with recent discussions in the field of machine intelligence [19], future efforts to model rockfall should be oriented to developing interpretable models.

Still, intelligibility may not be enough to produce useful models. To this end, it must be coupled with simplicity. In this sense, another important common characteristic found in
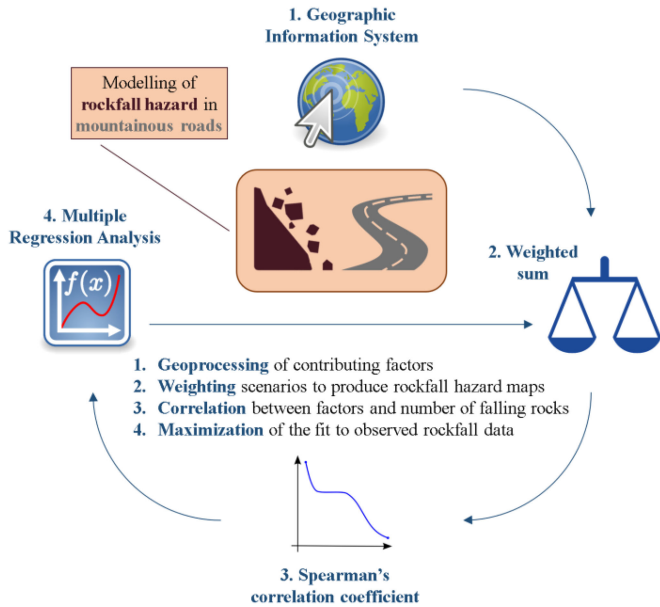
Fig. 1. Scheme of the approach taken for the spatial statistical modeling of rockfall hazard.
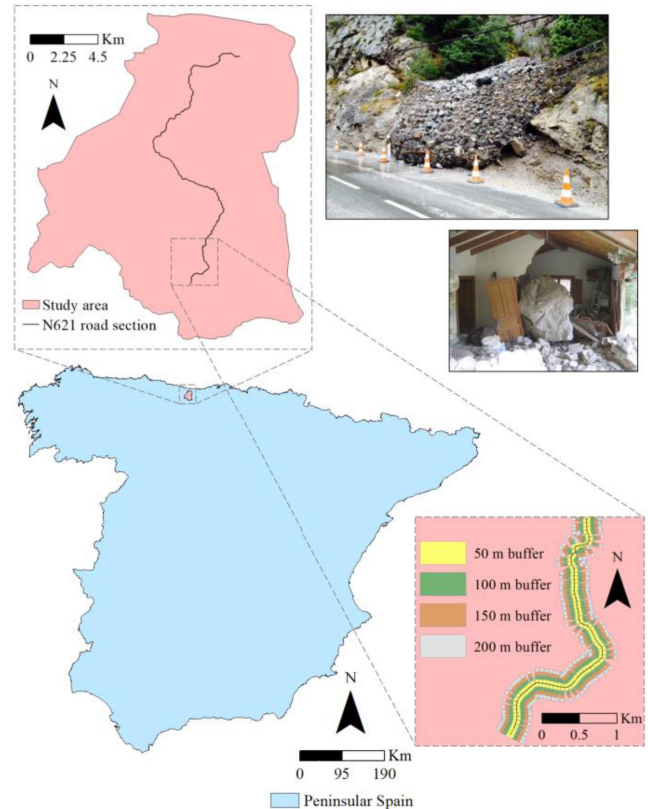


Fig. 2. Location of the study area and buffer distances considered to model the number of rockfall events reported along the monitored road section.

recent related literature is the use of 3-D numerical models and simulations to account for the kinematics and trajectories of rocks [20]–[23]. However, these models are often complex, demanding in computational terms and difficult to reconcile with spatial tools such as GIS.

Therefore, these approaches might be against the traditional promotion of simple and parsimonious predictive models with the potential to be generalized over time and space [24]. This aspect is also related to data needs, whereby some methods require rather specific geomorphological information to enable their application. In the end, both situations hinder the replicability of the models in other areas, thereby endangering their ultimate purpose as support tools for better preventing rockfall events.

As a result of the considerations extracted from the review of recent related literature, a research gap was identified in the development of simple, accessible, and interpretable methods to model rockfall hazards in mountainous roads. Hence, the aim of this research was to develop a spatial methodology combining all these characteristics to model rockfall frequency based on its relationship to a series of contributing factors. To this end, the proposed approach relied on easy-to-understand statistical techniques and globally available data, in order to facilitate its implementation in GIS and, therefore, its adoption by public entities used to handle spatial information for better planning rockfall hazard. The application of the methodology, which was tested in a mountainous road in Cantabria (Spain), also included the simulation of the impacts of climate change on rockfall hazard.

## II. METHODOLOGY

The approach taken for modeling rockfall susceptibility is summarized in Fig. 1, which illustrates its four main steps. First

was the spatial analysis of the data required to characterize a series of factors expected to contribute to rockfall via geoprocessing tools, in order to produce a series of hazard maps using weighted sum techniques. Then, the relationship between the values included in these hazard maps and the number of rocks removed from the road section monitored was determined using the Spearman's correlation coefficient. The statistical modeling step ended by building a multiple regression analysis (MRA) to determine the combination of weighted factors that maximized the fit to the observed rockfall data. This included the recalculations required to project the results when considering the impact of climate change. More details about the proposed methods are given below, once an overview of the characteristics of the study area and the list of proposed factors contributing to rockfall hazard has been provided.

### A. Study Area

The study site was located along a section of the N-621 road, in the surroundings of The Hermida Gorge, which is situated in the frontier between the provinces of Cantabria and Asturias, in the north of Spain (see Fig. 2). This is the longest gorge in Spain, with a length of 20 km, and belongs to the European Natura 2000 network as a Zone of Special Protection for Birds [25]. It is crossed by the Deva River and belongs to the Massif of Ándara, in the limits of the Picos de Europa National
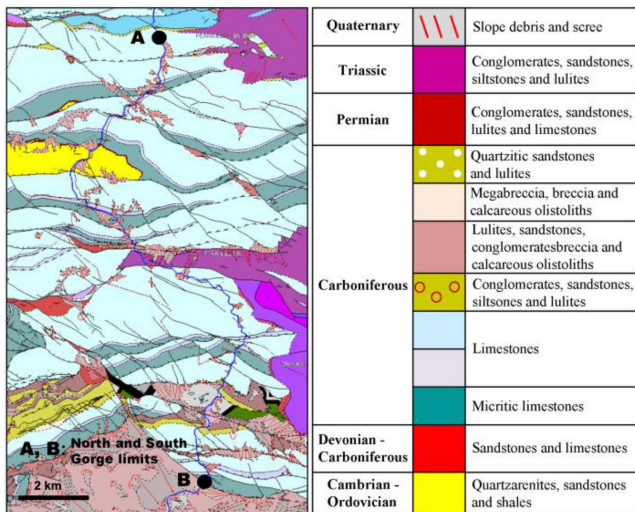
Fig. 3.   Geological map and stratigraphic column of The Hermida Gorge. Adapted from [30].

Park, the oldest (1918) and the third most visited National Park in Spain [26].

About 2500 vehicles per day circulate along the road section in The Hermida Gorge, a value that increases to 7500 in the summer season [27]. In contrast with the rest of Cantabria, which is subject to an Oceanic weather, the abrupt orography of this area favors a microclimate combining the Mediterranean, Oceanic and Alpine conditions. This results in values of annual mean precipitation of 1200 mm, with minimum and maximum mean temperatures of 8 and 28 °C, respectively [28].

The gorge is a combination of Variscan and Alpine thrusts that converge on a basal thrust in its southern edge. Its geological evolution took place under diagenetic conditions, with only a few zones experiencing a low degree of metamorphism [29]. The study area encompasses a series of stratigraphic levels as illustrated in Fig. 3. There is a succession of Palaeozoic rocks with ages ranging from the Lower Cambrian (550 myr) to the Permian-Triassic (250 myr), with an important stratigraphic lagoon that extends from the Lower Ordovician to the Upper Devonian [29].

On the top of these formations are Quaternary deposits generated by geomorphological agents on the substrate, on the top of which they are located discordantly [29]. Also, the mechanical weathering of steep slopes results in scree formed by the accumulation of free-falling blocks. When rockfall decreases, these deposits become inactive and are occupied by vegetation and soils. In addition to these units, the surface also includes anthropic deposits caused by backfilling tasks for construction purposes [29].

The study area is a fluvio-karstic gorge whose landform is the result of the combination of fluvial and fluvial-torrential incision processes, as well as of karstification and gravitational processes related to steep slopes. Its 20 km in length includes points of the unevenness of up to 1000 m, which favors the existence of a wide range of processes, some inherited from the past and some others still active. Thus, the geomorphological interest of The Hermida

Gorge is outstanding, including elements such as karst cavities, arches, and other landforms associated with slope debris, tuff formations, and ancient fluvio-torrential deposits [29].

Because of these particularities, numerous mass movements have been reported across the study area for more than a century, especially in the form of rockfall events [31]. Although the passage of time and the technological developments associated with it have helped improve both the preventive and reactive control over these instabilities, rockfall incidents are still recurrent in The Hermida Gorge [32]. Not in vain, the Spanish Ministry of Public Works has recently approved an 85 M€ project to improve the layout of the N-621 road [33].

As a precedent to this project, the State Road Demarcation in Cantabria has been monitoring the number of rocks removed from the road in The Hermida Gorge over the last years. This inventory collects data every 50 m from the N-621 road section between the kilometer points (KP) 154 and 174. In particular, the reports are structured as tables formed of the following fields: "KP," "Date," "Description," "Action," and "Number of rocks."

"Description" was a combination of the two last fields, since it indicates the action taken (remove or sweep) to deal with the falling rocks. In some cases, it also specified the size or weight of the rocks reported (e.g., "Remove 5 rocks of 20 cm" or "Remove 2 rocks of 2 kg"). However, this information was uneven throughout the inventory, which hinders its utilization for modeling purposes. In consequence, rockfall hazard was approached using frequency (number of rocks) as an indicator to represent whether this condition existed or not.

The data extracted from this inventory corresponded to 2007. This period was selected because it represented the least altered scenario by external mitigation measures installed afterward such as protection barriers, whose presence at specific sites might affect the reliability of the results to be achieved.

Overall, the number of rocks retired from the road during 2007 amounted to 14 717. This figure was unevenly distributed in temporal terms, yielding monthly average and standard deviation (STD) values of 1226.42 and 360.36 rocks, respectively. Almost 40% of these phenomena occurred between February and May, coinciding with the melting season in the study area. Regarding the location of the events reported, the data collected proved that almost 50% of them took place within the last 5 km of the section monitored, where the difference of altitude between the road and the surrounding mountains is particularly remarkable.

Instead of using complex numerical models to represent rock kinematics and trajectories, buffer areas were established at 200 m on each side of the monitored road section (see Fig. 2) to account for the source of the instabilities that caused the fall of rocks at every KP monitored [34]. Furthermore, to better represent the progressive influence of the slopes on the number of rocks removed from the road, these buffer areas were divided into four portions at 50, 100, 150, and 200 m from the road [35], as represented in Fig. 2. Hence, rock detachments from closer areas were more likely to end falling to the road, whereas the trajectory of further instabilities may be interrupted by the shape and relief of slopes before reaching the surface.

TABLE I
DESCRIPTION OF THE DATASETS COLLECTED AND FACTORS TO PRODUCE
THROUGH THEIR GEOPROCESSING

| Data | Format | Resolution | Factor | Units |
|---|---|---|---|---|
| Digital Elevation Model (DEM) | Raster | 5 m | Slope ($f_1$) | ° |
| Lithostratigraphic map | Vector | 1:200,000 | Lithology ($f_2$) | Score |
| Land cover map | Vector | 1:100,000 | Runoff threshold ($f_3$) | mm |
| Normalized Difference Vegetation Index (NDVI) | Raster | 30 m | NDVI ($f_4$) | [-1, 1] |
| Monthly precipitation | Raster | 0.93 km | Precipitation ($f_5$) | mm |
| Monthly maximum and minimum temperature | Raster | 0.93 km | Thermal amplitude ($f_6$) | °C |

## B. Description of the Factors Contributing to Rockfall Hazard

To boost the replicability of the proposed approach, the data required to model rockfall hazards in the monitored road section were selected to result in a collection of information widely available worldwide. Although elevation, lithological, and land cover data were collected from regional and national sources for the sake of precision, they can also be acquired using global repositories [36]–[38]. Table I provides a description of these data, their format, and resolution, as well as the rockfall-related factors to which they led and their corresponding units.

Apart from their potential to be replicated elsewhere, this set of factors was also selected for their suitability to capture both intrinsic (topography, geology, and hydrology) and extrinsic (climate and land cover changes) [39], [40] aspects, being similar to those used in previous related studies [41]–[43]. Other variables such as slope aspect and wind speed were not considered for this case study. The former is difficult to normalize and aggregate with other factors, whilst the values associated with the latter are low in the study area, ranging from 20 to 28 km/h during the period under analysis.

The first input was the digital elevation model (DEM), which was acquired from the Spanish Geographic Institute [44]. This variable might be used straightforward as a potential contributing factor to rockfall; however, the relationship between elevation and landslide hazard is not clear [45]. On the contrary, slope ($f_1$) has been found to be one of the most relevant factors for estimating these phenomena [46], since steeper slopes commonly result in higher stresses in the terrain [6]. Other DEM-derived variables, such as slope aspect or curvature, have been reported to provide a weak predictive ability [47] and were therefore discarded henceforth.

Lithology ($f_2$) stemmed from the lithostratigraphic map available at the spatial data infrastructure of the University of Extremadura [48]. The lithology and degree of alteration of the terrain influence its physical and mechanical properties, which eventually affect its potential stability [49]. Consequently, this factor was scored according to the mineralogical composition, texture, compaction, size, shape, and cementation of the particles forming the rocks or sediments [50]. Hence, materials with a low degree of cementation, fine grain size (silty clay or silty clayey sand), or wide gradation range are more likely to be instable. Moreover, lithological units with high contents of clay, especially in the form of montmorillonite, are very sensitive to water absorption, favoring the expansion of the terrain. The washing of salts contained in certain clay types produces a particle structure rearrangement that reduces their shear strength. In addition, the alternation of materials with different resistance, compaction, and permeability causes heterogeneities in the global behavior of the resulting mass, which also facilitates the existence of instabilities. Under these considerations, the scoring framework defined ranged from 0.1 (conglomerate) to 0.9 (clay), including intermediate values such as 0.2 (limestone), 0.4 (gravel), or 0.7 (blocks and boulders).

The map produced by the CORINE land cover project [51] was used to represent the runoff threshold of the terrain surface in the study area ($f_3$) according to the values approved by the Spanish Official State Gazette [52]. This factor was proportional to rockfall hazard, since high permeability facilitates rainfall infiltration, which endangers intergranular friction and cohesion and alters the suction and shear strength of the soil [53].

The normalized difference vegetation index (NDVI) ($f_4$) was used to represent the amount, quality, and development of vegetation in the study area [54]. The higher the value of NDVI, the greater both the protection of the terrain against erosion [55] and the growth of roots increasing the cohesion of the terrain [56]. This factor was determined as the quotient between the difference and the sum of the fourth (visible red) and fifth (near infrared) bands of the Landsat 8 satellite [57].

The last two variables were trigger factors symbolizing the meteorological conditions of the study area: precipitation ($f_5$) and thermal amplitude ($f_6$). On the one hand, precipitation can provoke water infiltration processes in the soil and modify the morphology of the terrain due to its erosive potential [58]. On the other hand, extreme maximum and minimum temperatures contribute to exacerbating freeze-thaw cycles, which in turn favors saturation and swelling processes [59] and increase the frequency of mass movements [3]. Both factors were determined from February to May, which represented the period with the most unfavorable conditions in these terms. The data required to compute these factors stemmed from the stationary climate maps of version 1.4 of WorldClim [60], which contain long-term average monthly precipitation and temperature data.

To further characterize the interaction between these meteorological factors and rockfall hazard, future variations in precipitation and temperature were explored by considering the effect of climate change through the projections made by the Intergovernmental Panel on Climate Change (IPCC) [61]. In particular, the downscaling and calibration of the outputs corresponding to three General Circulation Models (GCM) derived from this report, namely CNRM-CM5, HadGEM2-CC, and MIROC5, were considered for comparative purposes using the WorldClim 1.4 data for the years 2050 (average for 2041–2060) and 2070 (average for 2061–2080) [60]. Furthermore, variations in precipitation and maximum and minimum temperature were

TABLE II
GEOPROCESSING TOOLS USED TO GENERATE ROCKFALL HAZARD MAPS FROM A SERIES OF CONTRIBUTING FACTORS

| Output | Tools |
|---|---|
| Slope | Mosaic To New Raster → Extract By Mask → Slope |
| Lithology | Clip → Add Field → Select By Attributes → Calculate Field → Polygon To Raster |
| Runoff threshold | Clip → Raster Calculator → Raster To Polygon → Merge → Add Field → Select By Attributes → Calculate Field → Polygon To Raster |
| NDVI | Project Raster → Extract By Mask → Raster Calculator |
| Precipitation | Project Raster → Extract By Mask → Raster Calculator |
| Thermal insulation | Project Raster → Extract By Mask → Raster Calculator |
| Rockfall hazard map | Raster Calculator → Weighted Sum |
| Buffers' statistics | Zonal Statistics as Table |

TABLE III
WEIGHTING SCENARIOS PROPOSED TO FIT THE NUMBER OF ROCKFALL EVENTS REPORTED ALONG THE MONITORED ROAD SECTION

| Scenario | $w_{f_1}$ | $w_{f_2}$ | $w_{f_3}$ | $w_{f_4}$ | $w_{f_5}$ | $w_{f_6}$ |
|---|---|---|---|---|---|---|
| S1 | 0.167 | 0.167 | 0.167 | 0.167 | 0.167 | 0.167 |
| S2 | 0.500 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 |
| S3 | 0.100 | 0.500 | 0.100 | 0.100 | 0.100 | 0.100 |
| S4 | 0.100 | 0.100 | 0.500 | 0.100 | 0.100 | 0.100 |
| S5 | 0.100 | 0.100 | 0.100 | 0.500 | 0.100 | 0.100 |
| S6 | 0.100 | 0.100 | 0.100 | 0.100 | 0.500 | 0.100 |
| S7 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.500 |
| S8 | 0.750 | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 |
| S9 | 0.050 | 0.750 | 0.050 | 0.050 | 0.050 | 0.050 |
| S10 | 0.050 | 0.050 | 0.750 | 0.050 | 0.050 | 0.050 |
| S11 | 0.050 | 0.050 | 0.050 | 0.750 | 0.050 | 0.050 |
| S12 | 0.050 | 0.050 | 0.050 | 0.050 | 0.750 | 0.050 |
| S13 | 0.050 | 0.050 | 0.050 | 0.050 | 0.050 | 0.750 |
| S14 | 0.400 | 0.400 | 0.050 | 0.050 | 0.050 | 0.050 |
| S15 | 0.050 | 0.050 | 0.400 | 0.400 | 0.050 | 0.050 |
| S16 | 0.050 | 0.050 | 0.050 | 0.050 | 0.400 | 0.400 |

modeled according to representative concentration pathways (RCPs) with radiating forcing levels of 4.5 (RCP4.5) and 8.5 (RCP8.5) W/m$^2$ [62].

### C. Spatial Analysis

The first part of the methodology was devoted to the spatial processing and management of the data listed in Table I. Table II compiles the set of ArcGIS 10.1 [63] geoprocessing tools used to generate the maps associated with these factors, as well as that of rockfall hazard stemming from their aggregation and the summary statistics obtained across the four buffer distances considered.

The slope output determined the gradient of the terrain from the DEM. The characterization of the lithologic factor involved the scoring of the lithostratigraphic units according to the materials present in the zone, such that different soil groups were created and rated based on their characteristics. The determination of the runoff threshold required the previous two factors, since the calculation of this variable depends on the values of slope and soil permeability of the underlying soil associated with each land cover type in the study area [52].

With slight variations, the processing of the last three factors was the same. Unlike previous variables, all the data related to the NDVI and climate variables were produced on a worldwide scale, such that their reference system was WGS84. Consequently, these source data were transformed to ETRS89, which is the official reference system in Spain. Since these factors were provided in coarser resolutions than the DEM (see Table I), this projection step was also used to resample these data to a cell size of 5 m.

Then, these three factors were calculated in the study area. In the case of the NDVI, this task involved the visible red and near infrared bands, whilst precipitation was determined as the average of the values in the study area from February to May 2007. Thermal amplitude was obtained analogously but including an extra step to compute the difference between the maximum and minimum temperature.

Once all the factors $f_j$ were delineated, their values $v_j$ were normalized ($n_j$) by applying (1) using the "Raster Calculator"

tool. In addition, (2) was employed as a previous step in the case of the NDVI, since this parameter can be negative [64] (see Table I)

$$
n_j = \begin{cases} \dfrac{v_j}{\max\limits_j v_j}, \\ \text{if } f_j \text{ is directly proportional to rockfall hazard} \\ \dfrac{\min\limits_j v_j}{v_j}, \\ \text{if } f_j \text{ is inversely proportional to rockfall hazard} \end{cases}
\tag{1}
$$

$$
\bar{v}_j = v_j + \left| \min_j v_j \right| + 1.
\tag{2}
$$

The normalized factors were then aggregated, enabling their weighted combination to obtain a global score indicating rockfall hazard [65]. Consistent with the second part of the methodology, a series of weighting scenarios as shown in Table III was proposed to facilitate the statistical fitting of the rockfall events observed in the road section monitored. These combinations were designed to result in 1 balanced scenario, 14 scenarios oriented to prioritize one of the variables to different extents, and 3 scenarios in which two factors predominated over the others. The latter were conducted to highlight the importance of inherent (S14), sensitive to anthropogenic changes (S15), and trigger (S16) factors, respectively.

The maps generated according to these combinations of weights were used to extract the values of rockfall hazard in the buffer areas as indicated in Table II. The summary statistics computed in these sites included the minimum, maximum, range, mean, STD, and sum of the pixels enclosed by the buffer areas at distances of 50, 100, 150, and 200 m (see Fig. 2).

## D. Statistical Modeling

The second part of the methodology focused on the statistical modeling of the rockfall events reported and the hazard maps produced through geoprocessing. Unlike most previous related studies, which applied binary logistic regression analysis to model rockfall hazard due to the nature of the datasets available for validation (existence or absence of events), the inventory available in this investigation was in ordinal format (number of rocks removed from the road). Hence, the Spearman's rho ($\rho$) was used as expressed in (3) to measure the strength of the correlation between the number of falling rocks and hazard maps

$$\rho = 1 - \frac{6 \sum D^2}{N\left(N^2 - 1\right)} \qquad (3)$$

where $D$ is the difference between the ranks of the variables and $N$ is the number of observations. To transition to a more realistic spatial scale, the data corresponding to the segments into which the road section was divided in the monitoring campaign (50 m) were aggregated per KP, resulting in a value of $N = 20$ (from KP 154 to KP 174). Consequently, the scores obtained across the buffer areas considered were combined per KP too, weighting them according to their proximity to the road. Hence, the buffer zones of 50, 100, 150, and 200 m were multiplied by 0.4, 0.3, 0.2, and 0.1, respectively, in order to represent how the influence of the surrounding areas on rockfall hazard decreased proportionally with the distance from the road [35].

Since both the values of Spearman's rho and the combinations of weights proposed in Table III consisted of continuous data, the strength of their relationships was measured through MRA according to a significance level of 0.05 [66]. Thus, the Spearman's rho performed as the response ($r$) and the weights as the predictors ($p_i$) in the MRA model, which aimed at estimating the relationships between all the $t$ predictors involved. In addition to linear terms, first-order interactions were also included in the model as formulated in (4), since some of the variables considered in Table I were expected to produce combined effects

$$r = C_0 + \sum_{i=1}^{t} C_i p_i + \sum_{i=1}^{t} \sum_{k=1}^{t} C_{ik} p_i p_k + e \qquad (4)$$

such that $C_0$ is the constant, $C_i$ are the coefficients by which the predictors $p_i$ are multiplied and $e$ is the error of the model. The reliability of MRA was validated through a residual analysis, whereby the assumptions of normality, homoscedasticity, multicollinearity, and independence were verified as described in Table IV. Normality and homoscedasticity were evaluated through the $p$-values of the Ryan-Joiner [67] and Levene's tests [68], respectively. Multicollinearity was detected based on the value inflation factor (VIF) reached by the predictors according to a threshold of VIF $< 10$ [69]. Finally, the independence of residuals was corroborated using the significance tables presented in Savin and White [70], which required that the Durbin-Watson statistic was between the lower (dL) and upper bounds (dU) corresponding to the number of observations ($N$) and predictors ($t$) in the MRA model.

Moreover, the accuracy of MRA was measured through the joint consideration of the standard error of the regression ($S$) and different forms of the coefficient of determination: standard

### TABLE IV
REQUIREMENTS AND VERIFICATION PROCEDURES USED TO CHECK THE ASSUMPTIONS OF MRA

| Assumption | Requirement | Verification |
|---|---|---|
| Normality | The residuals must be normally distributed | p-value > 0.05 in the Ryan-Joiner test |
| Homoscedasticity | The error variance is similar across all the values of the predictors | p-value > 0.05 in the Levene's test |
| Multicollinearity | The predictors must not be highly correlated to each other | Value Inflation Factor (VIF) < 10 |
| Independence | The residuals must be independent to each other | dL < Durbin-Watson statistic < dU |

($R^2$), adjusted ($R^2_{adj.}$), and predicted ($R^2_{pr.}$). The $R^2_{adj.}$ coefficient adjusts the power of MRA to the number of predictors used, whilst $R^2_{pr.}$ consists of calculating the regression equation as each observation is removed from the dataset systematically, in order to validate the capacity of the model for making new predictions.

Since the eventual goal sought was the identification of the combination of weights that maximized the values of the Spearman's rho, the $R^2_{pr.}$ coefficient was given preference over the other goodness-of-fit measures due to its extrapolation potential. The weights were determined according to the contributions of the terms to the MLR model resulting in the highest $R^2_{pr.}$. Hence, the optimal weight $\bar{w}_{f_j}$ of a factor $f_j$ was computed as its contribution as a linear term ($c_{lin}$) and the sum of half of its contributions in the $q$ interaction terms ($c_{int_m}$), as described in the following equation:

$$\bar{w}_{f_j} = \frac{c_{lin} + \frac{1}{2} \sum_{m=1}^{q} c_{int_m}}{\sum_{i=1}^{t} \left(c_{lin} + \frac{1}{2} \sum_{m=1}^{q} c_{int_m}\right)} . \qquad (5)$$

Furthermore, the geoprocessing tasks summarized in Table II were replicated using these optimal weights and the values of precipitation and thermal amplitude under nonstationary conditions, in order to produce hazard maps indicating the impact of Climate Change on rockfall for the years 2050 and 2070 under the RCP4.5 and RCP8.5 trajectories.

## III. RESULTS AND DISCUSSION

This section presents and analyses the results produced through the spatial statistical modeling of rockfall hazards in The Hermida Gorge. The outputs derived from this case study are structured according to the two methodological steps into which the proposed approach was divided: spatial analysis and statistical modeling.

### A. Spatial Analysis

The geoprocessing of the contributing factors included in Table I yielded the maps represented in Fig. 4. Fig. 4(a) illustrates the slope conditions of the study area, which ranged from $0°$ to $86.98°$, with an average value of $27.89°$. This is equivalent
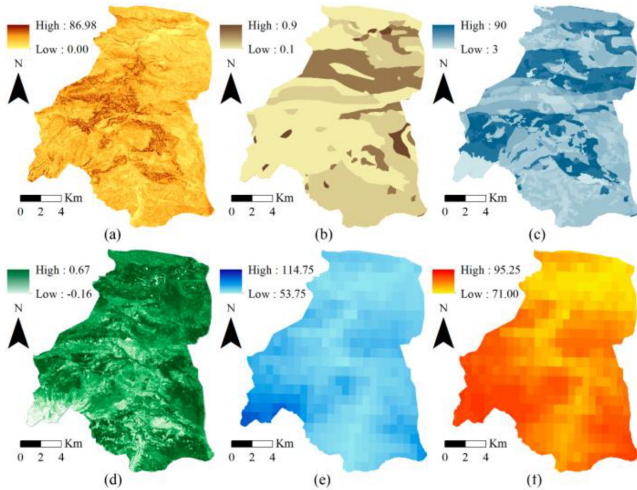
Fig. 4. Rockfall hazard-related factors in the study area. (a) Slope (°). (b) Lithology (Score). (c) Runoff threshold (mm). (d) NDVI (Score). (e) Precipitation (mm). (f) Thermal amplitude (°C∗10).

to approximately 53%, which highlights the abruptness of this mountainous area, characterized by having very steep slopes and even vertical walls in some localized spots.

The lithological map depicted in Fig. 4(b) revealed a predominance of limestone in the surroundings of The Hermida Gorge, with a notable presence of sandstone and combinations of gravel, sand, and silt too. The lithological units in the map were rated from 0.1 to 0.9 according to their contribution to producing instabilities based on their composition, texture, compaction, shape, and cementation and the classifications suggested in previous studies in some municipalities close to the study area [71], [72].

Lithology was also considered for the determination of the runoff threshold of the study area [see Fig. 4(c)], since this variable depended on both the permeability of the underlying soil and the slope of the terrain. As a result, the combination of materials such as clays or slates with steep slopes and surfaces with low percolation potential led to low values of runoff threshold and, consequently, low infiltration to the ground. The NDVI map shown in Fig. 4(d) was also related to the surface of the study area, such that those zones presenting the highest values of NDVI provided the greatest protection against erosion phenomena.

Finally, the maps corresponding to the two meteorological variables, represented in Fig. 4(e) and (f), exhibited a very similar pattern. The largest values of precipitation and thermal amplitude were located in the most elevated areas, where precipitation can reach up to 2000 mm and average minimum temperatures decrease to 4 °C approximately. On the contrary, the lowest values for both variables corresponded to valleys, including the vicinity of the road section monitored, in which weather extremes were smoother.

To address the evolution of rockfall hazard in the future due to climate change, precipitation and thermal amplitude maps were also prepared under nonstationary conditions, leading to the results represented in Figs. 5 and 6. These climate change projections corresponded to a series of combinations of RCP4.5

and RCP8.5, year (2050 and 2070), and GCM (CNRM-CM5, HadGEM2-CC, and MIROC5). The spatial patterns of the maps remained constant for both factors, such that the highest elevation areas kept being associated with the worst values of precipitation and thermal amplitude.

The slight differences between the stationary maps in Fig. 4 and those included in Figs. 5 and 6 laid on the ranges of values achieved for each scenario. In the case of precipitation, the scenarios involving an increase in rainfall concerned the CNRM-CM5 models, regardless of the year and RCP considered. The greatest increase in precipitation corresponded to the year 2070 and the RCP4.5 scenario, which represents medium-low greenhouse gases concentration. This is compatible with the Fifth Assessment Report (AR5) of the IPCC, whereby variations in cumulative precipitation are space-dependent and not necessarily proportional to GHC concentration levels, unlike extreme rainfall events [73].

As for thermal amplitude, the maps in Fig. 6 coincided with projecting an increase in the values taken by this variable under all scenarios, as a result of an intensification of maximum temperature due to climate change. These variations are in line with the AR5 of the IPCC, according to which more frequent and longer hot temperature extremes are expected in the future, particularly in the RCP8.5 scenario due to a projected rise in mean temperature from 2.6 to 4.8 °C [73]. This is consistent with the results in Fig. 6, since the greatest variations in thermal amplitude were associated with the year 2070 and the RCP8.5 scenario, especially for the MIROC5 model.

The maps produced to represent the factors contributing to rockfall hazard (see Fig. 4) were normalized using (1). In the case of NDVI, (2) was previously applied to deal with its negative values. In addition, precipitation and thermal amplitude were normalized according to both the stationary and climate change conditions, such that the values of $\max_i v_j$ in (1) corresponded to the maps depicted in Fig. 5(d) (114.75 mm) and Fig. 6(l) (95.25 °C∗10). This course of action enabled accounting for the impact of the variations between both situations (stationarity and climate change) on rockfall hazard.

Hence, the normalized versions of the maps shown in Fig. 4 were aggregated according to the weighting combinations proposed (see Table III), yielding the results depicted in Fig. 7. The predominance of some factors resulted in heterogeneous maps in which the values ranged from low to high hazard, especially those related to $f_3$ [see Fig. 7(d) and (j)] because of the variations in the filtration capacity of different land cover types. Instead, other maps such as Fig. 7(i) and (m) were rather homogeneous, either by including mainly low or high values, due to the distribution of $f_2$ and $f_6$ in relation to the remaining factors. Overall, the maps in Fig. 7 represented extreme hazard situations. Therefore, their subsequent interpolation through statistical modeling was aimed at maximizing the fit to the rockfall data reported in the study area.

### B. Statistical Modeling

The application of (3) yielded the results illustrated in Fig. 8, which represents the Spearman's correlation between the summary statistics of the hazard maps obtained from the weighting scenarios in Table III and the number of rocks removed from the
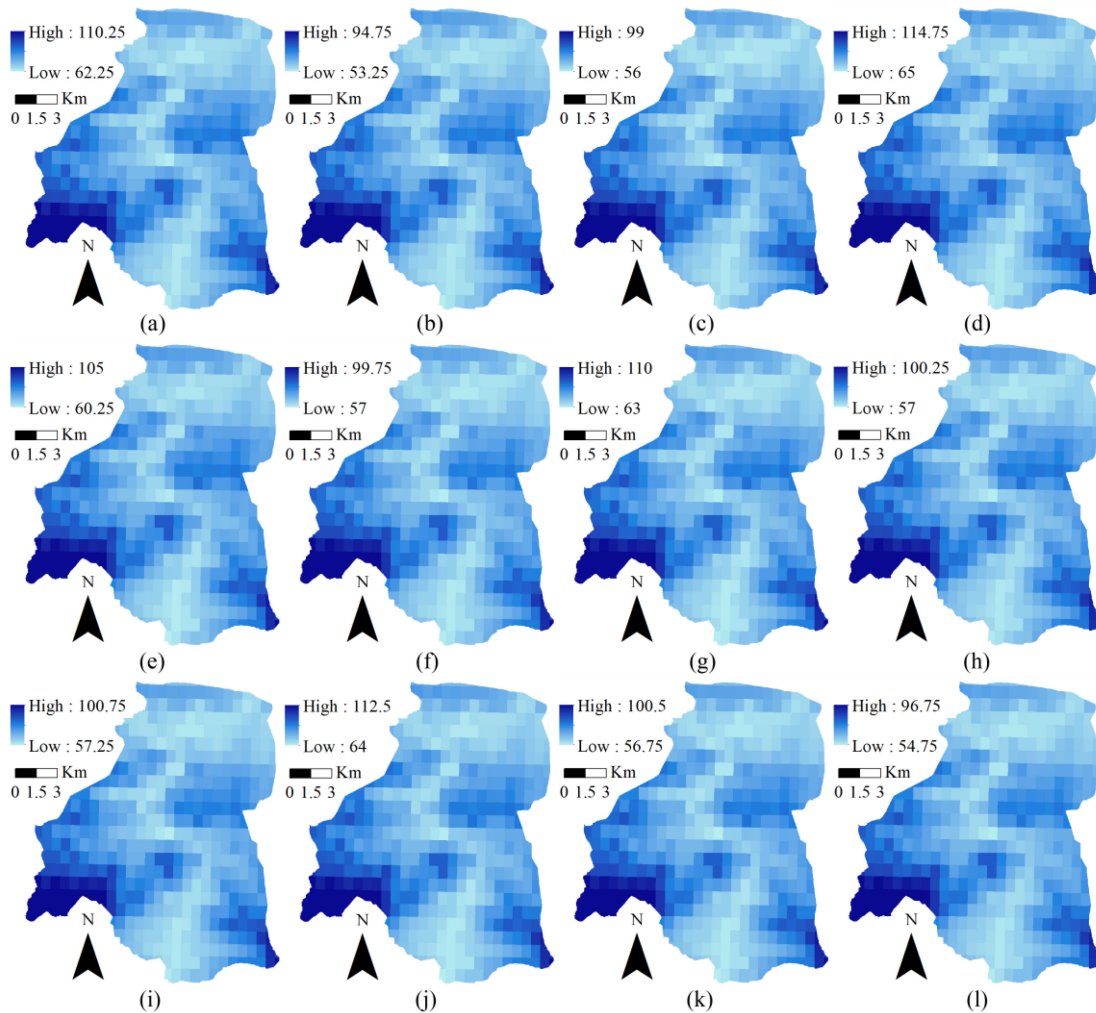
Fig. 5. Projections of precipitation (mm) obtained for different RCP, years and models. (a) RCP4.5; 2050; CNRM-CM5. (b) RCP4.5; 2050; HadGEM2-CC. (c) RCP4.5; 2050; MIROC5. (d) RCP4.5; 2070; CNRM-CM5. (e) RCP4.5; 2070; HadGEM2-CC. (f) RCP4.5; 2070; MIROC5. (g) RCP8.5; 2050; CNRM-CM5. (h) RCP8.5; 2050; HadGEM2-CC. (i) RCP8.5; 2050; MIROC5. (j) RCP8.5; 2070; CNRM-CM5. (k) RCP8.5; 2070; HadGEM2-CC. (l) RCP8.5; 2070; MIROC5.

monitored road section per KP. The values of range and STD per KP provided the worst fit to the observed data, with none of the correlation coefficients obtained proving to be statistically significant ($p$-values $< 0.05$). In contrast, the minimum (MIN), maximum (MAX), and mean yielded significant correlation coefficients under all weighting combinations, except for those scenarios in which lithology ($f_2$), NDVI ($f_4$), and thermal amplitude ($f_6$) were the predominant factors (S9, S11, and S13).

In consequence, the values of Spearman's rho suggested that $f_2$, $f_4$, and $f_6$ were the variables with less influence on rockfall hazard. On the contrary, in light of the results corresponding to S8, S10, and S12, slope ($f_1$), runoff threshold ($f_3$), and precipitation ($f_5$) were the most relevant factors to explain the number of rocks reported. In fact, the correlation coefficient reached in S8, which corresponded to the scenario where the slope predominated, was about 0.8 when using MIN (see Fig. 8). The strength of the relationship concerning this statistic was logical, since high minimum hazard values across a KP indicated a majority of areas contributing to rockfall events.

Based on these considerations, MRA was carried out using the statistically significant values of Spearman's rho as a response ($r$) and their corresponding MIN-based scenarios as predictors ($p_i$), in order to determine the combination of weights that best-fitted rockfall frequency. According to Fig. 8, the number of valid observations was 13, since the scenarios with a $p$-value $< 0.05$ were excluded. Table V summarizes the main features of the MRA built, which involved 3 linear terms and 5 interactions. The performance of the model was evaluated through the standard error of the Regression ($S$) and the $R^2$ coefficient, whose values highlighted the accuracy of the results, such that 99.7% of the variations in the Spearman's rho were explained by the weights of the factors. Furthermore, the value of $R^2_{adj.}$ corroborated the adequacy of the number of predictors included in the model, whilst $R^2_{pr.}$ guaranteed its predictive capability.

The validity of the MRA model was ensured by the outputs derived from the residual analysis. The results of the Ryan-Joiner test certified the normality of the residuals, since its $p$-value was above the significance level (0.05). The application of the
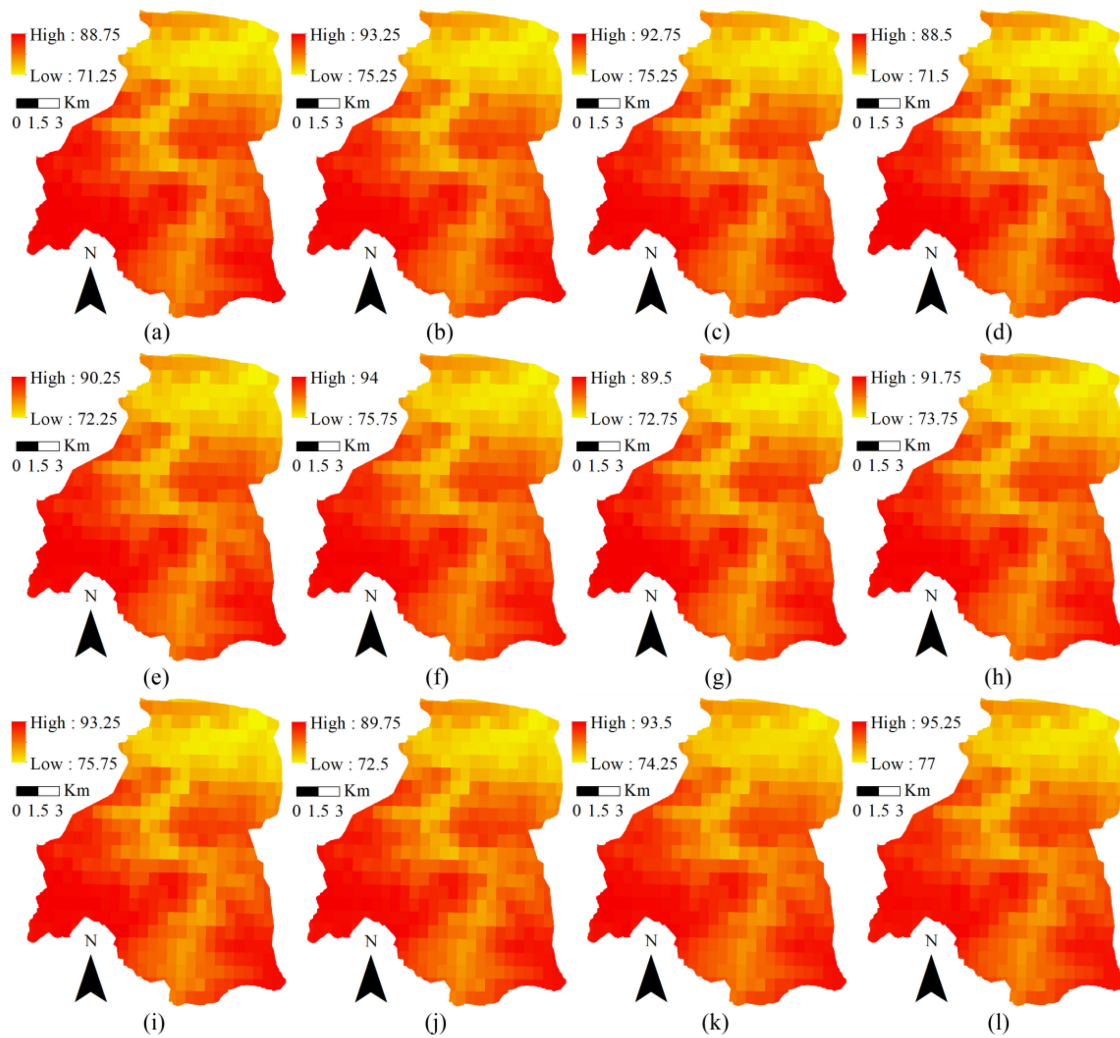
Fig. 6. Projections of thermal amplitude (°C∗10) obtained for different RCP, years and models. (a) RCP4.5; 2050; CNRM-CM5. (b) RCP4.5; 2050; HadGEM2-CC. (c) RCP4.5; 2050; MIROC5. (d) RCP4.5; 2070; CNRM-CM5. (e) RCP4.5; 2070; HadGEM2-CC. (f) RCP4.5; 2070; MIROC5. (g) RCP8.5; 2050; CNRM-CM5. (h) RCP8.5; 2050; HadGEM2-CC. (i) RCP8.5; 2050; MIROC5. (j) RCP8.5; 2070; CNRM-CM5. (k) RCP8.5; 2070; HadGEM2-CC. (l) RCP8.5; 2070; MIROC5.

Levene's test enabled verifying the assumption of homoscedasticity ($p$-value $> 0.05$), once the residuals were divided into three groups to compare their respective variances. The VIF was below 10 for all terms, which guaranteed that the predictors used were not significantly correlated to each other and, consequently, multicollinearity was not an issue. Finally, the value obtained for the Durbin–Watson statistic was between the lower (dL $= 0.090$) and upper bounds (dU $= 3.182$) associated with the number of observations (13) and predictors (8) used, which verified the independence of residuals.

The decimal representation of the contributions included in Table V was used to determine the set of optimal weights that best fitted the number of rocks removed from the study area. The application of (5) yielded the following weights for the factors involved in the MRA model: 0.668 ($f_1$), 0.112 ($f_3$), 0.055 ($f_4$), 0.142 ($f_5$), and 0.023 ($f_6$). Consistent with the trend observed in Figs. 7 and 8, these weights highlighted the predominant role played by the action of gravity to favor rockfall because of the existence of steep slopes ($f_1$), which are vertical in several KPs in the study area. This is in line with several previous studies pointing out slope gradient as the main contributing factor to rockfall hazard in mountainous areas [74], especially if elevation exceeds 500 m above the sea level as in this case [75].

The next most important factors were rainwater percolation in terms of mechanical degradation ($f_3$) and weathering ($f_5$). As pointed out by Maleki [76], the presence of voids favors the seepage of water, which results in forces acting on the slopes that can eventually lead to rockfalls. The role of precipitation as a triggering factor is exacerbated when coupled with temperature fluctuations ($f_6$) [77], which was in the next degree of relevance as indicated in Table V, due to the addition of water from thawing that seeps into the soil [78].

Although to a lesser extent, the role played by vegetation ($f_4$) in terms of protection of the terrain was also found to be statistically significant for rockfall stability. This outcome coincides with the findings of Wang *et al.* [79], who emphasized the negative correlation between vegetation depth and rockfall frequency. Other authors have also put a focus on the reforestation of hillslopes via native vegetation as a solution to reduce future rockfalls [80].
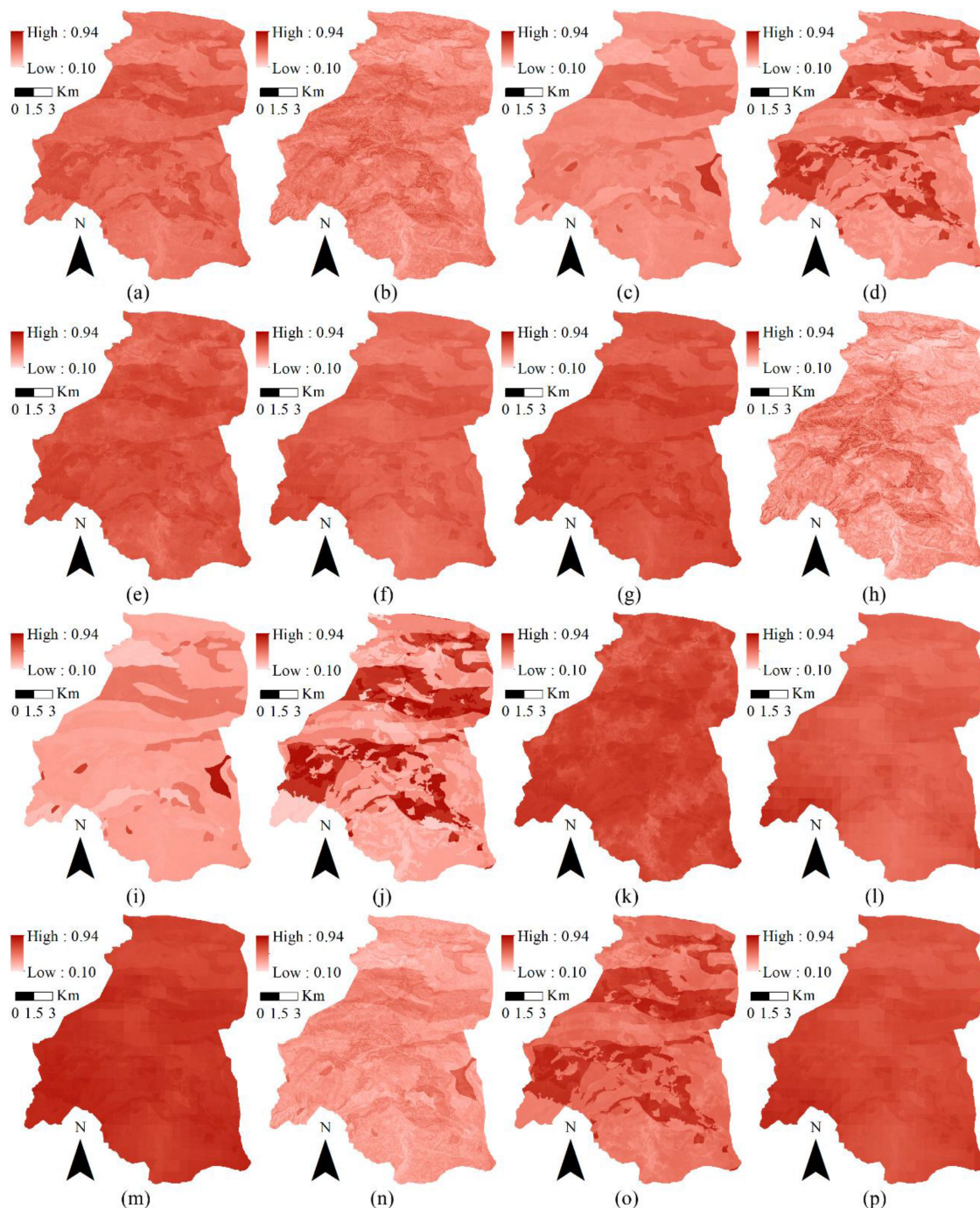
Fig. 7.     Rockfall hazard maps obtained through the aggregation of the factors according to the list of weighting scenarios proposed. (a) S1. (b) S2. (c) S3. (d) S4. (e) S5. (f) S6. (g) S7. (h) S8. (i) S9. (j) S10. (k) S11. (l) S12. (m) S13. (n) S14. (o) S15. (p) S16.

Lithology ($f_2$) was the only factor excluded from the model, probably due to the difficulties associated with the quantification of the physical properties of the rock units. This minor role played by lithology coincided with some models also built for Northern Spain, which highlighted that reasonably satisfactory susceptibility models might be achieved only from a good DEM and inventory data [81]. Still, this circumstance did not necessarily indicate that the lithological characteristics of the terrain were irrelevant for rockfall hazard. In fact, this variable was an

underlying factor in the calculation of the runoff threshold ($f_3$), as proven by the similarities between Fig. 4(b) and (c). Taking this into account, all factors can be concluded to contribute to rockfall hazard to a greater or lesser extent.

The reapplication of (3) using the optimal weights listed above yielded a Spearman's rho of 0.82 ($p$-value $< 0.05$), which improved the highest values achieved through the consideration of the initial combinations of weights (see Fig. 8). The aggregation of the significant factors according to these weights resulted in
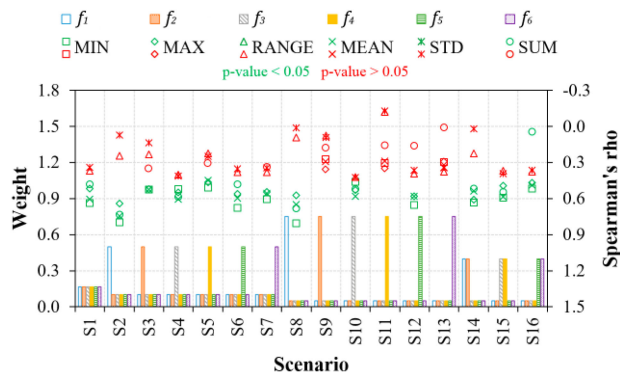
Fig. 8. Spearman's correlation coefficients reached between the hazard maps yielded by the weighting scenarios and the number of rocks reported along the monitored road section.

TABLE V
SUMMARY OF THE MULTIPLE REGRESSION MODEL BUILT FOR FITTING THE SPEARMAN'S CORRELATION COEFFICIENTS CALCULATED FROM THE NUMBER OF ROCKS REPORTED ALONG THE MONITORED ROAD SECTION

| Term | Contribution (%) | Coefficient | p-value | VIF |
|---|---|---|---|---|
| Regression | 99.73 | - | 0.000 | - |
| Constant | - | 0.443 | 0.000 | - |
| $f_1$ | 61.77 | 0.252 | 0.001 | 4.39* |
| $f_3$ | 6.48 | -0.077 | 0.021 | 2.73* |
| $f_5$ | 9.14 | 0.138 | 0.022 | 8.88* |
| $f_1 * f_4$ | 1.59 | -1.471 | 0.007 | 3.23* |
| $f_1 * f_5$ | 6.79 | 3.119 | 0.002 | 6.74* |
| $f_1 * f_6$ | 1.32 | 2.737 | 0.001 | 3.15* |
| $f_3 * f_4$ | 9.38 | 1.103 | 0.000 | 2.16* |
| $f_5 * f_6$ | 3.26 | -0.657 | 0.002 | 2.18* |
| $S$ | - | 0.010 | - | - |
| $R^2$ | - | 0.997 | - | - |
| $R^2_{adj.}$ | - | 0.992 | - | - |
| $R^2_{pr.}$ | - | 0.948 | - | - |
| Ryan-Joiner | - | 0.975 | > 0.100* | - |
| Levene | - | 0.070 | 0.933* | - |
| Durbin-Watson | - | 2.068* | - | - |

*Values used to verify the four assumptions about the residuals of the model.

the rockfall hazard map shown in Fig. 9, in which the number of rocks removed from each KP is shown for comparative purposes.

Overall, the map depicted in Fig. 9 was consistent with the need for monitoring this particular road section, since its previous and subsequent areas corresponded to values of rockfall hazard between very low and low. Instead, the road section selected for the monitoring campaign was subject to high values of hazard along most of its length. Furthermore, as highlighted in Fig. 9, the locations with the highest and lowest rockfall frequency coincided with the areas where the maximum and minimum values of rockfall hazard in the vicinity of the road section were
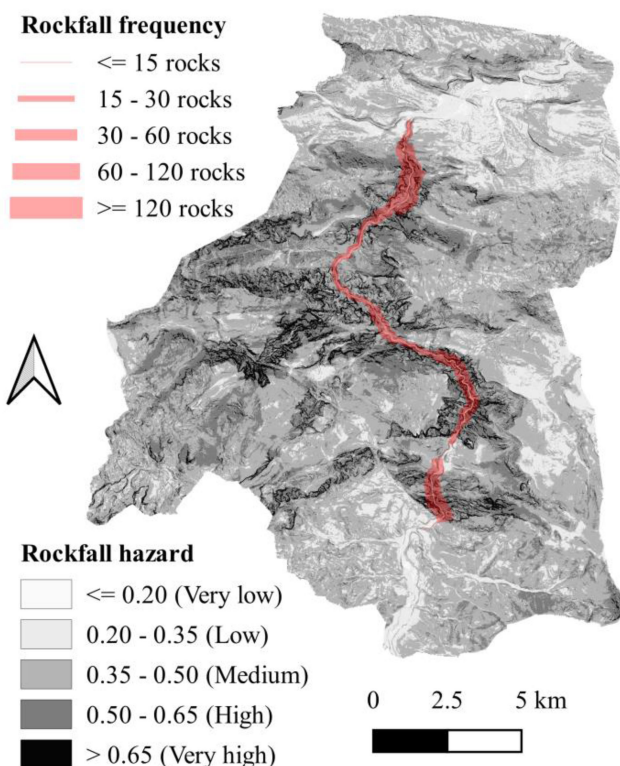


Fig. 9. Overlap between the hazard map produced using the weights determined through statistical modeling and the density of rockfall events reported along the monitored road section.

reached. This visual agreement further demonstrated that the proposed approach provided a reliable and accurate means to predict rockfall frequency.

Once the rockfall hazard map in Fig. 9 was validated both in statistical and visual terms, the results were computed again to consider the values of precipitation and thermal amplitude corresponding to the 12 climate change scenarios represented in Figs. 5 and 6. The percentage increases or decreases derived from these maps in relation to the stationary situation are provided in Table VI for comparative purposes.

The differences between the climate change and stationary scenarios were strongly affected by the variability observed in the projections of precipitation (see Fig. 5), which was the second most important variable for determining rockfall hazard ($w_{f_5} = 0.142$). Instead, the low weight of thermal amplitude ($w_{f_6} = 0.023$) prevented this factor from producing greater alterations in the results compiled in Table VI. Still, the worst scenario for the CNRM-CM5 model in terms of precipitation (RCP4.5 and 2070) resulted in an increment of 3.60% in the statistic used to build the MRA (MIN). Therefore, the impact of climate change was concluded to have a greater influence on time-series analyses involving severe storms and heat waves more explicitly. Instead, their inclusion in spatial assessments is limited by the use of average or accumulated precipitation and temperature data, which hinders the existence of large variations in the distribution and magnitude of susceptibility to experience natural disasters, such as rockfall events.

TABLE VI
Summary Statistics of the Rockfall Hazard Maps Corresponding to the Stationary and Climate Change Scenarios

| Scenario | Year | Model | MIN | MAX | RANGE | MEAN | SUM |
|---|---|---|---|---|---|---|---|
| Stationary | - | - | 0.139 | 0.912 | 0.773 | 0.412 | 4,265,482 |
| RCP4.5 | 2050 | CNRM-CM5 | +1.439% | +0.110% | -0.129% | +0.485% | +0.441% |
| RCP4.5 | 2050 | HadGEM2-CC | -6.475% | -0.768% | +0.259% | -2.427% | -2.426% |
| RCP4.5 | 2050 | MIROC5 | -4.317% | -0.329% | +0.388% | -1.699% | -1.552% |
| RCP4.5 | 2070 | CNRM-CM5 | +3.597% | +0.877% | +0.388% | +1.456% | +1.432% |
| RCP4.5 | 2070 | HadGEM2-CC | -0.719% | +0.219% | +0.388% | -0.243% | -0.238% |
| RCP4.5 | 2070 | MIROC5 | -2.878% | -0.219% | +0.259% | -1.214% | -1.256% |
| RCP8.5 | 2050 | CNRM-CM5 | +1.439% | +0.548% | +0.388% | +0.485% | +0.616% |
| RCP8.5 | 2050 | HadGEM2-CC | -2.878% | -0.329% | +0.129% | -1.214% | -1.244% |
| RCP8.5 | 2050 | MIROC5 | -2.878% | -0.110% | +0.388% | -1.214% | -1.086% |
| RCP8.5 | 2070 | CNRM-CM5 | +2.878% | +0.658% | +0.259% | +0.971% | +1.055% |
| RCP8.5 | 2070 | HadGEM2-CC | -3.597% | -0.219% | +0.388% | -1.214% | -1.250% |
| RCP8.5 | 2070 | MIROC5 | -4.317% | -0.548% | +0.129% | -1.942% | -1.854% |

## IV. Conclusion

This investigation demonstrated the suitability of combining spatial analysis and statistical modeling techniques to estimate rockfall hazard hotspots in mountainous areas. To this end, onsite reports accounting for the number of rocks removed from a road section in The Hermida Gorge (Cantabria, northern Spain) were used for validation. This approach, which also considered the effects of climate change, filled a knowledge gap in what concerns the development of interpretable and simple methodologies to prioritize rockfall hazard using globally available data. This contrasts with most previous related studies, which focused on modeling rockfall hazard under stationary conditions using complex and opaque numerical models.

The results achieved are highlighted by their accuracy, resulting in high correlation coefficients between the hazard maps produced through spatial statistical modeling and observed rockfall data. The most influential factors on this strong monotonic relationship included morphologic (slope), hydrogeological (runoff threshold), and meteorological (precipitation) variables, representing the three main aspects involved in the occurrence of natural disasters. Despite the notable importance achieved by the weather factors in the statistical model built, the effects of climate change on the spatial distribution and magnitude of rockfall hazard were not conclusive for this study area and only resulted in slight variations in its susceptibility to experience these events.

The findings of this research shed light on the relationships between the number of rockfalls reported in the road section under analysis and the factors contributing to their frequency, such that road managers can focus on the adoption of specific actions and measures at critical sites to mitigate these phenomena. In addition, either the methodology as a whole or some of its components can be used to map mass wasting risk at other mountainous locations around the globe, helping in the design of adaptation strategies to deal with natural disasters. This course of action might be undertaken in parallel with the examination of future lines of research derived from this study, such as the computerization of the methodology through an interface to facilitate its potential implementation or the testing of the proposed approach at other sites where spatial and monitoring data are available at different resolutions.

## References

[1] J. L. Zêzere et al., "DISASTER: A GIS database on hydro-geomorphologic disasters in portugal," Nat. Hazards, vol. 72, no. 2, pp. 503–532, Jun. 2014.

[2] C. Ye et al., "Landslide detection of hyperspectral remote sensing data based on deep learning with constrains," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 12, no. 12, pp. 5047–5060, 2019.

[3] U. Haque et al., "Fatal landslides in Europe," Landslides, vol. 13, no. 6, pp. 1545–1554, Dec. 2016.

[4] A. Decaulne, "Slope processes and related risk appearance within the icelandic westfjords during the twentieth century," Nat. Hazards Earth Syst. Sci., vol. 5, no. 3, pp. 309–318, 2005.

[5] S. M. F. Aghda and V. Bagheri, "Evaluation of earthquake-induced landslides hazard zonation methods: A case study of Sarein, Iran, earthquake (1997)," Arab. J. Geosci., vol. 8, no. 9, pp. 7207–7227, Sep. 2015.

[6] S. M. Fatemi Aghda, V. Bagheri, and M. Razifard, "Landslide susceptibility mapping using fuzzy logic system and its influences on mainlines in lashgarak region, Tehran, Iran," Geotech. Geol. Eng., vol. 36, no. 2, pp. 915–937, Apr. 2018.

[7] B. Damm and M. Klose, "The landslide database for Germany: Closing the gap at national level," Geomorphology, vol. 249, pp. 82–93, Nov. 2015.

[8] M. S. Alkhasawne, U. K. B. Ngah, T. L. Tien, and N. A. B. M. Isa, "Landslide susceptibility hazard mapping techniques review," J. Appl. Sci., vol. 12, no. 9, pp. 802–808, Sep. 2012.

[9] S. M. Arbanas and Z. Arbanas, Landslides: A Guide to Researching Landslide Phenomena and Processes. Hershey, PA, USA: IGI Global, 2014.

[10] S. Segoni, L. Piciullo, and S. L. Gariano, "A review of the recent literature on rainfall thresholds for landslide occurrence," Landslides, vol. 15, pp. 1483–501, Mar. 2018.

[11] W. Zhao, A. Li, X. Nan, Z. Zhang, and G. Lei, "Postearthquake landslides mapping from Landsat-8 data for the 2015 nepal earthquake using a pixel-based change detection method," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 10, no. 5, pp. 1758–1768, 2017.

[12] J. Corominas, O. Mavrouli, and R. Ruiz-Carulla, "Rockfall occurrence and fragmentation," in Proc. Adv. Culture Living Landslides, 2017, pp. 75–97.

[13] I. Yilmaz, M. Yildirim, and I. Keskin, "A method for mapping the spatial distribution of rockfall computer program analyses results using ArcGIS software," Bull. Eng. Geol. Environ., vol. 67, no. 4, pp. 547–554, 2008.

[14] L. Losasso and F. Sdao, "The artificial neural network for the rockfall susceptibility assessment. A case study in Basilicata (Southern Italy)," Geomatics, Nat. Hazards Risk, vol. 9, no. 1, pp. 737–759, 2018.

[15] J. Cao, Z. Zhang, J. Du, L. Zhang, Y. Song, and G. Sun, "Multi-geohazards susceptibility mapping based on machine learning—A case study in Jiuzhaigou, China," Nat. Hazards, vol. 102, no. 3, pp. 851–871, 2020.

[16] A. M. Fanos, B. Pradhan, A. Alamri, and C. W. Lee, "Machine learning-based and 3d kinematic models for rockfall hazard assessment using LiDAR data and GIS," Remote Sens., vol. 12, no. 11, 2020, Art. no. 1755.

[17] A. M. Fanos and B. Pradhan, "A novel rockfall hazard assessment using laser scanning data and 3D modelling in GIS," CATENA, vol. 172, pp. 435–450, Jan. 2019.

[18] C. B. Azodi, J. Tang, and S.-H. Shiu, "Opening the black box: Interpretable machine learning for geneticists," *Trends Genet.*, vol. 36, no. 6, pp. 442–455, 2020.

[19] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nat. Mach. Intell.*, vol. 1, no. 5, pp. 206–215, 2019.

[20] G. Matas, N. Lantada, J. Corominas, J. A. Gili, R. Ruiz-Carulla, and A. Prades, "RockGIS: A GIS-based model for the analysis of fragmentation in rockfalls," *Landslides*, vol. 14, no. 5, pp. 1565–1578, 2017.

[21] M. Calista, V. Menna, V. Mancinelli, N. Sciarra, and E. Miccadei, "Rockfall and debris flow hazard assessment in the SW escarpment of montagna del morrone ridge (Abruzzo, Central Italy)," *Water (Switzerland)*, vol. 12, no. 4, 2020, Art. no. 1206.

[22] H. J. Park, J. H. Lee, K. M. Kim, and J. G. Um, "Assessment of rock slope stability using GIS-based probabilistic kinematic analysis," *Eng. Geol.*, vol. 203, pp. 56–69, 2016.

[23] N. Depountis, K. Nikolakopoulos, K. Kavoura, and N. Sabatakakis, "Description of a GIS-based rockfall hazard assessment methodology and its application in mountainous sites," *Bull. Eng. Geol. Environ.*, vol. 79, no. 2, pp. 645–658, 2020.

[24] F. Huettmann and T. Gottschalk, "Simplicity, model fit, complexity and uncertainty in spatial prediction models applied over time: We are quite sure, aren't we?," in *Predictive Species and Habitat Modeling in Landscape Ecology*. New York, NY, USA: Springer, 2011.

[25] BOC, "Ley 4/2006, de 19 de mayo, de Conservación de la naturaleza de Cantabria," *Nature Conservation Cantabria*, 2006. Act 4/2006 May 19, Accessed: Jun. 19, 2018. [Online]. Available: http://www.territoriodecantabria.es/-/ley-4-2006-de-19-de-mayo-de-conservacion-de-la-naturaleza-de-cantabria

[26] elEconomista.es, "Picos de europa, un centenario de parque natural," 2018. Accessed: Jun. 19, 2018. [Online]. Available: http://www.eleconomista.es/evasion/viajar/noticias/9187922/06/18/Picos-de-Europa-un-centenario-de-Parque-Natural.html

[27] ifomo Cantabria, "The route of The Hermida Cillorigo-Panes Gorge is taken to public information in terms of expropriations," 2018. Accessed: May 31, 2018. [Online]. Available: https://www.ifomocantabria.es/cantabria/municipios/liebana/a-informacion-publica-a-efectos-de-expropiaciones-el-trazado-del-desfiladero-de-la-hermida-cillorigo-panes

[28] R. Ancell Trueba and R. Célis Díaz, "Cantabria's thermo-pluviometry during the period 1981–2000," Madrid, Spain, 2013.

[29] IGME, "Spanish Geological Survey maps. Cabreña - Cabrales," Madrid, Spain, 1984.

[30] IGME, "IELIG - CA068: The Hermida Gorge," *Spanish Inventory Geological Sites Interest*, . Accessed: Mar. 24, 2021. [Online]. Available: http://info.igme.es/ielig/LIGInfo.aspx?codigo=CA068#documentacion

[31] ValledeLiebana.info, "Desprendimientos de piedras en el Desfiladero de La Hermida," *Píldora de historia del Valle de Liébana*, Rockfalls in The Hermida Gorge. History of the Liébana Valley, Cantabria, Spain , 2013. Accessed: May 31, 2018. [Online]. Available: http://www.valledeliebana.info/historias/igual.html

[32] El Diario Montañes, "Un desprendimiento de rocas bloquea el desfiladero de la hermida," A rockfall event blocks The Hermida Gorge, 2017. Accessed: May 29, 2018. [Online]. Available: http://www.eldiariomontanes.es/cantabria/desprendimiento-rocas-bloquea-20171109130650-nt.html

[33] ALERTA, "The route of The Hermida Cillorigo-Panes Gorge is taken to public information in terms of expropriations," 2018. Accessed: May 31, 2018. [Online]. Available: http://www.eldiarioalerta.com/articulo/cantabria/informacion-publica-efectos-expropiaciones-trazado-desfiladero-hermida-cillorigo-panes/20180530104929029525.html

[34] D. Stead, M. A. Brideau, T. Paddington, T. Millard, and K. Turner, "Characterization of debris slope failures: An integrated field based numerical modelling study," in *Proc. 17th Int. Conf. Soil Mech. Geotech. Eng.*, 2009, vol. 5, pp. 3495–3499.

[35] G. D. Bathrellos, H. D. Skilodimou, K. Chousianitis, A. M. Youssef, and B. Pradhan, "Suitability estimation for urban development using multi-hazard assessment map," *Sci. Total Environ.*, vol. 575, pp. 119–134, Jan. 2017.

[36] LP DAAC, "*ASTGTM: ASTER Global Digital Elevation Model V002*," 2014. Accessed: Jul. 04, 2018. [Online]. Available: https://lpdaac.usgs.gov/node/1079

[37] J. Hartmann and N. Moosdorf, "The new global lithological map database GLiM: A representation of rock properties at the earth surface," *Geochem., Geophys. Geosyst.*, vol. 13, no. 12, pp. 1–37, Dec. 2012.

[38] C. Jun, Y. Ban, and S. Li, "Open access to earth land-cover map," *Nature*, vol. 514, no. 7523, pp. 434–434, Oct. 2014.

[39] C. Y. Wu and S. C. Chen, "Integrating spatial, temporal, and size probabilities for the annual landslide hazard maps in the shihmen watershed, Taiwan," *Nat. Hazards Earth Syst. Sci.*, vol. 13, no. 9, pp. 2353–2367, Sep. 2013.

[40] R. K. Dahal, S. Hasegawa, N. P. Bhandary, P. P. Poudel, A. Nonomura, and R. Yatabe, "A replication of landslide hazard mapping at catchment scale," *Geomatics, Nat. Hazards Risk*, vol. 3, no. 2, pp. 161–192, Jun. 2012.

[41] S. Rampone and A. Valente, "Neural network aided evaluation of landslide susceptibility in southern Italy," *Int. J. Mod. Phys. C*, vol. 23, no. 1, Jan. 2012, Art. no. 1250002.

[42] S. Lee and B. Pradhan, "Landslide hazard mapping at Selangor, Malaysia using frequency ratio and logistic regression models," *Landslides*, vol. 4, no. 1, pp. 33–41, Feb. 2007.

[43] A. N. Matori, A. Basith, and I. S. H. Harahap, "Study of regional monsoonal effects on landslide hazard zonation in cameron highlands, malaysia," *Arab. J. Geosci.*, vol. 5, no. 5, pp. 1069–1084, Sep. 2012.

[44] CNIG, "Centro de descargas del CNIG (IGN)," *National Centre Geographic Informat. Download Center*, 2018. Accessed: May 28, 2018. [Online]. Available: http://centrodedescargas.cnig.es/CentroDescargas/locale?request_locale=en

[45] F. Zhang, W. Chen, G. Liu, S. Liang, C. Kang, and F. He, "Relationships between landslide types and topographic attributes in a loess catchment, China," *J. Mountain Sci.*, vol. 9, no. 6, pp. 742–751, Dec. 2012.

[46] G. C. Ohlmacher and J. C. Davis, "Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA," *Eng. Geol.*, vol. 69, no. 3–4, pp. 331–343, Jun. 2003.

[47] W. Chen *et al.*, "GIS-based landslide susceptibility evaluation using a novel hybrid integration approach of bivariate statistical based random forest method," *Catena*, vol. 164, pp. 135–149, May 2018.

[48] IDE UNEX, "Mapa litoestratigráfico de España 1:200.000," *Spanish lithostratigraphic map*, 2015. Accessed: Feb. 22, 2018. [Online]. Available: http://ide.unex.es/conocimiento/index.php?/article/AA-00216/0/Mapa-litoestratigrfico-de-Espaa-1%3A200.000.html

[49] M. G. Srapionyan and S. Y. Vartanyan, "Physico-mechanical properties of rocks and structural-lithologic conditions of mineralization in copper-molybdenum deposits of Armenia," *Int. Geol. Rev.*, vol. 21, no. 7, pp. 833–838, Jul. 1979.

[50] V. Summa, S. Margiotta, R. Colaiacovo, and M. L. Giannossi, "The influence of the grain-size, mineralogical and geochemical composition on the verdesca landslide," *Nat. Hazards Earth Syst. Sci.*, vol. 15, no. 1, pp. 135–146, Jan. 2015.

[51] EEA, "CORINE land cover—European environment agency," 1995. Accessed: May 28, 2018. [Online]. Available: https://www.eea.europa.eu/publications/COR0-landcover

[52] BOE, "Orden FOM/298/2016, de 15 de febrero, por la que se aprueba la norma 5.2 - IC drenaje superficial de la instrucción de carreteras," in *Official State Gazette*, Surface Drainage Standard 5.2-IC of the Spanish Roads Instruction (FOM/298/2016), 2016. Accessed: May 29, 2018. [Online]. Available: https://www.boe.es/buscar/doc.php?id=BOE-A-2016-2405

[53] I. Egeli and H. F. Pulat, "Mechanism and modelling of shallow soil slope stability during high intensity and short duration rainfall," *Sci. Iran.*, vol. 18, no. 6, pp. 1179–1187, Dec. 2011.

[54] P. Teillet, "Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indices of forested regions," *Remote Sens. Environ.*, vol. 61, no. 1, pp. 139–149, Jul. 1997.

[55] W.-T. Lin, C.-Y. Lin, and W.-C. Chou, "Assessment of vegetation recovery and soil erosion at landslides caused by a catastrophic earthquake: A case study in central Taiwan," *Ecol. Eng.*, vol. 28, no. 1, pp. 79–89, Nov. 2006.

[56] E. Arnone, D. Caracciolo, L. V. Noto, F. Preti, and R. L. Bras, "Modeling the hydrological and mechanical effect of roots on shallow landslides," *Water Resour. Res.*, vol. 52, no. 11, pp. 8590–8612, 2016.

[57] NASA, "Landsat 8 bands," 2013. Accessed: May 28, 2018. [Online]. Available: https://landsat.gsfc.nasa.gov/landsat-8/landsat-8-bands/

[58] L.-W. Wei *et al.*, "The mechanism of rockfall disaster: A case study from Badouzih, Keelung, in northern Taiwan," *Eng. Geol.*, vol. 183, pp. 116–126, Dec. 2014.

[59] T. Dahl-Jensen *et al.*, "Landslide and Tsunami 21 November 2000 in Paatuut, West Greenland," *Nat. Hazards*, vol. 31, no. 1, pp. 277–287, Jan. 2004.

[60] R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, "Very high resolution interpolated climate surfaces for global land areas," *Int. J. Climatol.*, vol. 25, no. 15, pp. 1965–1978, Dec. 2005.

[61] K. E. Taylor, R. J. Stouffer, and G. A. Meehl, "An overview of CMIP5 and the experiment design," *Bull. Amer. Meteorol. Soc.*, vol. 93, no. 4, pp. 485–498, Apr. 2012.

[62] D. P. van Vuuren *et al.*, "The representative concentration pathways: An overview," *Clim. Change*, vol. 109, no. 1, pp. 5–31, Nov. 2011.

[63] ESRI, "ArcGIS desktop," 2013. Accessed: Jun. 07, 2018. [Online]. Available: https://www.esri.com/en-us/store/arcgis-desktop

[64] R. Ginevicius and V. Podvezko, "Some problems of evaluating multicriteria decision methods," *Int. J. Manage. Decis. Making*, vol. 8, no. 5/6, pp. 527–539, 2007.

[65] H. Zhang, G. Zhang, and Q. Jia, "Integration of analytical hierarchy process and landslide susceptibility index based landslide susceptibility assessment of the Pearl river delta area, China," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 11, pp. 4239–4251, Nov. 2019.

[66] R. A. Fisher, *Statistical Methods for Research Workers*. Edinburgh, Scotland: Cosmo, 1925.

[67] J. Zolgharnein and A. Shahmoradi, "Characterization of sorption isotherms, kinetic models, and multivariate approach for optimization of hg(ii) adsorption onto *fraxinus* tree leaves," *J. Chem. Eng. Data*, vol. 55, no. 11, pp. 5040–5049, Nov. 2010.

[68] J. L. Gastwirth, Y. R. Gel, and W. Miao, "The impact of Levene's test of equality of variances on statistical theory and practice," *Stat. Sci.*, vol. 24, no. 3, pp. 343–360, Aug. 2009.

[69] D. W. Marquaridt, "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation," *Technometrics*, vol. 12, no. 3, pp. 591–612, Aug. 1970.

[70] N. E. Savin and K. J. White, "The Durbin-Watson test for serial correlation with extreme sample sizes or many regressors," *Econometrica*, vol. 45, no. 8, pp. 1989–1996, 1977.

[71] R. M. Duarte and J. Marquínez, "The influence of environmental and lithologic factors on rockfall at a regional scale: An evaluation using GIS," *Geomorphology*, vol. 43, no. 1–2, pp. 117–136, 2002.

[72] J. Marquínez, R. Menéndez Duarte, P. Farias, and M. Jiménez Sánchez, "Predictive GIS-based model of rockfall activity in mountain cliffs," *Nat. Hazards*, vol. 30, no. 3, pp. 341–360, 2003.

[73] IPCC, "Climate change 2014: Synthesis report. Summary for policymakers," Geneva, Switzerland, 2014.

[74] F. Baillifard, M. Jaboyedoff, and M. Sartori, "Rockfall hazard mapping along a mountainous road in switzerland using a GIS-based parameter rating approach," *Nat. Hazards Earth Syst. Sci.*, vol. 3, no. 5, pp. 435–442, 2003.

[75] A. Omran, K. Fahmida, D. Schröder, M. O. Arnous, A. E. El-Rayes, and V. Hochschild, "GIS-based rockfall hazard zones modeling along the coastal gulf of Aqaba region, Egypt," *Earth Sci. Informat.*, vol. 327, pp. 691–709, 2021.

[76] M. R. Maleki, "Study of the engineering geological problems of the Havasan dam, with emphasis on clay-filled joints in the right abutment," *Rock Mech. Rock Eng.*, vol. 44, no. 6, pp. 695–710, 2011.

[77] R. Paranunzio, F. Laio, M. Chiarle, G. Nigrelli, and F. Guzzetti, "Climate anomalies associated with the occurrence of rockfalls at high-elevation in the Italian alps," *Nat. Hazards Earth Syst. Sci.*, vol. 16, no. 9, pp. 2085–2106, 2016.

[78] O. Sass, "Temporal variability of rockfall in the Bavarian Alps, Germany," *Arctic, Antarct. Alpine Res.*, vol. 37, no. 4, pp. 564–573, 2005.

[79] R. Wang *et al.*, "Rockfall hazard identification and assessment of the Langxian- Milin section of the transmission line passage of central Tibet grid interconnection project," in *Proc. IOP Conf. Series, Earth Environ. Sci.*, 2018, vol. 189, Paper 052046.

[80] J. W. Borella, M. Quigley, and L. Vick, "Anthropocene rockfalls travel farther than prehistoric predecessors," *Sci. Adv.*, vol. 2, no. 9, 2016, Art. no. e1600969.

[81] J. Remondo, A. González, J. R. Díaz de Terán, A. Cendrero, A. Fabbri, and C.-J. F. Chung, "Validation of landslide susceptibility maps; examples and applications from a case study in northern Spain," *Nat. Hazards*, vol. 30, no. 3, pp. 437–449, 2003.

**Daniel Jato-Espino** was born in Santander, Cantabria, Spain, in 1986. He received the B.S., M.S., and Ph.D. degrees in civil engineering from the University of Cantabria in 2009, 2012, and 2016, respectively.

From 2013 to 2019, he was a Research Assistant with the Construction Technology Applied Research Group (GITECO) and a Visiting Researcher with the Centre for Agroecology, Water, and Resilience (CAWR). In 2019, he worked as a Data Scientist and Chief Research Officer with Método Sáncal S.L. Then, he returned to the University of Cantabria, where he combined his job as a Postdoctoral Researcher in the INGEPRO Research Group with his role as a Research Coordinator in Green Fortland Foundation, a Ghanaian NGO. He currently works as a Senior Researcher with the Universidad Internacional de Valencia. He has authored more than 35 articles and holds one international patent. His research interests include water management, urban hydrology, and flood risk assessment adapted to climate change, with emphasis on the use of green infrastructure.

Dr. Jato-Espino is a Topic Editor of the journal *Infrastructures* and has served as a Special Issue Editor for the journals *Sustainability* and *Frontiers in Water*. Hewas the recipient of two awards as the author of the best Ph.D. Thesis in the area of engineering and architecture (University of Cantabria, Spain) and Best Ph.D. Thesis in the area of engineering and technologies (Juan María Parés Social Council, Spain).

**Alejandro Roldan-Valcarce** was born in Marbella, Andalucía (Spain). He received the Bachelor's degree in environmental sciences from the University of León, León, Spain, 2014, and the Master's degree in meteorology from the Complutense University of Madrid, Madrid, Spain, in 2018.

His Master's thesis was focused on the use of hysplit, an atmospheric dispersion model for simulating radioactive pollution. In 2019, he joined the Construction Technology Applied Research Group (GITECO), the Department of Transport and Project and Processes Technology, the University of Cantabria, working on the study of the risks associated with the impact of extreme events on transport infrastructures in Europe under the European project FORESEE. In 2020, he received the Ph.D. grant from the Spanish Government to develop a tool for the integration of models to support the location and selection of low impact development practices for the MELODRAIN project, which aims at the characterization of filter sections for intelligent and sustainable management of urban drainage systems in cities.

**Felipe Collazos-Arias** was born in Caceres, Spain, in 1978. He received the M.S. degree in civil engineering from the University of Cantabria, Santander, Spain, and the B.S and Ph.D. degrees from the University of Extremadura, Badajoz, Spain.

He works for the Spanish Road Directorate Ministry of Transport in Cantabria, Spain, and he is an Associate Professor with the the Department of Transport and Project and Processes Technology, the University of Cantabria.

Dr. Collazos-Arias is a member of the Construction Technology Applied Research Group (GITECO), the International Association for Bridge Maintenance and Safety (IABMAS ICM 102), the Transversal Working Group on "Climate Change and Resilience in Roads" recently created within the Technical Committee on Environment of the Technical Association of Roads- World Road Association (ATC-PIARC), and a nominated member of the TC 202 Transportation Geotechnics of the International Society for Soil Mechanics and Geotechnical Engineering (ISSGME), representing the Spanish Society of Soil Mechanics and Geotechnical Engineering (SEMSIG).

**Jorge Rodriguez-Hernandez** was born in Reinosa, Cantabria, Spain, in 1978. He received the degree of Professional Engineer in roads, canals and ports ("Ingeniero de Caminos, Canales y Puertos"), and the Ph.D. degree focused on permeable pavements and sustainable construction, from the University of Cantabria, Santander, Spain, in 2003 and 2008, respectively.

Since then, he has developed his work in the Construction Technology Applied Research Group (GITECO), the Department of Transport and Project and Processes Technology of that university, in collaboration with other institutions like Coventry University. He has participated as a Principal Investigator (PI) in several competitive R&D projects dealing with sustainable urban pavements and risk assessment in infrastructures, some of them financed by the Spanish Government. Moreover, he has experience in European calls like INFRAVATION, FP VII or LIFE, and noncompetitive contracts with companies. With 5 cosupervisions of doctoral thesis (2010, 2014, 2016, 2017, and 2020), more than 40 contributions to conferences, and 40 publications indexed in the JCR, of which more than 15 are in the first quartile (Q1), he has received around 500 citations according to WoS with a h-index of 12. Regarding technology transfer, he has produced 3 patents whose owners are private companies, 2 of them at an international scale. All his work is mainly focused on the sustainable construction of urban and interurban transport and drainage infrastructure.