Managing Rockfall Hazard on Strategic Linear Stakes: How Can Machine Learning Help to Better Predict Periods of Increased Rockfall Activity?

Marie-Aurélie Chanut, Hermann Courteille, Clara Lévy, Abdourrahmane Atto, Lucas Meignan, Emmanuel Trouvé, Muriel Gasc-Barbier

sustainability

MDPI

Article

Managing Rockfall Hazard on Strategic Linear Stakes: How Can Machine Learning Help to Better Predict Periods of Increased Rockfall Activity?

Marie-Aurélie Chanut [1,*], Hermann Courteille [2], Clara Lévy [3], Abdourrahmane Atto [2], Lucas Meignan [4], Emmanuel Trouvé [2] and Muriel Gasc-Barbier [1,5]

1 Team GéoCoD, Cerema, 69674 Bron, France; muriel.gasc@cerema.fr
2 LISTIC, Université Savoie Mont-Blanc, 74944 Annecy, France; hermann.courteille@irisa.fr (H.C.); abdourrahmane.atto@univ-smb.fr (A.A.); emmanuel.trouve@univ-smb.fr (E.T.)
3 BRGM, 45060 Orléans, France; c.levy@brgm.fr
4 Géolithe, 38920 Crolles, France; lucas.meignan@geolithe.com
5 GERS, Université Gustave Eiffel, 77454 Marne-la-Vallée, France
* Correspondence: marie-aurelie.chanut@cerema.fr

Abstract: When rockfalls hit and damage linear stakes such as roads or railways, the access to critical infrastructures (hospitals, schools, factories . . .) might be disturbed or stopped. Rockfall risk management often involves building protective structures that are traditionally based on the intensive use of resources such as steel or concrete. However, these solutions are expensive, considering their construction and maintenance, and it is very difficult to protect long linear stakes. A more sustainable and effective risk management strategy could be to account for changes on rockfall activity related to weather conditions. By integrating sustainability principles, we can implement mitigation measures that are less resource-intensive and more adaptable to environmental changes. For instance, instead of solely relying on physical barriers, solutions could include measures such as restriction of access, monitoring and mobilization of emergency kits containing eco-friendly materials. A critical step in developing such a strategy is accurately predicting periods of increased rockfall activity according to meteorological triggers. In this paper, we test four machine learning models to predict rockfalls on the National Road 1 at La Réunion, a key road for the socio-economic life of the island. Rainfall and rockfall data are used as inputs of the predictive models. We show that a set of features derived from the rainfall and rockfall data can predict rockfall with performances very close and almost slightly better than the standard expert model used for operational management. Metrics describing the performance of these models are translated in operational terms, such as road safety or the duration of road closings and openings, providing actionable insights for sustainable risk management practices.

Keywords: rockfall; rainfall; machine learning; predictive models; risk management; transient mitigation measures

Citation: Chanut, M.-A.; Courteille, H.; Lévy, C.; Atto, A.; Meignan, L.; Trouvé, E.; Gasc-Barbier, M. Managing Rockfall Hazard on Strategic Linear Stakes: How Can Machine Learning Help to Better Predict Periods of Increased Rockfall Activity? *Sustainability* **2024**, *16*, 3802. https://doi.org/10.3390/su16093802

Academic Editors: Fernando António Leal Pacheco and Yaxun Xiao

Received: 14 February 2024
Revised: 27 March 2024
Accepted: 26 April 2024
Published: 30 April 2024

1. Introduction

Rockfalls are gravity-driven instabilities affecting steep slopes of soft and hard rocks [1–3]. Rockfall volumes span several orders of magnitude at the site scale, from about 10^{-2} m^3 up to 10^2 m^3 or more [4–6]. Rockfalls can hit and damage linear stakes, such as roads or railways, thus having an impact on the access to critical infrastructures (hospitals, schools, factories . . .) [7]. Building protective structures (hanging nets, gabion-walls. . .) is the main chosen solution to manage this risk [8]. However, the real cost of installing protective structures is often underestimated, as their maintenance is often not carried out correctly. In addition, the design of the structures is based on current hypotheses, but considering the prospective impacts of climate change, rockfall activity and hazard may change significantly in the coming years. When built, protective structures are difficult to change. Furthermore,

for managers who have to protect a significant length of infrastructure and have a limited budget, it often makes no sense to install protection structures in one area without being able to protect the area next to it. Faced with this situation, new risk management practices for linear stakes are being developed, particularly taking into account the increase in the level of hazards when implementing mitigation measures (restricting access, monitoring, mobilizing emergency kits, etc.). This approach could be a more sustainable and effective way of managing risks. This solution is more economical, saving on raw materials such as steel and concrete, and does not alter the ground in situ. Nevertheless, a critical step in developing such a strategy is the proper prediction of periods with increased rockfall activity according to identified triggers. Figure 1 illustrates the two approaches for rockfall risk management: the main chosen approach using protective works and the proposed approach using temporary mitigation measures. Table 1 compares, for each approach, their effects on risk, their relative building cost and maintenance cost, their relative CO_2 emissions, as well as the constraints for implementation and maintenance.

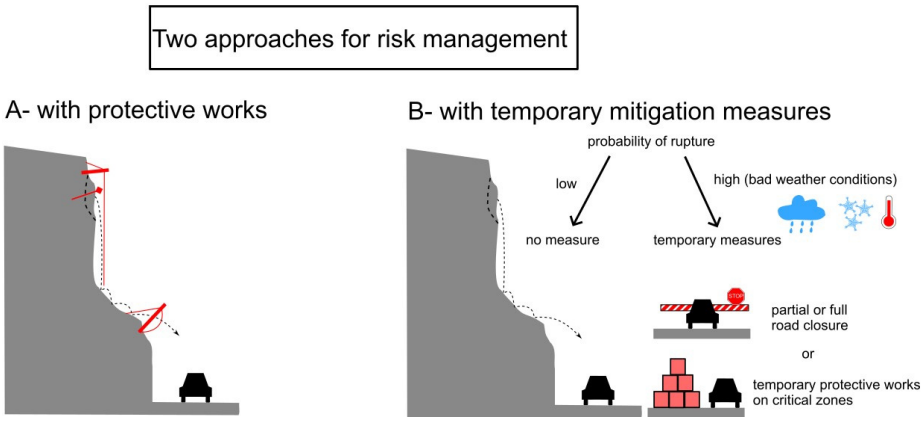


Figure 1. Two approaches for rockfall risk management: the main chosen approach, using protective works (**A**) and the proposed approach, using temporary mitigation measures (**B**).

Table 1. Characteristics of two approaches for rockfall risk management.

	A	**B**
Effects on risk	Decrease in: - Probability of rupture - Probability of propagation (on the road) Increase in: - Exposure ratio for rope workers	Decrease in: - Exposure ratio during road closure - Probability of propagation with temporary protective work
Build cost	high	low
Maintenance cost	high	low
$C0_2$ emission	high	low
Implementation	- Need access and space to build protective works - Protectives works cannot prevent high magnitude events - Generally too costly to apply all along linear stakes	- Need data to train the predictive model - Need permanent staff to implement temporary measures - Risk level depends on the accuracy of the predictive model - Risk never tends to be zero - Need unexposed alternate road during full road closure
Maintenance	- Difficulties in adapting to hazard changes due to global warming - Difficulties in maintaining protective works	- Adapt to hazard changes due to global warming with predictive model training on new data

To carry out this work, it is necessary to understand the physics relative to rockfalls. It was evidenced that the first-order control on the rate of rock wall retreat are the geological and geotechnical characteristics of the cliff (e.g., the mechanical resistance of the rock mass, the discontinuity network, etc.) when studying similar environments, such as alpine mountain slopes [5] or coastal cliffs [9]. These observations are consistent with many descriptions of rockfall scars. These surfaces consist mostly of pre-existing discontinuities (fractures, cracks, joints, strata limits, . . .), and only small rock bridges maintain the rock compartments prior to their collapse [10–14]. Thus, the evolution of a rock compartment toward its collapse depends on the degradation of the mechanical characteristics of theses rock bridges and their reduction in number and section. Without human intervention, this evolution spreads over a very long period of time under the influence of environmental factors such as meteorological factors [10], earthquakes [15–17], volcanic eruptions [18,19] or sea waves [20–22]. As reported in [23], several physical mechanisms can be involved by the same meteorological factor (for example, rainfall can induce water pressure in rock joints, clay swelling in rock joints, rock dissolution, leaching).

The processes that most significantly influence rockfall activity depend on the geographical context. For an alpine cliff of average altitude, the rockfall frequency can be multiplied by 7 during freeze-thaw episodes and by 26 when the mean rainfall intensity is higher than 5 $mm \cdot h^{-1}$ [23]. For high alpine cliffs, the permafrost melting is the major factor [24,25]. In La Reunion Island, located in a tropical environment, the only evidenced meteorological trigger with statistical significance is rain [26]. Thus, it is possible to establish a link between rockfall activity and environmental triggers when databases of rockfall occurrences covering long periods of time exist. A review of the available statistical methods to define a rainfall-rockfall relationship is proposed [27], allowing for a better understanding of its stochastic nature. Circular von Mises distributions are used to show the correlation between rockfall frequencies and meteorological conditions and show the weighting of each meteorological trigger [28–30]. The authors of [31] adopt a consolidated statistical approach defining rainfall thresholds based on both the cumulated rainfall and the duration of rainfall events related to each rockfall.

Statistical tools could provide real-time forecast of rockfall hazards and the associated risk [27]. Indeed, our ability to predict periods of increased activity according to meteorological triggers already provides some benefits for operational risk management. Infrastructure managers (roads, railways, navigable waterways, bridges, etc.) should be able to implement a risk management strategy that includes the following:

- Ensuring the safety of users and staff by preventing accidents and secondary accidents;
- Maintaining continuity of service (the road must remain open for as long as possible);
- Justifying the money spent on infrastructure maintenance;
- Guaranteeing their own legal certainty.

If we focus only on the two first points, traffic has to be maintained for as long as possible and, in the meantime, the road user (or other linear infrastructure) should remain safe. When a rockfall hits a road, the road often has to be closed, even partially, to clear and secure the cliff below. If a traffic alternator or if a diversion is needed, the risk of users' exposure should not be greater in the waiting area or on the diversion than on the road. In the meantime, preventive road closures can be decided when periods of increased rockfall activity due to meteorological triggers are predicted, but such closures may not be well accepted if they occur too frequently without any accompanying rockfall incidents. Thus, if an automatic management system of the traffic is required, this system should be able to minimize the unneeded closures and, if possible, ensure road closure as shortly as possible before rockfall events. To achieve this, the site should be described very precisely (e.g., its geological and geotechnical settings, the spatial variability of rockfall hazard), as well as the triggering factors of rockfalls.

In geoscience domains, such as landslide analysis, many machine learning (ML) models have been proposed in the last decade, in parallel with the densification of satellite, lidar or radar acquisitions. A general survey is presented in [32]. For landslides, many devel-

opments using ML and Deep Learning (DL) are focused on change detection, landslides inventory and susceptibility mapping [33] or rockfalls detection [34]. Up to now, to the best of our knowledge, ML has not been used for rockfall prediction. Statistical studies [23,26,27] have found some triggering factors of rockfalls, such as rainfall amounts and freeze-thaw cycles and have provided some predictions, but the performance of these statistical models has not been compared until now. The novelty of our article lies in using ML tools to predict rockfalls, defining tools to quantify model performance, comparing these models, and interpreting the results for the operational management of transport infrastructures.

We therefore propose using ML algorithms to improve the prediction of periods of increased rockfall activity, as they are more automatic and more likely to look for causal links in rare events. Here, the approach is only temporal. The suitability of 4 ML models to describe rockfalls triggered by meteorological conditions, especially rainfall, is evaluated. A 13 km-long road on La Réunion Island (Indian Ocean), which is prone to recurrent rockfalls, has been selected because of the road's strong socio-economic role, the large amount of protection works carried out, the long-term rockfall database available and the high rate of rockfall production.

After describing the study site (Section 2) and the data used to build predictive models (Section 3), the ML models obtained are described in Section 4. The models' predictions and the support provided for operational management of the road are finally discussed in Section 5.

2. Study Site

2.1. Location, Topography and Stakes

The study site is the RN1 road on La Réunion, a volcanic island located near the south-east coast of Africa. It is 82 km-long and connects the north to the south of the island through the West coast. It is a two-way, four-lane road where the traffic has increased from 45,000 vehicles/day in 2006 [35] to about 60,000 vehicles/day in 2018 [36]. The studied section is 13 km long and stretches between La Possession and Saint Denis (Figure 2a). This section lies between 90 to 200 m-high subvertical cliffs and the sea (Figure 2b). The cliffs are partly human-made, cut in the 1950' to build the road. Thus, several of the original valleys are now situated in the middle of the cliffs (Figure 2d). Two rivers still flow directly to the sea: Jacques, at landmark KP 7.5 km (or Kilometre Point 7.5), and Grande Chaloupe, at landmark KP 8.5 km (Figure 2c).

2.2. Geology

The cliffs are composed of one to three different geological units, named inferior, intermediate and superior (Figure A1, [37–40]). The inferior unit corresponds to a lithological structure composed of alternating layers of basalt, crystallized magma, scoriae and sedimentary rocks. It was deposited during the volcanic activity of the Piton des Neiges, about 2 million years ago. Dykes are abundantly present on this structure due to volcanic activity. The intermediate unit was deposited afterward, during a stage of volcanic dormancy. This unit corresponds to sediments produced by erosion that filled valley beds. The superior unit corresponds to a lithological structure very similar to the inferior unit. It was deposited during the volcanic activity of the Piton de la Fournaise volcano, about 1 million years ago. Its basal interface accommodates the morphologies of paleo valleys.

The inferior unit is present all along the road. As it is the oldest geological unit, it is also the most altered. The intermediary unit appears only in some segments. The superior unit appears at KP 2 km, but is mainly present between KP 5 km and KP 10 km (Figure A1).

All these geological units produce rockfalls that can reach volumes up to 5000 m^3 (Figure 2b). The significant rockfall activity justified the collection of rockfall events reaching the road since 2000 (until 2018) by the department in charge of road management (e.g., 18 years of observations and 827 reported rockfall events).

Different rupture mechanisms can be observed, such as edge failures, toppling of rock columns, overhang failures, etc. Rockfalls occur along pre-existing discontinuities

(dykes, sills, diaclases, decompression fractures, the flanks of gully structures, such as paleo-valleys), especially in weathered sites.

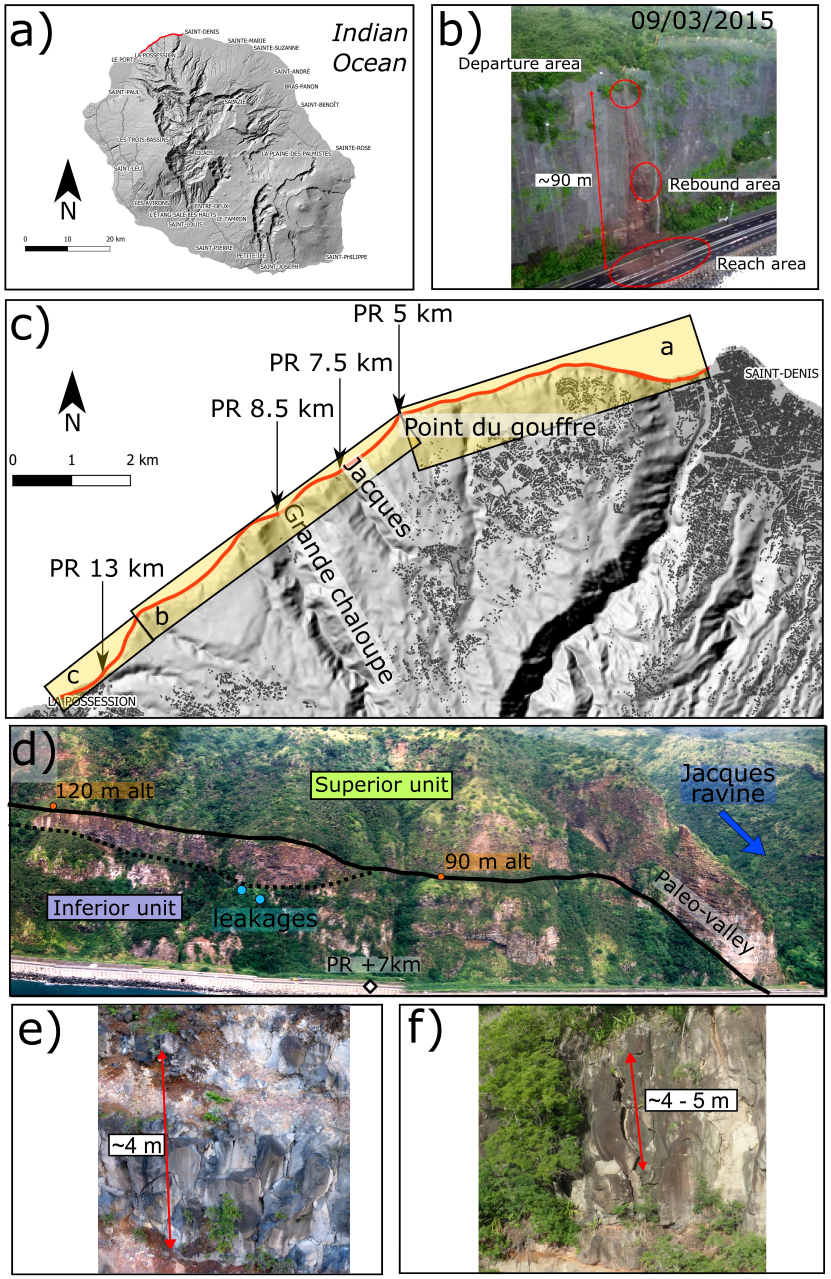


Figure 2. (**a**) Localization of the RN1 road (red line) along the coast of La Reunion Island, linking the cities of Saint Denis and La Possession. (**b**) Photograph of the RN1 road after a rockfall that reached the road on the 09/03/2015, breaking steel nets and gabions in its path (image from report BRGM/RP-64556-FR). (**c**) Zoom in on the RN1 settings, with indications of several landmarks (KP) along the road. Two important ravines reach the road: Jacques, at KP 7.5 km, and Grande Chaloupe, at KP 8.5 km. The yellow rectangles correspond to the figures presenting the detailed geology of the site in Appendix A. (**d**) Photograph of the cliff around KP 7 km, with the indication of the limits between several geological units and other hydro-geomorphological features. Pictures (**e**,**f**) represent outcrops of the superior unit, with variable basalt flow thickness and roughness. Plot (**f**) also shows an instable rock volume.

2.3. History: The Construction of Protective Works

The exposure of the stakes evolved through time, as some protective works have been carried out on the RN1. Indeed, the road was progressively expanded toward the sea, and gabions were placed between the cliff and the road to avoid accidents due to rockfalls. From 1994 to 1996, protective steel nets were installed in strategical areas. Such protective works were completed in 2006–2008 with the addition of numerous protective nets (rockfall drapery systems, defector nets, flexible barriers, gabion structures), now covering about 57% of the 13 km route (Figure A2).

The number of rockfalls hitting the road drastically decreased after the completion of the 2006–2008 protective works (696 rockfalls occurred between 2000 and 2006 and "only" 131 rockfalls between 2009 and 2018). However, despite all the above-mentioned protective works, rockfalls still hit the road, often reaching the two lanes near the cliff (Figure 2b) and sometimes also impacting the other two.

2.4. Meteorology and Rockfall Activity

Thanks to the above-mentioned database, Ref. [26] studied the link between rockfall activity and meteorological triggers on RN1. As the mean temperatures of this coastal tropical site vary from 24° in winter to 30° in summer, they did not look at the influence of freeze-thaw cycles on rockfall activity at this site but focused instead on the influence of rainfall, cumulated rain, maximum and minimum daily temperature and their variations. In [26], it is evidenced that rain and cumulated rain are the only statistically significant meteorological triggers of rockfalls on the RN1 road (assuming that the apparent relationship with temperature variations is mainly attributable to the fact that temperature and rain are both seasonal events and do not vary independently). The effect of a rainfall event on rockfall activity is maximum one day after the rainfall and remains significant for up to 3 days after the rainfall. The dependency of rockfall activity on rain is particularly obvious when compared to other sites, which is explained by the fact that rain can be very intense in this tropical environment affected by cyclones. La Réunion island holds world records for rainfalls, ranging from 500 mm in 3 h to 6083 mm in 15 days during Cyclone Hyacinthe in 1980 (source: Météo France Event).

2.5. Operational Management of the Study Site

In addition to protective works, operational rules have been implemented since 1976 to limit the level of risk and have been optimized continuously ever since [41]. Thus, on days of excessive rainfall or when rainfall accumulates, the road is partially closed (two lanes along the cliff) or completely closed (four lanes) to avoid accidents. The closure is decided for a fixed duration whenever the accumulated rainfall within 24 h, or over 48 to 72 h, exceeds a fixed threshold. Although the values of the rain thresholds have changed depending on the protective works set up, the general principles of the rule remain the same. A detailed description of these rules is given in Section 3.5.

3. Data

3.1. Original Data

Three rain gauges were installed in 1998 along the road (Pointe du Gouffre KP 5.5, Grande Chaloupe KP 8.5, La possession, KP 13, see Figure 2c) to contribute to the operational management of the RN1 road. The rainfall time series (R_t, Table 2) is composed of the maximum values recorded by the three daily rain gauges. These three values are highly correlated, and R_t shows a high correlation with each series of measured rainfall (Table 3). We chose to use the maximum daily value rather than an average value or the value of the rain gauge closest to the event because of the following reasons: (1) the operational rule for road management has been based on the maximum value of the three rain gauges since 1999; (2) the value of the rain gauge closest to the landslide can be used to build the model but cannot be used for prediction (as we cannot predict where the rockfall will occur).

Table 2. Raw rainfall and rockfall time series.

	Rainfall Time Series	Rockfall Time Series		
Time	R_t	B_t	N_t	M_t
t_1	r_1	b_1	n_1	m_1
...				
t_i	r_i	b_i	n_i	m_i
...				
t_N	r_N	b_N	n_N	m_N

Table 3. Correlation matrix for rainfall measurements and the maximum daily rainfall at 3 rain gauges located along the RN1 road.

	Possession	Pte_du_Gouffre	Gde_Chaloupe	Max Value
Possession	1	0.85	0.9	0.93
Pte_du_Gouffre	0.85	1	0.93	0.94
Gde_Chaloupe	0.9	0.93	1	0.95
Max value	0.93	0.94	0.95	1

In the database, only the rockfalls that reached the road are referenced, including the following information (Figure A3):

- The date of the event;
- Its approximate mass: each rockfall event can consist of one stone, one rock, several rocks or large rocky landslides;
- Its location, i.e., the landmark KP at which the rockfall reached the road, as well as which side of the road was hit (mountain lanes or sea lanes).

From this raw database, rockfall time series (B_t, N_t, M_t) are formed (Table 2) by the series of three properties b, n and m, defined as follows:

- A Boolean b indicating the occurrence of rockfalls (0: no rockfall, 1: at least one rockfall);
- The number n of rockfalls (from 0 to 22);
- The sum m of their masses.

3.2. Descriptive Analysis of the Raw Rainfall and Rockfall Time Series

Figure 3 shows the daily time series of rainfall R_t and the number of rockfalls N_t. We can see data gaps. Between 26 February 2004 to 1 January 2005, no rockfall data were recorded. Then, as explained in Section 2.3, major works were achieved from August 2006 to December 2007 and from March 2008 to December 2008 to secure the cliff. Thus, the data between August 2006 and December 2008 are not displayed. We can notice that, as expected, the number of rockfalls hitting the road decreased considerably after the building of protective works. As the behavior of the cliff has changed considerably due to the building of protective structures, two separate analyses are needed: the first analysis will cover the period between 1 January 2000 and 31 July 2006, prior to the building of protective works; and the second analysis will cover the period between 1 January 2009 and 17 May 2018, after the building of protective works. Since many more rockfalls impacted the road during the first period, we started to build predictive models on this period. We decided to process the data from the second period in a second step to evaluate the capabilities of the tested models on a dataset with fewer events.

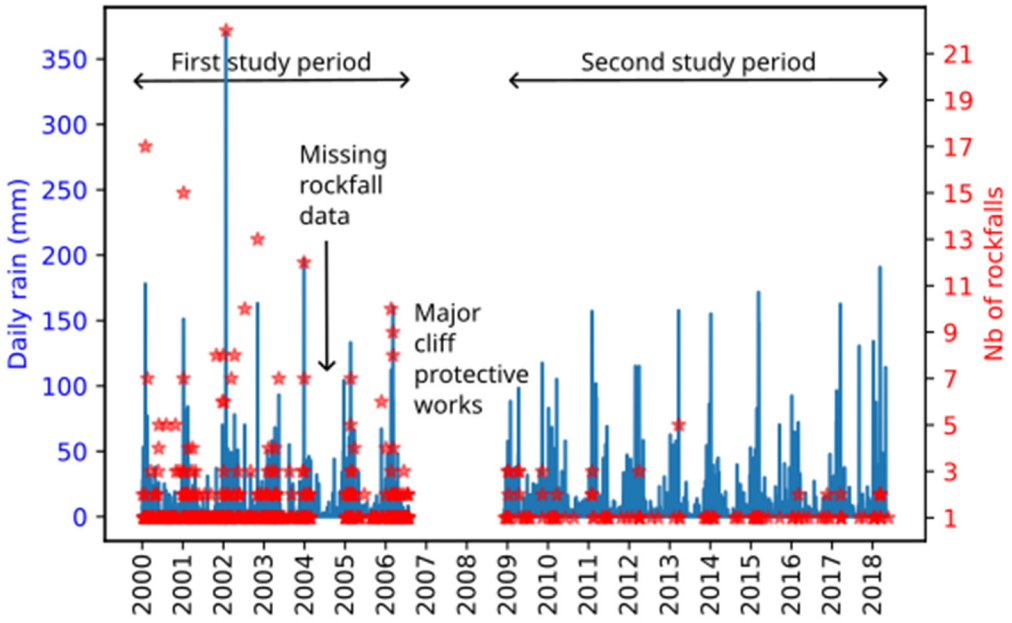


Figure 3. Rainfall (blue) and rockfall (red star) time series (respectively, R_t and N_t) from 2000 to 2018.

Most days are free of rockfalls. Rockfalls are therefore rare events, and we work in this context. On most days with rockfalls, there is only one rockfall. In a few cases, there may be as many as 22 rockfalls on the same day during the first period and 5 rockfalls on the same day during the second period. During the first period, days with rockfalls with a mass between 10 and 50 kg are the most common. These masses are associated with volumes between 4 and 20 L, large enough to cause accidents. There are 24 days with rockfalls with a mass higher than 1000 kg and 99 with rockfalls weighing less than 5 kg. During the second period, we also observed 24 days with rockfalls weighing more than 1000 kg and 25 with a rockfalls weighing less than 5 kg. Over the second period, there are relatively more days with higher masses.

Over both study periods, a classical analysis is performed, examining the cross-correlation between the number of rockfalls and rainfall. It shows a significant correlation for delays, respectively, between 0 and 3 days and between 0 and 5 days (Figure 4). The maximum correlation is obtained for the rain of the day before, as evidenced in [26]. The high correlation with the rain of the current day is explained by the daily scale of the data and the fact that when rockfalls occur at the end of the day, they can be linked to the rain that fell at the beginning of the day.

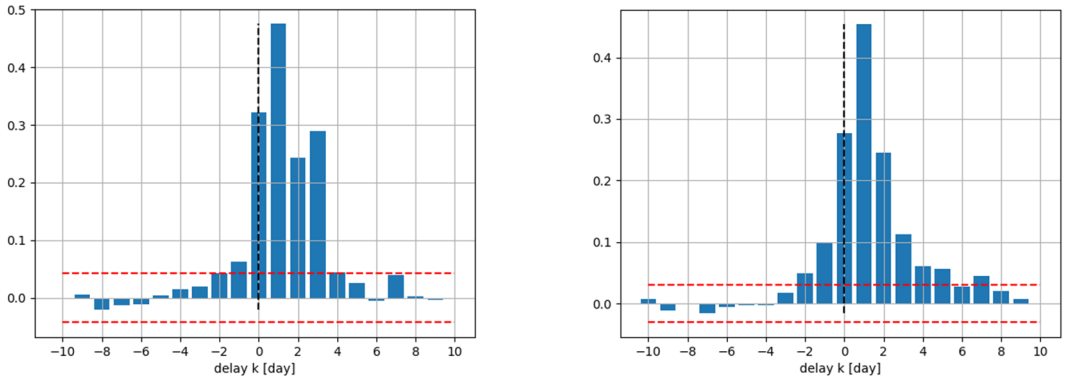


Figure 4. Cross-correlation between daily rockfalls and rainfall delayed for several delays: (**a**) for the 2000–2006 period and (**b**) for the 2009–2018 period.

Even if rockfalls and the rainfall of previous days are correlated, some rockfalls occur during periods with low rainfall, as specified in the following. In Figure 5, an example of the distribution of the maximal value of the rainfall of the three previous days, i.e., max(r_{i-1}, r_{i-2}, r_{i-3}), is displayed only for the days with at least one rockfall. Between 2000 and 2006, 372 days with rockfalls are recorded, including 220 days for which max(r_{i-1}, r_{i-2}, r_{i-3}) is less than 15 mm and 155 days for which max(r_{i-1}, r_{i-2}, r_{i-3}) is less than 5 mm. Over the period 2009–2018, there are 100 days with rockfalls, including 27 days for which max(r_{i-1}, r_{i-2}, r_{i-3}) is less than 5 mm and 43 days for which max(r_{i-1}, r_{i-2}, r_{i-3}) is less than 15 mm. This result confirms the work of [41], covering the period 1998–2002. Thus, some rockfalls are not related to rain triggers, and this constitutes noise in the rockfall time series. It will therefore not be easy to predict the occurrence of such rockfalls.

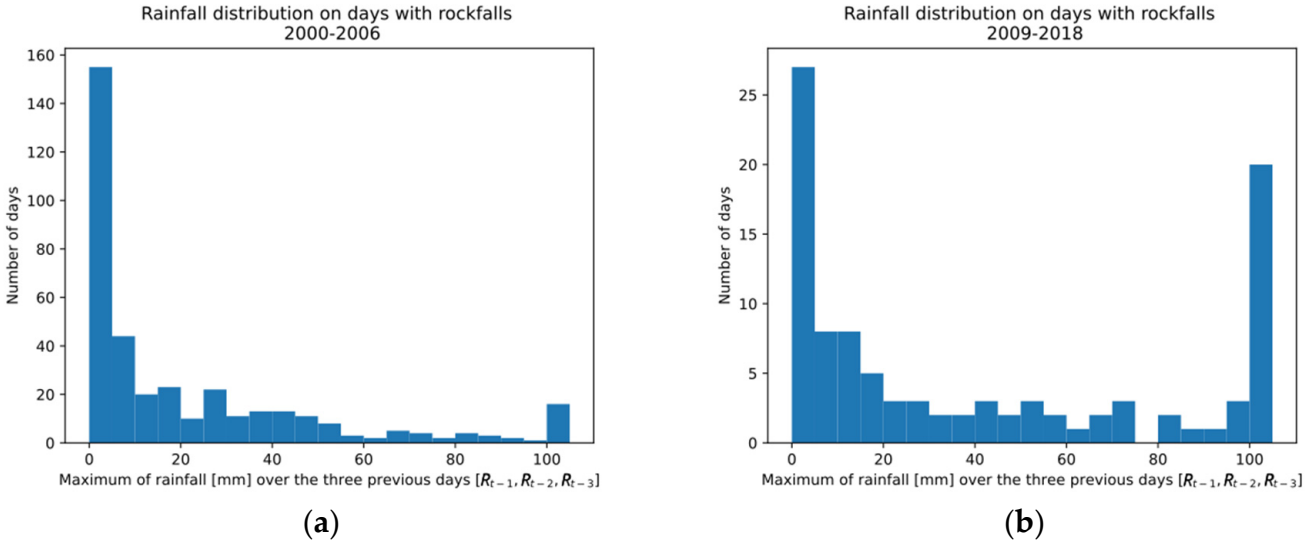


Figure 5. Rainfall distribution on days with rockfalls: (**a**) from 2000 to 2006, (**b**) from 2009 to 2018.

Figure 6 displays the rainfall conditions (max(r_{i-1}, r_{i-2}, r_{i-3})) of days with rockfalls for the different rockfall numbers per day (Figure 6a,c) and for the rockfall mass classes (Figure 6b,d). To facilitate understanding, we provide two examples: (1) between 2000 and 2006, 66% of days with one rockfall occurred when the max(r_{i-1}, r_{i-2}, r_{i-3}) was less than 15 mm (Figure 6a); (2) between 2009 and 2018, 25% of days with rockfalls with a mass between 5 and 10 kg occurred when the max(r_{i-1}, r_{i-2}, r_{i-3}) was less than 10 mm (Figure 6d). A significant change of behavior between the two periods can be observed:

- Period 2009–2018 presents a lower percentage of days with rockfalls for a specified value of max(r_{i-1}, r_{i-2}, r_{i-3}) than the period 2000–2006.
- During periods of low rainfall (values less than 10 mm), up to five rockfalls can occur on the same day for the period 2000–2006 and only one for the period 2009–2018.
- No conclusion can be drawn regarding the masses of the rockfall: for both periods, rockfalls can have various masses, and some are greater than 1000 kg.

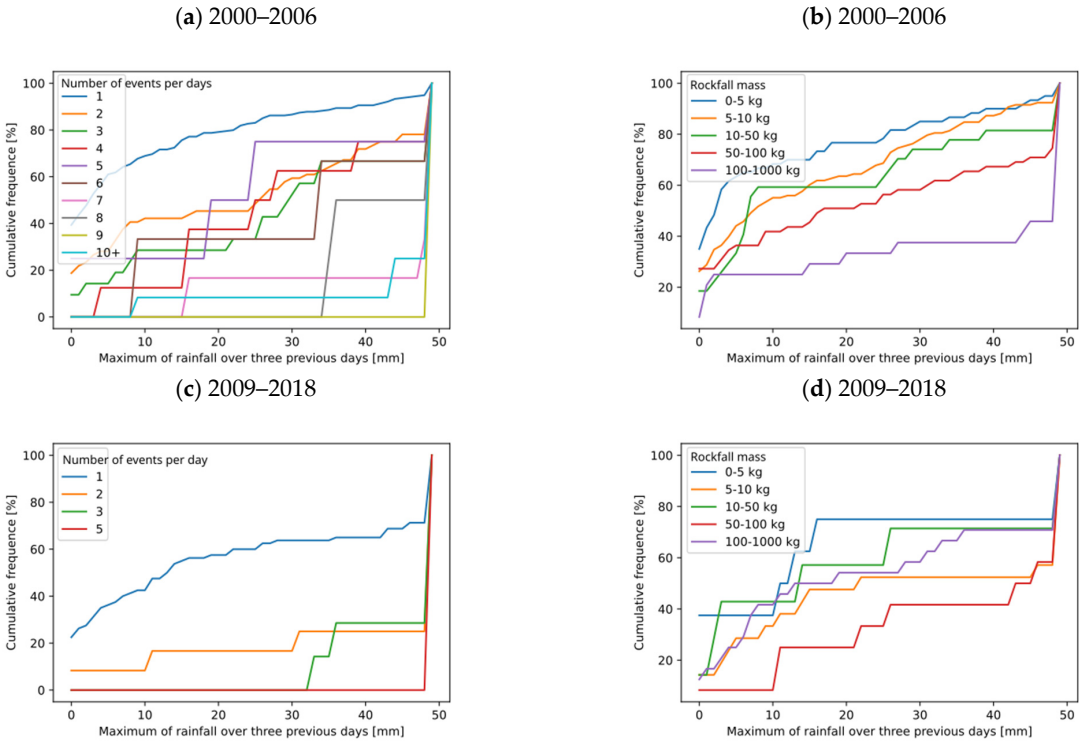


Figure 6. Rainfall conditions on days with rockfalls depending on the number of events per days, as observed in (**a**) the 2000–2006 period and (**c**) the 2009–2018 period. Additionally, they are dependent on the rockfall mass in (**b**) the 2000–2006 period and (**d**) the 2009–2018 period.

3.3. Features Selection

Here, ML models aim to predict rockfall occurrence (B_t). The number of rockfall events N_t and their mass M_t will be used to qualitatively evaluate predictions and are not used to build models. Later on, these data could be used to predict rockfall masses.

As proposed in [42], raw data must be structured to be integrated as input data for ML models. From the raw time series derived from the available databases and according to the results of the descriptive analysis, we select $p + k$ features to predict the occurrence of rockfall B_t (Figure 7), with $p \geq 1$, $k \geq 1$:

- p rainfall features obtained by lagging the rainfall time series R_t with delays between $[1..p]$. The p lagged rainfall time series are denoted as $R_{t-1}, \ldots, R_{t-p}$.
- k rockfall features obtained by lagging the rockfall time series B_t with delays between $[1..k]$. The k lagged rockfall time series are denoted as $B_{t-1}, \ldots, B_{t-k}$.

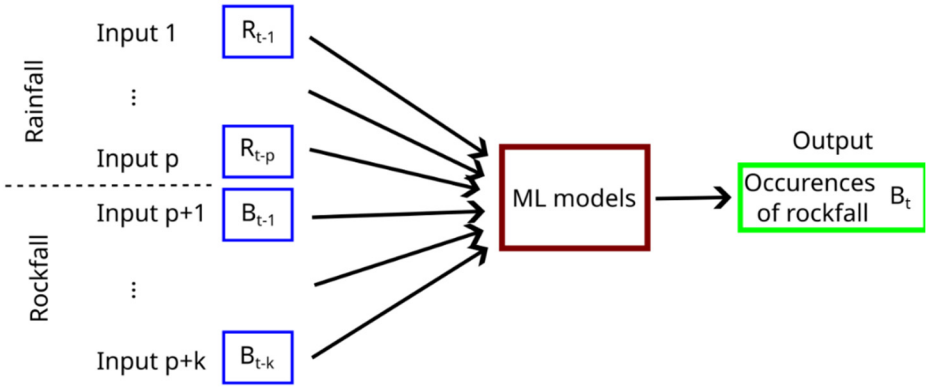


Figure 7. Selected features for ML models.

Other features could be engineered, such as cumulative rainfall over several days or differences in rainfall compared to the previous day. Here, the choice has been made to use the rainfall of the p previous days and the occurrence of rockfalls of the k previous days to predict rockfall occurrences (Figure 7). The p and k values can differ from one another and from one period to the next.

3.4. Definition of Training, Validation and Test Sets

To be used properly, ML models need a set to be trained on (the so-called "training or learning data") and then a set to compare the ML results with the measured ones (the so-called "test data"). See Appendix B.5 for further explanations. Thus, we need to split the data from any study period into a learning set and a test set to optimize and build the model and compute its performance. For the period 2000–2006, the lack of data between 26 February 2004 and 1 January 2005 is a natural cut to define the learning set from 1 January 2000 to 25 February 2004 and the test set from 2 January 2005 to 31 July 2006. The length of the test set represents 27% of the length of the period. In addition, a Kolmogorov–Smirnov test is performed on these two sets to determine if they have similar distributions and to determine the relevance of using these data as training and test sets. The Kolmogorov–Smirnov test is performed for the daily number of rockfalls averaged per month. The test statistic is equal to 0.24, and the corresponding p-value is 0.34. Since the p-value is higher than 0.05, we can assume that the two sets come from the same distribution and can be used as learning and test sets.

The study period 2009–2018 is also split into a learning set from 1 January 2009 to 31 December 2015 and a test set from 1 January 2016 to 17 May 2018. In this case, the length of the test set represents 25% of the overall period. The Kolmogorov–Smirnov test is also performed to compare the distribution of both sets. The test statistic is equal to 0.05, and the corresponding p-value is 0.99. As previously mentioned, the p-value is higher than 0.05, so we can assume that the two sets have similar distributions and can be used as learning and test sets.

Table 4 presents the following information for each data set: the period duration, the total number of rockfalls, the number of days without rockfall (class 0) and with at least one rockfall (class 1). The two classes (class 0 and class 1) are imbalanced. During the first period (2000–2006), rockfalls occur on about 18% of the days, and during the second period (2009–2018), they occur on only 3% of the days. For the development of predictive models, this imbalance should be taken into account.

Table 4. Study period characteristics.

Study Period		Start and End Date	Period Duration (Days)	Total Number of Rockfalls	Number of Days without Rockfalls (Class 0)	Number of Days with Rockfalls (Class 1)
1	Total	1 January 2000–31 July 2006	2092	696	1722	370
	Learning set	1 January 2000–24 February 2004	1516	527	1237	279
	Test set	2 January 2005–31 July 2006	576	169	485	91
2	Total	1 January 2009–17 May 2018	3424	131	3323	101
	Learning set	1 January 2009–31 December 2015	2556	98	2482	74
	Test set	1 January 2016–17 May 2018	868	33	841	27

3.5. Expert Models

As explained in Section 2, operational rules have been implemented since 1976 to limit and avoid accidents and deaths. The operational rules were optimized in 1983 with the setup of rain gauges, and further enhancements were made in 1999, 2002 and 2009 with the setup of protective works and the development of statistical studies [36]. The operational rules are defined by a delay of closure of the two lanes along the cliff whenever the 24 h cumulative rainfall exceeds a fixed threshold (Table 5). In 2004, if the cumulated rain over

the last 24 h was between 15 and 30 mm, the two mountainside lanes were held closed for the next 24 h, and if the cumulative rainfall was greater than 30 mm, the lanes remained closed for 72 h. Considering the setup of numerous protective structures, the cumulated rain thresholds were both raised, and the closure times reduced in 2009. The expert rules used in 2004 and 2009 are, respectively, used as reference models to compare the results of the ML models, respectively, over the first and second periods.

Table 5. Description of expert rules used for the management of the road during the study periods.

Definition Date	Cumulated Rain in 24 h (CR_{24h})	Delay of Closure
2004	15 mm < CR_{24h} < 30 mm	24 h
	30 mm < CR_{24h}	72 h
2009	30 mm < CR_{24h} < 50 mm	24 h
	50 mm < CR_{24h}	48 h

4. Calculation and Results

The features and predictions of the ML models were explained in Section 3.3. We recall that the input features are the p previous rainfall quantities $\{R_{t-i},\ i = 1..p\}$ and k previous occurrences of rockfalls $\{B_{t-j},\ j = 1..k\}$, with $p, k > 0$. The ML models are aimed to predict the occurrence of rockfalls B_t. The structure of each ML model and the hyperparameter optimization method are presented in Appendix B. In this section, we will detail selected parts of the calculations and results.

4.1. Period 2000–2006

The performance of both the expert model and the four ML models is specified in the following section, with a focus on optimizing their architecture. The delay values of p and k tested for the optimization of the four models of the first study period are the following: $p \in [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15]$ days and $k \in [1, 3, 5, 7, 9]$ days. These values are chosen inside the physical range of these parameters.

4.1.1. Expert Model

The expert rule is already defined (see Section 3.5) and can be directly computed for both the training and test sets. The results are given in Table 6. We can see that the F1-score (F1-Sc) and the balanced accuracy (BA) values are similar for both the training and test sets. The recall (Rec.) value is higher for the test set, and, conversely, the precision (Prec.) value is higher for the training set.

Table 6. Performance of the expert model—study period 2000–2006.

Data Set	TP	FP	FN	TN	F1-Sc	Rec.	Prec.	BA
training	98	83	181	1154	0.42	0.35	0.54	0.64
test	38	43	53	442	0.44	0.42	0.47	0.66

4.1.2. Bagged Trees

As explained in Appendix B.2, the optimization of the bagged trees is mainly conducted for two hyperparameters: the max depth (called max_depth) of each decision tree and the minimum value of sample per leaf (called min_samples_leaf). The number of estimators for the bagging approach is set to 200 because it is a commonly accepted value for bagging [43]. The range of variation of the maximum depth is [2,3,4,5,6,7,8], and that of the minimum values of samples per leaf is [5,10,20]. The values of the ranges are chosen to be coherent with the length of data (3000 samples maximum in the training or test sets).

In addition, the two classes (0: no event of rockfall, 1: at least one rockfall) are imbalanced because there are significantly more days without rockfalls (~80%) than with

rockfalls (~20%). We test two configurations of class weights $\{0 : w_0, 1 : w_1\}$: no weight, i.e., $w_0 = w_1 = 1$ or weights proportional to real proportions of days with and without rockfalls $w_0 = 1$ and $w_1 = 4$. Bagged trees have been implemented with the scikit-learn package in Python [44].

The best results are obtained with the class weights inversely proportional to their frequency ($w_1 = 4$) and are explained below. The best F1-score (0.495) is obtained for a rain delay of $p = 4$ days and a rockfall delay of $k = 9$ days, max_depth = 4 and min_samples_leaf = 20 from the training/validation set. With this optimized bagged decision tree, we obtain the following values for metrics: 0.43 for the F1-score, 0.66 for balanced accuracy, 0.42 for precision and 0.44 for recall for the test set. In this case, we take into account the previous occurrences of rockfall. The same optimization without taking the previous occurrences of rockfall ($k = 0$) into account led to similar results: the best F1-score (0.479) is obtained for a rain delay of $p = 5$ days, max_depth = 2 and min_samples_leaf = 20 from the training/validation set, and we obtain 0.43 for the F1-score, 0.66 for balanced accuracy, 0.44 for precision and 0.43 for recall for the test set (Table 7).

Table 7. Evaluation of optimized models—value computed on testing data (period 2000–2006).

Model	Optimum Hyperparameters	TP	FP	FN	TN	F1-sc	Rec.	Prec.	BA
Expert model		38	43	53	442	0.44	0.42	0.47	0.66
Nearest Neighbors	p = 1, k = 7, n_neighbors = 1 weight = max	38	207	53	278	0.23	0.42	0.15	0.50
Bagged Trees	p = 4, k = 9, n_depth = 4, min_samples_leaf = 20, w1 = 4	40	56	51	429	0.43	0.44	0.42	0.66
Logistic Regression	p = 3, weighted loss	45	66	46	419	0.45	0.49	0.41	0.68
DNN	p = 15, k = 7 weighted loss	49	70	42	415	0.47	0.54	0.41	0.70

4.1.3. Nearest Neighbors

We test variations of the two hyperparameters for the nearest neighbors: the number of neighbors (n_neighbors) and the method used to weight the data in the neighborhood. The number of neighbors varies from 1 to 20 to be coherent with the length of data (3000 samples maximum in the training or test sets). The explored weights are 'max', 'distance' and 'uniform'.

The best F1-score (0.414) is obtained for a rain delay of $p = 1$ day and a rockfall delay of k = 7 days, weight = 'max' and n_neighbors = 1 from the training/validation set. With this optimized nearest neighbor model, we obtain 0.23 for the F1-score, 0.495 for balanced accuracy, 0.15 for precision and 0.42 for recall on the test set. We can note that several values of the hyperparameters provide a high F1-score (between 0.4 and 0.41) on the validation set: either a low number of neighbors (1) with similar recall and precision values (0.6–0.7) or a higher number of neighbors (more than 8) with high recall (0.81) and low precision (0.28). The optimization, without taking the previous occurrences of rockfall (k = 0) into account, led to other results: the best F1-score (0.392) is obtained for a rain delay of p = 3 days, n_neighbors = 6 and weight = 'max' from the training/validation set and we obtain 0.38 for the F1-score, 0.625 for balanced accuracy, 0.5 for precision and 0.31 for recall for the test set. We can see that the metrics values on the test set, without taking into account the previous occurrences of rockfall, are higher than when taking them into account, even if the F1-score is lower on the training/validation set during the optimization process. The choice of hyperparameters is based on the value of the F1-score, calculated on the validation set, but it is possible to achieve better performances on the test set with other hyperparameter values.

4.1.4. Logistic Regression

In addition to time delays (p, k), we investigate class weights $\{0 : w_0, 1 : w_1\}$ in the log-likelihood to optimize logistic regression: no weighting, i.e., $w_0 = w_1 = 1$ or weighting proportional to the inverse of each class frequency in the training set: $\frac{1}{f_i \times nbclass}$. For the 2000–2006 period, such class weights are $\{0 : 0.6, 1 : 2.7\}$.

The function Logistic Regression from the scikit-learn package [44] is used to integrate class weight. When using unweighted cross-entropy loss, models give poor results, as they always predict the majority class. When using the weighted version, this method gives better performances with time delays p = 3 and no previous rockfall, i.e., 0.45 for the F1-score, 0.68 for balanced accuracy, 0.41 for precision and 0.49 for recall.

4.1.5. Neural Networks (NNs)

As mentioned in Appendix B, three types of NNs were examined for rockfall predictions: DNNs, CNNs and RNNs. As both the training and test sets have few positives events (i.e., days with rockfalls), we build simple networks with only one hidden layer to minimize parameters, using a dense layer for DNNs, a convolutional one for CNNs or a recurrent one for RNNs (the architecture is summarized in Table A1).

Before optimizing the models, classical hyperparameters are calibrated during some training phases: the number of epochs, size of the batch, learning rate, number of neurons in the hidden layer for DNNs.

In addition to time delays (p, k), different configurations of these NNs are tested:

- Rescale each series R_t and B_t between 0 and 1 when both are considered;
- No dropout or dropout (0.3) to prevent overfitting;
- Weighting or unweighted cross-entropy loss, similar to logistic regression.

The results indicate that better performance is achieved without rescaling and with weighted loss. Among all experiments with a weighted loss, without dropout and no rescale, we can keep the best configuration: a DNN model with a rain delay of p = 15 days and a rockfall delay of k = 7 days; the F1-score reaches 0.47, balanced accuracy is 0.70, recall is 0.54 and precision is 0.41.

4.1.6. Results Synthesis

The performances of both the expert model and the four ML models are summarized in Table 7. The expert model can identify 38 rockfalls out of 91 rockfalls (recall = 0.42) and predict 43 false rockfalls (precision = 0.47). Bagged trees, logistic regression and dense neural networks are able to identify more observed rockfalls (resp. 40, 45 and 49). This involves higher recalls and means that the security is improved for the road users. However, at the same time, more rockfall events are predicted, whereas no rockfalls are observed. These FP are higher (resp., 56, 66 and 70); that is why the three precisions are lower (resp., 0.42, 0.41 and 0.41). Thus, these models recommend closing the road more often than the expert model when there is no rockfall, and this penalizes more the economic interests. On the other hand, the nearest neighbor model predicts the same number of "true" rockfalls (38) but many more "false" rockfalls (207) compared to the expert model. This model also exhibits a lower recall, a significantly lower precision (0.15) and a lower F1-score. This model would greatly penalize the economic interests by closing the two mountainside lanes much more often. Thus, it is difficult to improve both recall and precision, and we can say that the three ML models (bagged trees, logistic regression and dense neural networks) provide similar performances to those of the expert model, with different compromises between recall and precision, with the expert rule falling between them.

In the expert rule, the relationship between rain and rockfall is simple but allows for the prediction of many rainfall-triggered rockfalls. However, this condition is neither sufficient nor necessary: on the one hand, these conditions do not always trigger rockfalls (FP), and on the other hand, other rainfall conditions can trigger rockfalls (FN). It was expected that ML models would find more specific relationships between rainfalls and

rockfalls so as not to increase the number of FPs. An increase of the number of TPs leads to an increase in FPs as well. This can be explained by the fact that the rockfall data are biased. We considered only rockfalls that reached the road and did not conduct a complete census of rockfalls from the cliff, which includs those that stopped before the road naturally or by protective structures. Some FPs could probably be TPs if all rockfalls were known.

4.2. Period 2009–2018

Results for the study period between 2009 and 2018 are shown in Table 8. The four ML models are optimized in the same way as for period 2000–2006. The expert model is able to identify 11 rockfalls over the 27 observed (recall = 0.41) and predict 42 false rockfalls (precision = 0.21). The bagged tree model has higher precision (0.31) and the same recall (0.41) as the expert model. Thus, it is as secure as the expert model for road users and, at the same time, penalizes less economic interests. In the first study period, the bagged trees exhibit higher recall but lower precision compared to the expert model. This performance trend is reversed in the second study period.

Table 8. Evaluation of optimized models—value computed on testing data (period 2009–2018).

Model	Optimum Hyperparameters	TP	FP	FN	TN	F1sc	Rec.	Prec.	BA
Expert model		11	42	16	799	0.28	0.41	0.21	0.68
Nearest Neighbors	p = 8, k = 7, weights = distance N_neighbors = 6	4	2	23	839	0.24	0.15	0.67	0.57
Bagged Trees	p = 8, k = 0, n_depth = 6, min_samples_leaf = 10, w1 = 10	11	24	16	817	0.35	0.41	0.31	0.69
Logistic Regression	p = 1, k = 5, weighted loss	14	98	13	743	0.2	0.52	0.12	0.70
DNN	p = 1, k = 7 weighted loss	11	50	16	791	0.25	0.41	0.18	0.67

The dense neural network also identifies the same number of rockfalls (11) as the expert model, but with a slightly higher number of "false" rockfalls (50). Its recall is the same and its precision is lower (0.18). Thus, it penalizes the economic interests by closing more often the road when no rockfall occurs. Contrary to the first period, the optimal dense neural network does not manage to predict more rockfalls compared to the expert model.

The logistic regression model identifies more rockfalls (14) and more "false" rockfalls (98) compared to the expert model. Its recall is higher (0.52) and its precision (0.12) is very low. It increases road safety by closing more of the two mountainside lanes when rockfall occurs, but it also penalizes economic interests by closing the road more frequently when no rockfall occurs. The prediction trend observed in the first period is amplified with respect to false positives.

The nearest neighbor model identifies only four rockfalls and predicts two false rockfalls: the recall is thus very low (0.15), and the precision is high (0.67). This performance is the opposite of the first study period. For this second period, the nearest neighbors exhibit a too small recall, i.e., they are not secure enough to be used operationally.

In the context of few rockfalls, the logistic regression model predicts most of the rockfalls (TPss) but at the expense of many false positives (FPs). The decision tree model and the dense neural network predict the same number of rockfalls (TPs) as the expert model while predicting, respectively, fewer and more FPs. The performance of the ML models is limited in this context due to the small number of rockfalls (3% of days with rockfalls) during the second study period with 3500 input data (2500 for the training set and 900 for the test set). Even though the architecture of the ML models has been defined based on the input data (1 single hidden layer with a maximum of 64 neurons for a DNN, for example), there are not enough events to allow the ML models to generalize their learning.

5. Discussion

Rockfall occurrences tend to cluster, both in time and in space. Small rockfalls were reported very early in the literature as precursors to larger rockfalls (a database with events from the 15th century is presented in [45]). In this case, the clustering consists of a spectacular increase in the number of small rockfalls, generally occurring in the days to minutes preceding a rockfall. On La Reunion Island, there is a specific French term to designate this increase in activity: it is said that the cliff "graine". These events are mainly located around the perimeter of the future scar of the main event. Longer and more detailed monitoring of cliffs has made it possible to highlight a temporal clustering of rockfalls over longer periods (i.e., several months to years, see [6,20,46]). This clustering is attributed to the redistribution of stresses in the rock mass following a rockfall. Consequently, the spatial clustering related to this phenomenon is generally very limited in space (at most, within a radius equal to the height of the cliff).

Given the amount and quality of the available data, investigating the long-term clustering of RN1 rockfalls seems unlikely. However, the short-term clustering of RN1 rockfalls was taken into account by using the rockfall delay (k). Indeed, precursory rockfalls, generally of small volumes, would normally be more systematically stopped in their propagation before reaching the road by the protective works. Thus, a different behavior was expected and is obtained between the first study period (prior the building of most protective works) and the second period (after their construction).

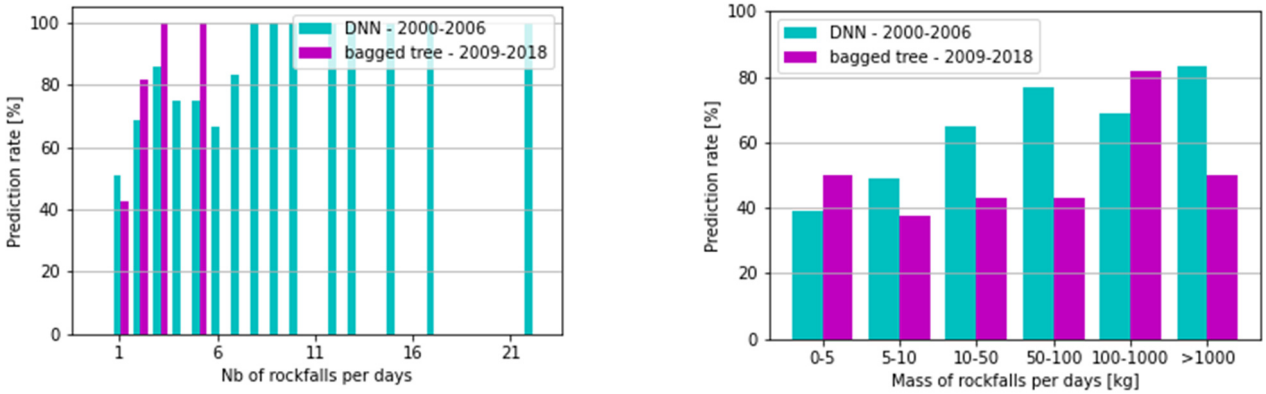
In the following section, the predictions of the ML models are analyzed, and some suggestions are proposed to improve the predictive ability of the ML models. Because of the poor performance of the nearest neighbor model, we decided not to use it anymore. We first investigate the characteristics of DNN-predicted rockfalls for 2000–2006 and bagged trees for 2009–2018, and then compare the predictions of the models (ML and experts) with the observed rockfalls only for 2000–2006. Finally, we analyze the temporal distribution of the predictions made by the ML models to see their implications for the operational management of the road in terms of the opening/reopening of the mountain side lanes.

5.1. Analysis of Rockfalls Predicted by the DNN (2000–2006) and the Bagged Tree (2009–2018)

We evaluate the whole dataset, assessing prediction performance based on TPs and miss-detections (FNs), according to both the mass of the events and the number of events per day. However, it is not possible to perform the same analysis for over-detections (FPs) because neither the mass nor the number of events are available since no rockfall has been observed.

As the DNN and the bagged tree models exhibit the highest F1-scores, respectively, in the first and the second periods, we test their ability to predict rockfalls for their respective period (Figure 8). The DNN and the bagged tree models are trained over the learning period as defined above, and we analyze the performance of the models evaluated over the entire study period (respectively, 16 January 2000–31 July 2006 and 16 January 2009–17 May 2018), which is different from the Section 4, where we evaluated the performance only on the test set. The prediction rate of the model is displayed as a function of the number of rockfalls per days (Figure 8a) and of the mass of rockfalls (Figure 8b).

Over the first study period, the DNN predicts all the days with rockfalls when the number of rockfalls is higher than seven and predicts better rockfalls when their masses are high. Only 39% of days with rockfalls between 0 and 5 kg are predicted whereas about 83% of days with rockfalls higher than 1000 kg are predicted. The greater the mass of the rockfalls, the better the prediction.



(**a**) (**b**)

Figure 8. Prediction rate of the DNN over 2000–2006 and bagged tree over 2009–2018 compared with the number of rockfalls per days (**a**) and the mass of rockfalls (**b**).

These results are in a good agreement with the results presented in Figure 6., both for the number of rockfalls and for the masses. We succeed in predicting most of the events triggered by rainfall, and a part of the unpredicted events corresponds to the events that occur on days with low or no cumulative rainfall (i.e., days for which the cumulative rain over the 3 previous days is inferior to 10 mm). During the initial period, the proportion of days with low rainfall decreases as the mass of rockfalls increases (Figure 6a,b). For the class 0–5 kg, there are 60% of days with rockfalls occurring during days with low (<10 mm) or no cumulated rain. For the class 10–50 kg, this percentage drops to 50% and for the class >1000 kg, it further drops to 25%. These values are nearly complementary to our predictions: for the class 0–5 kg, we predict rockfalls on 39% of days, for the class 10–50 kg, this percentage goes up to 65% of days and for the class >1000 kg, it further rises to 83% of days. Thus, small rockfalls occurring during dry periods (e.g., not triggered by rain) cannot be predicted by the DNN model.

Over the second study period, the bagged tree predicts all the days with events when the number of rockfalls is higher than two and predicts about 37% to 82% of days with rockfalls regardless of their masses. For this period, the percentage of correct predictions does not follow a monotonic increase with the rockfall masses. During the second study period, the percentage of days with rockfalls occurring during dry periods oscillates between 8 and 42% regardless of the rockfall masses (Figure 6c,d). Such behavior might explain why the model's ability to predict events does not seem dependent on the event masses during the second period. Indeed, the protective works built before the second period manage to prevent the propagation of small rockfalls down to the road, when most of the small events are not triggered by rainfall. Conversely, the protective works fail to stop the propagation of large rockfalls down to the road when most of them are triggered by rainfall.

Our results indicate that our models fail to relate a fair amount of the observed rockfall occurrences to any obvious trigger. This feature is consistent with findings in the worldwide literature (for example, about 50% of the rockfall occurrences in [23] are not related to any obvious trigger). Several hypotheses can explain our current inability to predict this residual rockfall activity:

1. It is difficult to demonstrate in a statistically reliable manner the link between rockfall occurrences and secondary triggering factors, since such events are relatively rare and represent a very small proportion of the analyzed databases. This argument is reinforced by the in situ instrumentations of [47] or [48], which showed that strong and cyclic temperature variations actively contribute to the propagation of cracks. It is therefore relevant to expect that this physical process contributes to the aging of

the La Réunion cliffs, even though this could not be demonstrated in a statistically robust manner.

2. The time delay between the start of an irreversible evolution towards failure following a triggering event and the actual time of failure is highly variable: from less than a few hours to several days [49]. The duration of this delay can vary depending on factors such as the failure mechanism, the volume of the instability and its constituent materials. The longer this delay, the more difficult it is to formally establish the link between a triggering factor and the occurrence of a rockfall, especially when no monitoring of precursory signs (opening of cracks, acoustic emissions, etc.) is available.

5.2. Comparison of the Model Predictions (ML and Expert Ones) with the Observed Rockfalls

In this part, we compare the detections of rockfalls (TPs) and the miss-detections (FNs) of the models among themselves and with the observed events in terms of both the number of rockfalls per day and their masses. Unfortunately, such an analysis cannot be conducted for over-detections (FPs) in the same way as the previous part. Figure 9 allows for the generalization of the conclusions obtained in the previous section for the DNN model (2000–2006) and bagged tree (2009–2018). Overall, the higher the hazard level (number of rockfalls per day), the more reliable are the predictions. This allows us to confirm that the triggering of rockfalls by rainfall is better predicted when several rockfalls occur on the same day.

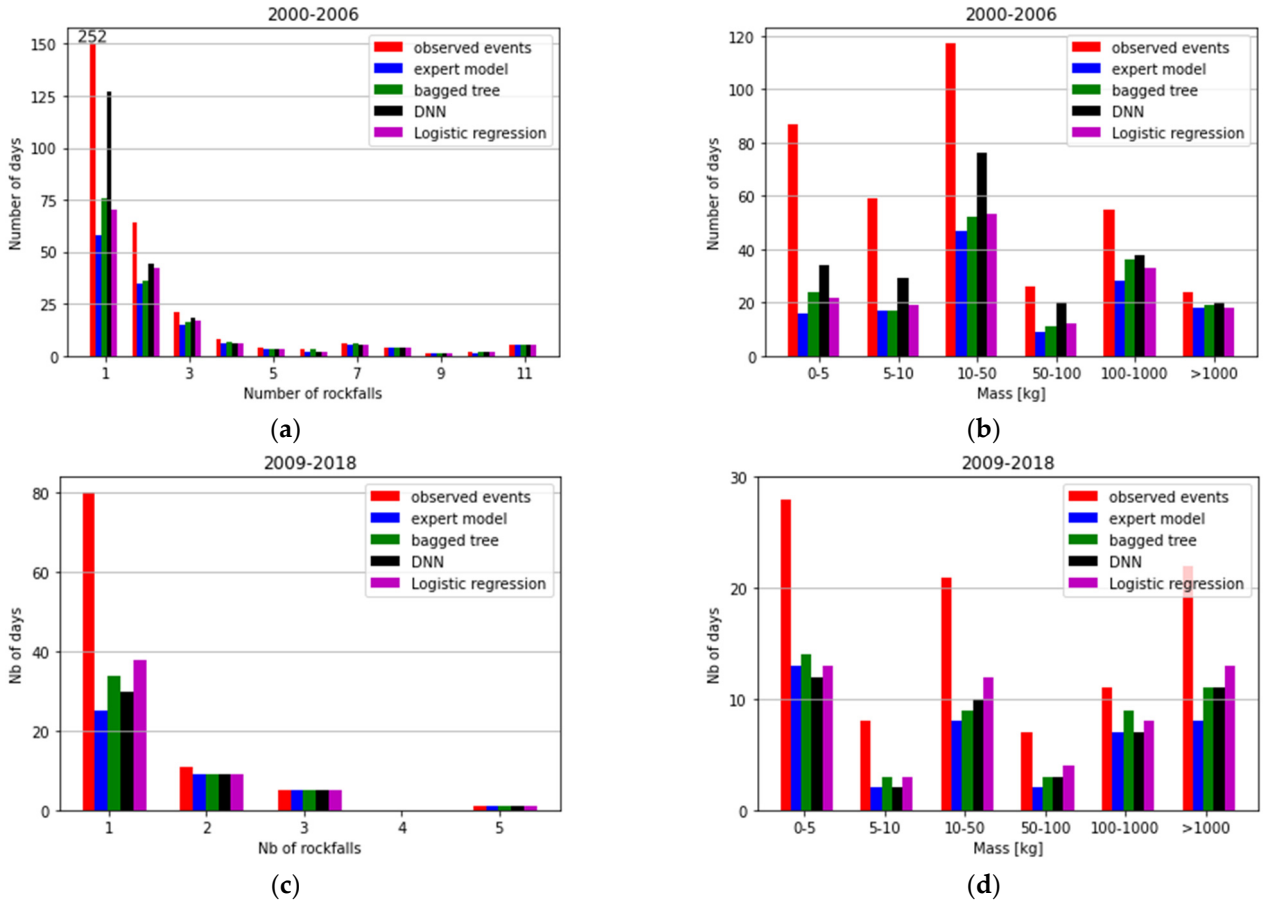


Figure 9. Comparison of the distribution of the number of rockfalls per day over the period 2000–2006 (**a**) and 2009–2018 (**c**) and comparison of the distribution of the mass of rockfalls over the period 2000–2006 (**b**) and 2009–2018 (**d**).

In detail, Figure 9 shows that bagged tree, DNN and logistic regression are slightly better in terms of TPs compared to the expert model (green, black and purple histogram bars are higher than the blue ones) regardless of the study period and the analyzed characteristics

(number of rockfalls or mass), except for only one class and one model: class 0–5 kg and the DNN model for the period 2009–2018 (12 TPs for DNN instead of 13 TPs for the expert model).

The DNN model gives the highest number of TPs for the first period. The prediction of rockfalls is significantly better for low numbers of rockfalls per day (1) or low masses (<100 kg), compared to the expert model or the other ML models. The relative complex structure of the DNN model allowed us to detect most rockfalls (TPs) in this context.

In the second period, the bagged tree and logistic regression display the highest number of TPs. In this case, there is no clear trend concerning the strengths of each model.

It can be noted that the expert model is a very simple decision tree, very constrained in its definition. The bagged tree can better capture the structure of the relationship between rainfall and rockfall with this general decision tree architecture and the choice of F1 score optimization.

From these results and analysis, we can propose some suggestions to improve the performance of the models:

1. The use of daily-scale data could hide some features of the relationships between rainfall and rockfalls. In particular, the triggering of rockfalls might occur only a few hours after the rainfalls. A test was performed by adding rainfall time series (R_t) to build bagged trees, even if this is not possible in real time. It showed better predictions of rockfalls for the F1-score (0.45), recall (0.47) and precision (0.43). This is in good agreement with the cross-correlation results in Figure 4, which show a significant correlation between the daily number of rockfalls and the rainfall of the same day. In addition, the ability of rainfall events to trigger rockfalls highly depends on the intensity and duration of these events. Two rainfall events of similar intensities but of different durations would not lead to the same consequences. For example, an event lasting only a few hours is more likely to trigger a rockfall compared to a rainfall spread across a whole day. In this context, using hourly data displays several benefits: (a) it allows for the separation of these two patterns of rainfalls, (b) it highlights the short-term triggering, (c) it provides more data to build ML models, and other models might be suitable in this case: SVM, gradient boosting or RNN.
2. Only a detection (binary classification) has been implemented so far. A classification into multiple classes, such as three classes (no rockfall, one rockfall, at least two rockfalls per day), could better describe the patterns and mechanisms associated with each class and allow for an increase in good detections (TPs) and a decrease in over-detections (FPs).
3. As mentioned before, the rockfalls are only those which reach the road. With a more exhaustive description of the rockfalls along the road, it would be easier to account for the relationships between rainfall and rockfalls and therefore predict more rockfalls (TPs).

5.3. Consequences for the Road Operational Management: Opening/Reopening of the Mountainside Lanes

In addition to qualifying predicted rockfall events in terms of the number of rockfalls per day or their masses, we can analyze the temporal distribution of the ML models' predictions. For this purpose, we define rockfall periods as consecutive days with rockfalls. The analysis is carried out during the period 2000–2006 because we have shown that it is more difficult for ML models to make reliable predictions when there are fewer data. The distribution of rockfall periods and the distribution of the time interval between rockfall periods are shown in Figure 10. Most rockfall periods inferred from the observed events are one day long, i.e., rockfalls do not occur on consecutive days, but the time interval between periods may be brief: only one or two days. The rockfalls predicted by the DNN model follow the same pattern, with a significant number of one-day rockfall periods. However, the number of rockfall periods longer than 3 days is slightly higher than that for observed events, and the time interval between rockfall periods is slightly shorter than for observed

events. The bagged tree tends to predict longer and more continuous rockfall periods, with about three times fewer one-day periods than for the observed events and slightly more rockfall periods longer than 3 days compared to the observed events. The time interval between rockfall periods is also higher than that for the observed events. This trend is accentuated for the expert model, with three times fewer one-day periods than for the observed events and very few two-day periods but many more periods of three days or more. The logistic regression falls between the DNN model and the bagged tree in terms of both the duration of periods with consecutive days of rockfalls and the time interval between these periods.

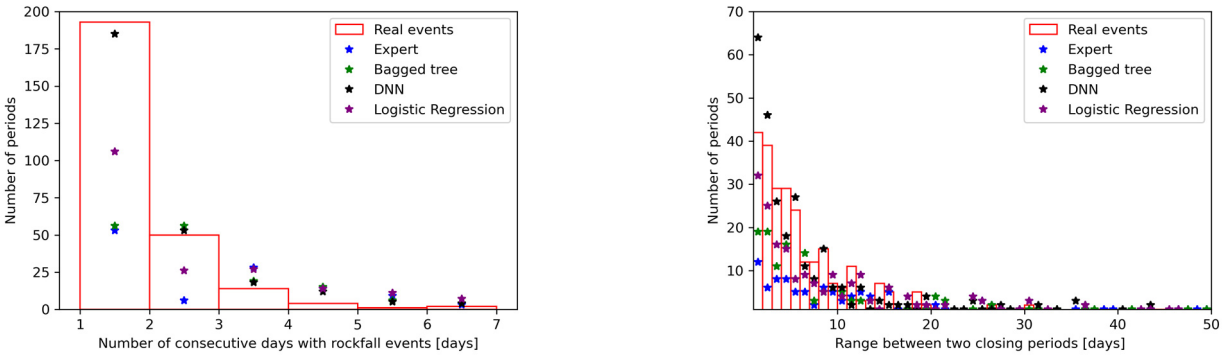


Figure 10. Distribution of the duration of periods with consecutive days with rockfalls (**left**) and the range between these periods (**right**) during 2000–2006.

From these observations, we can derive three descriptors related to the operational management of the road (Table 9): (1) the cumulative number of days with closures (related to the indirect socio-economic costs), (2) the number of closure periods (related to the operational cost) and (3) the mean duration of closure (related to the involvement time of the intervention team). The rockfall period predicted by the models is now called the closing period to consider this management objective. The cumulative number of closing days is obtained by adding the duration of all closure periods. The mean duration of road closures is obtained by dividing the cumulative number of days when the road was closed by the number of closure periods. In the case of real events, out of 368 days with rockfalls, there are 264 periods of closing. We can see that the predictions of the expert model are more continuous, with fewer closure periods and a higher mean closing duration (2.47 days). The bagged tree, DNN and logistic regression models have lower mean closure durations (between 1.92 and 2.13 days). The bagged tree provides the same number of closing days as the real events (about 368 days) but with fewer closure periods; the predictions are also more continuous. The DNN model and logistic regression predict more days of closure than the real events.

Table 9. Comparison of predictive models with the real events concerning the closing and reopening of the two mountainside lanes over 2000–2006.

2000–2006	Events Data	Expert Model	Bagged Tree	DNN	Logistic Regression
Nb of closure days	368 d	285 d	366 d	444 d	411 d
Nb of closure periods	264	115	159	277	193
Mean duration	1.39 d	2.47 d	1.92 d	1.98 d	2.13 d

Comparing the temporal distributions of rockfall predicted by each model can help the road manager to choose the model that fits best his strategy: more short closure periods or fewer long closure periods, depending on the way and cost of lane closing.

Inevitably, the models do not predict all events (FN) and predict more rockfalls (FPs), leading to road closing when rockfalls do not occur. But even if rockfall does not occur

one day, trigger conditions may exist, with a high hazard of rockfall, and it is relevant and wise to keep both lanes closed, especially when there is only one day between two closure periods. Many FPs are predicted by the ML models on these days between two real rockfalls. It may be worthwhile to add a new category of days called "warning days", which are days without rockfalls between two days with rockfall, to highlight the high risk and mitigate misclassification in ML models.

The purpose of the predictive models is to predict future rockfalls based on rainfall forecasts. Thus, from these rain forecasts, rockfall predictions can help to decide the following:

- When the road is open, close the road if the rainfall patterns identified by the ML models are present.
- When the road is closed, decide whether to reopen it depending on the potential occurrence of future rockfalls the following day or the day after.

The ML models thus allow us to anticipate the implementation of devices for the closing or the reopening of the road according to the criteria of the road manager. Opening a road one day and closing it the next can be costly due to the involvement time of the intervention team and equipment and can be misunderstood and misperceived by users who do not know which road to take. The results provided by ML models are obviously raw and can thus be post-processed to fit these management constraints.

6. Conclusions and Perspectives

Statistical models are sometimes used for the operational management of rockfalls, but, to our knowledge, the comparison of model predictions and their performances has not yet been formalized. In this paper, we adapted ML models to rockfall event prediction and compared their performance not only with a numerical approach but also by taking into account the operational consequences.

In detail, four ML models (nearest neighbors, bagged tree, logistic regression and dense neural networks) were built to predict periods of increased rockfall on National Road 1 on La Réunion island in the context of its high socio-economic stakes. Based on the hypothesis that rainfall is a unique meteorological trigger, features consisting of several lagged rainfall and rockfall time series were used to feed the ML models to predict the occurrence of rockfall. The predictions were compared with those of the expert model used for the operational management of the road. The performances of the ML models are similar or slightly better in terms of recall than the expert model, and we evidenced that ML models allow to correctly predict rockfalls actually triggered by rainfall events. Events with several rockfalls on the same day are rather well predicted. The analysis over two periods with 20% and 3% of days with rockfalls showed that the number of days with rockfall must be sufficient for the models to be able to generalize the pattern of the relationship between rainfalls and rockfalls from the training set to the test set.

The ML models have been straightforwardly implemented and optimized according to the classical criteria (F1-score). It allows for taking into account both road security (recall) and socio-economic costs (precision). However, sooner or later it will be possible to use the more general F_β-score to weight recall and precision differently according to the manager's strategy or to incorporate more complex criteria based on quantitative risk level assessment and the concept of acceptable risk, as developed by [50]. Regardless of the optimization criteria and the type of ML models used, tools have been developed to deduce the operational consequences of predictive models (duration and frequency of closure, duration and frequency of on-site interventions). The results can be directly integrated into road management to help decision making processes. In particular, they enable the implementation of mitigation measures whenever the hazard level is expected to increase (restriction of access, monitoring, mobilization of emergency kits...).

The ML models have a large number of FPs compared to the observed rockfalls. We can suggest several explanations. First, the rockfall database is biased and does not include rockfalls that did not reach the road. Second, some FPs reflect a risk situation based on rain

conditions, even though no rockfall occurs. This is, for example, the case of days without rockfall that occurs between two days with rockfalls. The FPs are not misinformative, but indicate that the risk may be high and can justify some transient mitigation measures, such as lanes closures.

Different ways can be proposed to improve the prediction of increased rockfall activity: (1) As the dynamics of rockfall are very different in low or high rainfall activity periods, it could be better to investigate switching models (from dry to rainy periods). (2) Developing a classification into multiple classes could better describe the patterns and mechanisms associated with each class. (3) Other triggering factors could be added as input to better describe triggering conditions for rockfalls. More information (temperature, opening of cracks, acoustic emissions, 3D surface displacements) could be helpful to assess the delay between the trigger and the failure of the rockfall. As said before, having more complete data concerning rockfalls from a site should allow for a better description of the meteorological patterns that trigger the rockfalls, thereby improving the models' performance. (4) Hourly data would also be better suited for the identification of rainfall triggering patterns (short intense rainfall or long low rainfall). (5) Concerning feature engineering, as ML models are designed to investigate more than linear correlations, average values between the three rain gauges per day, standard deviation within a day, differences in rainfall compared to the previous day or cumulative rainfall over larger time periods (until 1 month) could be used. (6) A spatio-temporal approach could be investigated by integrating new features, such as geology, soil properties, slope characteristics or the Topographic Wetness Index along the route. (7) Other methods, such as the SMOTE method, could be also explored to address the imbalance between classes.

Finally, how can machine learning help with the operational management of rockfall hazard along strategic linear stakes? We have seen that ML models can provide complementary prediction rules that can be fused with the expert ones. Machine learning approaches do provide a black box model which cannot be used directly for operational risk management rules, but it is possible to understand the underlying physics of the model using explainable AI methods to investigate which factors are influential and how. Then, when we are confident in the predictive ability of the models, we can translate the neural network results into potential socio-economic costs and operational risk management rules. Machine learning makes it possible to predict periods of increased rockfall activity, which can enable managers to implement agile risk management measures, such as restricting access or monitoring.

Author Contributions: Conceptualization, M.-A.C. and C.L.; methodology, M.-A.C., H.C., C.L., A.A., E.T. and M.G.-B.; software, M.-A.C. and H.C.; validation, C.L. and M.G.-B.; formal analysis, A.A.; investigation, C.L.; resources, L.M.; data curation, H.C.; writing—original draft preparation, M.-A.C. and H.C.; writing—review and editing, C.L. and M.G.-B.; visualization C.L.; supervision, M.G.-B.; project administration, M.-A.C.; funding acquisition, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FEREC Fondation.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is the property of the Conseil Régional La Réunion.

Acknowledgments: They also thank the Conseil Régional La Réunion for their support and for sharing their data. This work was carried out within the framework of the National Project C2ROP (www.c2rop.fr). Finally, the authors thank Julien Borderon from Cerema for his friendly and careful review.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A.

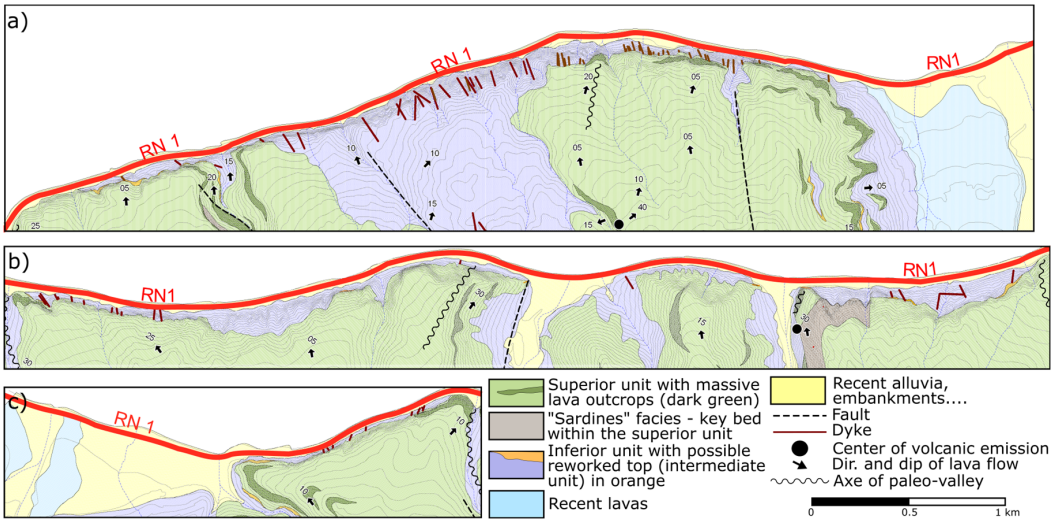


Figure A1. Geological map of the area around the RN1 road. For the location of the different sub-figures (**a**–**c**), see Figure 2. After [37–40].

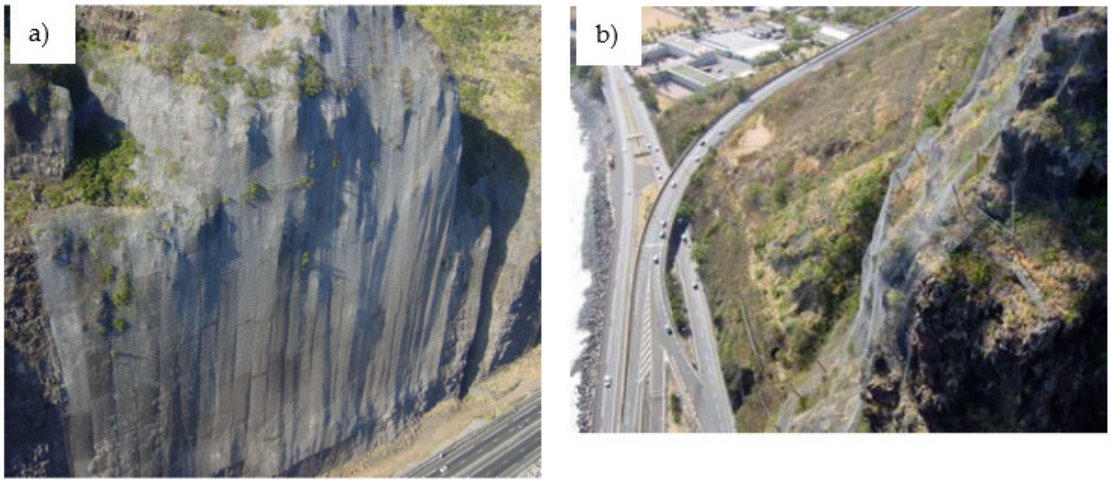


Figure A2. (**a**) Drapery systems set up in 1990 and (**b**) deflector net set up in 2016 along the national road 1 in La Réunion.

			CLIFF SIDE		SEA SIDE	
Date	**Time of entry**	**KP**	**Estimated mass kg**	**very small stones**	**Estimated mass kg**	**very small stones**
6 January 2000	02h00	11+700	15.0		1.0	
6 January 2000	02h00	12+800	40.0			
11 January 2000	8h00	8+100	3.0			
14 January 2000	10h15	10+900	2.0			
18 January 2000	9h20	4+600			5.0	
25 January 2000	05h30	12+700	6.0			
29 January 2000	02h15	4+400	40.0		20.0	
29 January 2000	10h55	7+900	3100.0			
30 January 2000	18h00	3+600	10.0			
30 January 2000	18h00	7+800	10.0			
30 January 2000	18h00	6+100		very small stones		

Figure A3. Extract from the raw rockfall database (source: RN1 Management Services).

Appendix B. Machine Learning Models

In this section, we present the families of models we used for event prediction, the way for optimizing the ML model architectures and the metrics used to evaluate and compare their performance.

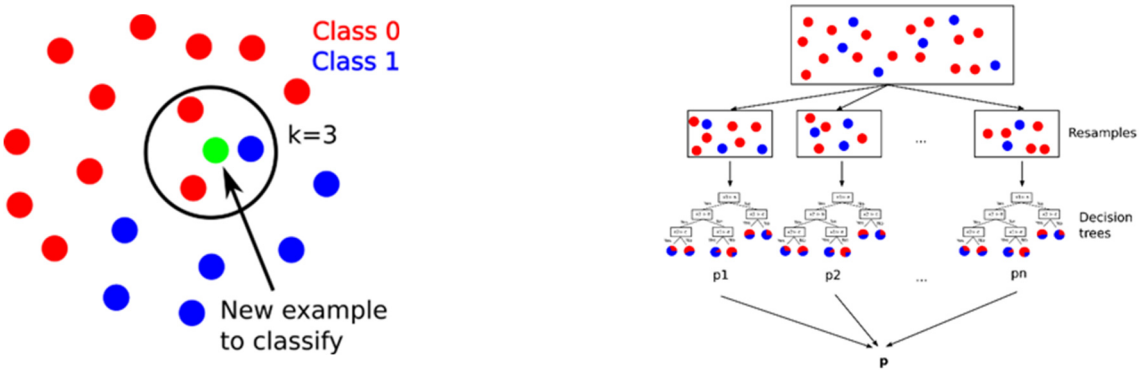
Appendix B.1. Predictive Models

In addition to the expert rule used for the RN1 management, four ML models originally developed for binary classification were adapted to perform rockfall event predictions: the nearest neighbor algorithm, bagged decision trees, logistic regression and neural networks. Random forests and gradient-boosted trees were also tested, but, with our data, the results were less satisfactory. In order to be concise and for the sake of designing an explainable model, we chose to limit our paper to the four mentioned models.

The architecture of each model is defined by hyperparameters that must be adjusted to avoid overfitting, where the model fits too perfectly to the training data and fails to find general trends, or underfitting, where the model is not able to find the general trends of the data. Then, once the architecture is defined, the parameters of the model are optimized to maximize the chosen performance during the training stage. The hyperparameters of the models are explained below, and the methods for optimizing the architecture of the models are described in Appendix B.5.

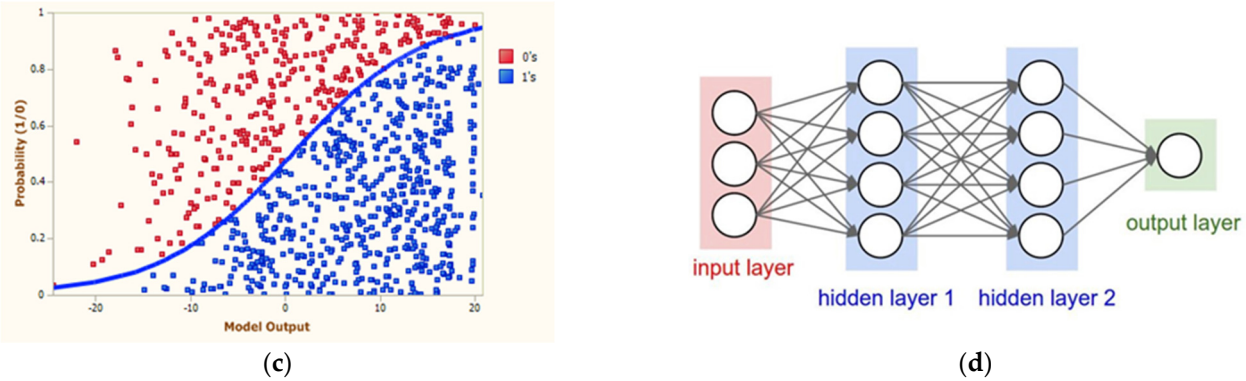
Appendix B.1.1. Nearest Neighbors

Nearest neighbors is a simple algorithm that stores all the available cases and classifies the new data or case based on a similarity measure with respect to the neighbors. Thus, when new data are added, the model will analyze the former data and observe a specified number of points which are closer to it. An example is displayed in a 2D space with three neighbors. The predictions change according to the definition of the measure. The simplest one is predicting the point chosen to be the most current among its neighbors. In this case, the weights of points are uniform (weights = uniform). In Figure A4a, for example, the prediction will be 0. Another possibility is weighting the points inversely to their distance (weights = distance). This means that closer points will have a larger weight. In the example, the prediction will be 1. The third measure consists in weighting the classes in order to give more weight to cases of the rarest class. Many weights can be used, but we test an infinite value for the rare case, i.e., if there is one case of this class in the neighborhood, the predicted value is the value of the rare class (weights = max). In Figure A4a, if we consider that class 1 is rarer than class 0, the prediction will be 1. Finally, to define the architecture of the nearest neighbor models, there are two hyper parameters to tune: the number of neighbors and the method of weighting.



(**a**) (**b**)

Figure A4. *Cont.*



(**c**) (**d**)

Figure A4. Principle of the four tested ML models: (**a**) nearest neighbor algorithm, (**b**) bagged decision trees, (**c**) logistic regression, (**d**) dense neural network with 2 hidden layers.

Appendix B.2. Bagged Trees

A decision tree model consists in a series of binary logical decisions in order to choose which category is the relevant output for the input data. At a node, each decision is simply based on the part of features transferred to this node. The decision rule can lead to other branches or to the final decision. The hyperparameters of decision tree models are the maximum depth, the minimum number of samples and possibly class weights. The maximum depth of the model stops it from growing excessively and limits overfitting. The minimum number of samples of a leaf stops branches extending with few samples, once more avoiding the model overfitting. The class weights are a way of taking into account imbalanced classes.

A bagged tree is an ensemble of decision trees (Figure A4b). The bagging method starts by separating the training data into smaller sets randomly created. Decision trees are formed for any of the subsets. The model is then a combination of those trees, and the prediction will be the combination of the predictions. Bagging aims to reduce the complexity of models and also prevents overfitting. It is a technique to improve predictive accuracy. The hyperparameters of bagged trees are those of decision trees in addition with the number of trees composing the ensemble and called number of estimators. As mentioned, it is possible to use hyperparameters of decision trees that can overfit elementary trees because we do not analyze each tree individually.

Appendix B.3. Logistic Regression

Logistic regression (Figure A4c) is a special case of the so called General Linear Model (GLM). GLM aims to model a target variable Y with respect to explanatory variables $[X_1, X_2, \ldots, X_p]$. More precisely, its conditional expectation, noted $\mathbb{E}[Y|\boldsymbol{X}]$, is given by $\mathbb{E}[Y|\boldsymbol{X}] = g(\beta \boldsymbol{X})$, where β is a vector of coefficients and $g(.)$ is the link function adapted to the probability law of Y. For instance, if $Y \epsilon \{0,1\}$ follows a Bernouilli law, it GLM corresponds to a logistic regression ($g = \frac{1}{1+e^{-\beta \boldsymbol{X}}}$). Different estimators, based on log-likelihood, are used to estimate coefficients β. When data are imbalanced, we can introduce class weights to balance sample importance in log-likelihood. The logistic regression model can also be seen as a specific case (one single neuron having p input synapses and a scalar output) of the general neural network framework described below.

Appendix B.4. Neural Networks (NNs)

From a statistical point of view, NNs are non-linear functions $f_\theta(.)$ estimated from data, which attempt to map inputs to targets, i.e., $\hat{Y} = f_{\hat{\theta}}(X)$. As illustrated in Figure A4d for a dense neural network (DNN), they are structured in multiple layers of neurons, each neuron of one layer being connected to the previous layer with a weight parameter, and followed by non-linear functions called "activation functions" (Table A1).

Table A1. Structure of used networks.

DNNs, CNNs or RNNs
• Inputs: X • hidden layer: – dense – or convolution 1d – or recurrent Cells • relu activation • dropout layer • dense final layer: B_t

Unlike classical estimators, neural networks possess millions of parameters, $card(\theta)$, which can be tuned thanks to their structure and effective parallel computing. A neural network is said to learn from data because of its operating steps. First, all learnable parameters θ of the model, such as connection weights and biases, are randomly initialized. Then, the model self-improves by gradually, seeking to find the best (X, Y) relationship by minimizing loss function $L\left(Y, \hat{Y}\right) = L_\theta$, which measures the error between predictions and true labels. Starting from model configuration θ_n, we can learn better weights θ_{n+1} from one labelled data (X, Y) of the training dataset:

- A forward pass: given the input X and the actual model configuration θ_n, a prediction $\hat{Y} = f_{\theta_n}(X)$ is computed;
- A backward pass: given the label Y and the prediction $\hat{Y}$, the loss L_θ and its gradient $grad(L_\theta)$ with respect to all trainable parameters θ are computed. Model configuration θ_n is updated, in the opposite direction of gradient to lower the loss function (therefore errors between prediction and labels): $\theta_{n+1} = \theta_n - l_r grad(L_\theta)$, where l_r, the learning rate is a scaling factor, tuned during learning.

These two points, forming a learning step, are repeated over the entire dataset, hence the so called training epoch. After one epoch, model capabilities are evaluated on the validation dataset. Depending on the results, some hyperparameters could be adjusted, such as the learning rate. To learn correctly, many epochs are repeated to reach a good predictive model. Other types of NNs have been investigated, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs,). In these networks, some layers involve multiple convolutions applied to inputs, with kernels learned from labelled data, even if those methods were designed for image or video analysis.

Appendix B.5. ML Models Architecture

The available data are divided into learning data to build the model and test data to evaluate the performance of the model. As the training of the model architecture (i.e., the value of the hyperparameters) cannot be performed using the test data, the learning data is therefore subdivided into training data (the first 80% of the data) and validation data (the last 20% of the data). Indeed, as we deal with time series, we prefer to train the models on data prior to those we are going to predict and for this reason, we do not do cross-validation with a division of the data into n parts. We evaluate the performance of the models by varying the value of the hyperparameters and represent the chosen metrics on the training and validation data with respect to these hyperparameters. The trained model is obtained when loss is minimal or when the chosen metrics are the highest on the validation set while being good on the training dataset. In that case, the model has surely learned, being neither overfitted nor underfitted.

Appendix B.6. Metrics

The metrics presented below are used to select the best hyperparameters for each of the four types of ML models and to compare their performance. Some of these metrics are

also used to analyze the operational consequences of using each model, in order to choose an acceptable compromise between the level of security for the stakes and the duration of road closures. This point is addressed in more detail in Section 5.

First, let recall classical metrics for hypothesis testing:

- Confusion matrix (see example in Figure A5) to have all information on correct and incorrect predictions. TP means true positive and refers to the points where class 1 is predicted as 1 (in our case study, days with rockfalls well detected). TN means true negative and refers to the points of class 0 predicted 0 (days without rockfalls well predicted). FP means false positive and refers to the points of class 0 predicted 1 (days without rockfalls detected as days with rockfalls). FN means false negative and refers to the points of class 1 predicted 0 (days with rockfalls detected as days without rockfalls). The training of the models aims at making TP and TN tend towards 1 and FP and FN towards 0.
- Accuracy, which measures the proportion of correct classification among all, $accuracy = \frac{TP+TN}{TP+TN+FN+FP}$,
- Precision of each class which measures the proportion of well classified samples among all predicted in this class, $precision = \frac{TP}{TP+FP}$. Poor class precision means that there are many over-detection (FP for binary) in this class,
- Recall of each class which measures the proportion of well classified samples among all real labelled in this class. In binary classification, recall usually indicates recall of positive class, $recall = \frac{TP}{TP+FN}$. Poor class recall means that there are missed detections (FN for binary) in this class.

Accuracy, precision and recall values vary between 0 and 1, and the optimal value is 1, but it is difficult to optimize recall and precision simultaneously.

When classes are unbalanced, other metrics are more suitable. In binary cases, like ours, negative samples (0 rockfall) are much more frequent in our dataset than positive sample (1 or more rockfalls).

- Balanced accuracy is the mean of positive and negative recall, $balanced\ accuracy = \frac{\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)}{2}$,
- F_β score, $F_\beta = \left(1+\beta^2\right)\frac{precision \times recall}{\beta^2(precision + recall)}$. When $\beta = 1$, F_1 is the harmonic mean of precision and recall and weights precision and recall equally. F_1 is the measure most often used when learning from imbalanced data [42]. For $\beta > 1$, F_β score puts less weight on recall and more weight on precision, if $0 < \beta < 1$ F_β score puts more weight on recall and less weight on precision.

Following the definitions above, an increase in recall, means an increase in road safety because more days with rockfalls are detected. An increase in precision means a decrease in road closure on days without rockfalls, which has direct consequences on the economy. Since infrastructure managers consider both aspects (safety and economy), we want to maximize both recall and accuracy. Here, we chose to select ML models from the F1 score in the following data processing, as it allows us to take into account both recall and precision.

		Predicted		
		0	**1**	Recall
Ground truth	**0**	TN	FP	$\frac{TN}{TN+FP}$
	1	FN	TP	$\frac{\mathbf{TP}}{\mathbf{TP+FN}}$
Precision		$\frac{TN}{TN+FN}$	$\frac{\mathbf{TP}}{\mathbf{TP+FP}}$	

Figure A5. Binary confusion matrix: 1 stands for at least one rockfall, 0 for no rockfall. TN = true negative, FP = false positive.

References

1. Varnes, D.J. Slope Movement Types and Processes. 1978. Available online: https://onlinepubs.trb.org/onlinepubs/sr/sr176/176-002.pdf (accessed on 25 April 2024).
2. Cruden, D.M.; Varnes, D.J. Landslides: Investigation and Mitigation. 1996. Available online: https://onlinepubs.trb.org/Onlinepubs/sr/sr247/sr247.pdf (accessed on 25 April 2024).
3. Hungr, O.; Leroueil, S.; Picarelli, L. The Varnes classification of landslide types, an update. *Landslides* **2014**, *11*, 167–194. [CrossRef]
4. Dussauge-Peisser, C.; Helmstetter, A.; Grasso, J.-R.; Hantz, D.; Desvarreux, P.; Jeannin, M.; Giraud, A. Probabilistic approach to rock fall hazard assessment: Potential of historical data analysis. *Nat. Hazards Earth Syst. Sci.* **2002**, *2*, 15–26. [CrossRef]
5. Hantz, D.; Dewez, T.; Levy, C.; Guérin, A.; Jaboyedoff, M. Rockfall frequency in different geomorphological conditions. In *International Symposium Rock Slope Stability*; Lyon, France, 2016; Available online: https://c2rop.fr/wp-content/uploads/2023/01/RSS2016_Proceedings.pdf (accessed on 25 April 2024).
6. Rosser, N.; Lim, M.; Petley, D.; Dunning, S.; Allison, R. Patterns of precursory rockfall prior to slope failure. *J. Geophys. Res. Earth Surf.* **2007**, *112*, 2006JF000642. [CrossRef]
7. Guzzetti, F.; Reichenbach, P.; Ghigi, S. Rockfall Hazard and Risk Assessment Along a Transportation Corridor in the Nera Valley, Central Italy. *Environ. Manag.* **2004**, *34*, 191–208. [CrossRef] [PubMed]
8. Volkwein, A.; Schellenberg, K.; Labiouse, V.; Agliardi, F.; Berger, F.; Bourrier, F.; Dorren, L.K.A.; Gerber, W.; Jaboyedoff, M. Rockfall characterisation and structural protection—A review. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 2617–2651. [CrossRef]
9. Prémaillon, M.; Regard, V.; Dewez, T.J.B.; Auda, Y. GlobR2C2 (Global Recession Rates of Coastal Cliffs): A global relational database to investigate coastal rocky cliff erosion rate variations. *Earth Surf. Dyn.* **2018**, *6*, 651–668. [CrossRef]
10. Frayssines, M.; Hantz, D. Modelling and back-analysing failures in steep limestone cliffs. *Int. J. Rock Mech. Min. Sci.* **2009**, *46*, 1115–1123. [CrossRef]
11. Paronuzzi, P.; Serafini, W. Stress state analysis of a collapsed overhanging rock slab: A case study. *Eng. Geol.* **2009**, *108*, 65–75. [CrossRef]
12. Paronuzzi, P.; Bolla, A.; Rigo, E. 3D Stress–Strain Analysis of a Failed Limestone Wedge Influenced by an Intact Rock Bridge. *Rock Mech. Rock Eng.* **2016**, *49*, 3223–3242. [CrossRef]
13. Levy, C.; Jongmans, D.; Baillet, L. Analysis of seismic signals recorded on a prone-to-fall rock column (Vercors massif, French Alps): Analysis of seismic events preceding a rockfall. *Geophys. J. Int.* **2011**, *186*, 296–310. [CrossRef]
14. De Vilder, S.J.; Rosser, N.J.; Brain, M.J. Forensic analysis of rockfall scars. *Geomorphology* **2017**, *295*, 202–214. [CrossRef]
15. Kobayashi, Y.; Harp, E.L.; Kagawa, T. Simulation of rockfalls triggered by earthquakes. *Rock Mech. Rock Eng.* **1990**, *23*, 1–20. [CrossRef]
16. Malamud, B.D.; Turcotte, D.L.; Guzzetti, F.; Reichenbach, P. Landslides, earthquakes, and erosion. *Earth Planet. Sci. Lett.* **2004**, *229*, 45–59. [CrossRef]
17. Yin, Y.; Wang, F.; Sun, P. Landslide hazards triggered by the 2008 Wenchuan earthquake, Sichuan, China. *Landslides* **2009**, *6*, 139–152. [CrossRef]
18. Hale, A.J.; Calder, E.S.; Loughlin, S.C.; Wadge, G.; Ryan, G.A. Modelling the lava dome extruded at Soufrière Hills Volcano, Montserrat, August 2005–May 2006. *J. Volcanol. Geotherm. Res.* **2009**, *187*, 69–84. [CrossRef]
19. DeRoin, N.; McNutt, S.R. Rockfalls at Augustine Volcano, Alaska: The influence of eruption precursors and seasonal factors on occurrence patterns 1997–2009. *J. Volcanol. Geotherm. Res.* **2012**, *211–212*, 61–75. [CrossRef]
20. Rosser, N.J.; Brain, M.J.; Petley, D.N.; Lim, M.; Norman, E.C. Coastline retreat via progressive failure of rocky coastal cliffs. *Geology* **2013**, *41*, 939–942. [CrossRef]
21. Young, A.P.; Guza, R.T.; O'Reilly, W.C.; Burvingt, O.; Flick, R.E. Observations of coastal cliff base waves, sand levels, and cliff top shaking. *Earth Surf. Process. Landf.* **2016**, *41*, 1564–1573. [CrossRef]
22. Brain, M.J.; Rosser, N.J.; Norman, E.C.; Petley, D.N. Are microseismic ground displacements a significant geomorphic agent? *Geomorphology* **2014**, *207*, 161–173. [CrossRef]
23. D'Amato, J.; Hantz, D.; Guerin, A.; Jaboyedoff, M.; Baillet, L.; Mariscal, A. Influence of meteorological factors on rockfall occurrence in a middle mountain limestone cliff. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 719–735. [CrossRef]
24. Ravanel, L.; Deline, P. Climate influence on rockfalls in high-Alpine steep rockwalls: The north side of the Aiguilles de Chamonix (Mont Blanc massif) since the end of the 'Little Ice Age'. *Holocene* **2011**, *21*, 357–365. [CrossRef]
25. Ravanel, L.; Magnin, F.; Deline, P. Impacts of the 2003 and 2015 summer heatwaves on permafrost-affected rock-walls in the Mont Blanc massif. *Sci. Total Environ.* **2017**, *609*, 132–143. [CrossRef]
26. Delonca, A.; Gunzburger, Y.; Verdel, T. Statistical correlation between meteorological and rockfall databases. *Nat. Hazards Earth Syst. Sci.* **2014**, *14*, 1953–1964. [CrossRef]
27. Macciotta, R. Review and latest insights into rock fall temporal variability associated with weather. *Proc. Inst. Civ. Eng. Geotech. Eng.* **2019**, *172*, 556–568. [CrossRef]
28. Macciotta, R.; Hendry, M.; Cruden, D.M.; Blais-Stevens, A.; Edwards, T. Quantifying rock fall probabilities and their temporal distribution associated with weather seasonality. *Landslides* **2017**, *14*, 2025–2039. [CrossRef]
29. Pratt, C.; Macciotta, R.; Hendry, M. Quantitative relationship between weather seasonality and rock fall occurrences north of Hope, BC, Canada. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3239–3251. [CrossRef]

30. Mitchell, S.A. *An Assessment of Rockfall Triggers and Seasonal Weather Trends through an Examination of Railway Slope Management Procedures*; Queen's University: Kingston, ON, Canada, 2021.
31. Melillo, M.; Gariano, S.L.; Peruccacci, S.; Sarro, R.; Mateos, R.M.; Brunetti, M.T. Rainfall and rockfalls in the Canary Islands: Assessing a seasonal link. *Nat. Hazards Earth Syst. Sci.* **2020**, *20*, 2307–2317. [CrossRef]
32. Bergen, K.J.; Johnson, P.A.; De Hoop, M.V.; Beroza, G.C. Machine learning for data-driven discovery in solid Earth geoscience. *Science* **2019**, *363*, eaau0323. [CrossRef] [PubMed]
33. Ghorbanzadeh, O.; Blaschke, T.; Gholamnia, K.; Meena, S.; Tiede, D.; Aryal, J. Evaluation of Different Machine Learning Methods and Deep-Learning Convolutional Neural Networks for Landslide Detection. *Remote Sens.* **2019**, *11*, 196. [CrossRef]
34. Blanco, L.; García-Sellés, D.; Guinau, M.; Zoumpekas, T.; Puig, A.; Salamó, M.; Gratacós, O.; Muñoz, J.A.; Janeras, M.; Pedraza, O. Machine Learning-Based Rockfalls Detection with 3D Point Clouds, Example in the Montserrat Massif (Spain). *Remote Sens.* **2022**, *14*, 4306. [CrossRef]
35. Berthet-Rambaud, P.; Guillemin, P. *Protocole et Banc D'essais en vue de Tester Différents Types de Filets Pendus ou Plaqués; Journées Nationales de Géotechnique et de Géologie de L'Ingénieur*; Lyon, France, 2006. Available online: https://www.cfmr-roches.org/sites/default/files/jngg/JNGG%202006%20s3%20pp%20169-176%20Berthet.pdf (accessed on 25 April 2024).
36. Azemard, P. The RN1 in Reunion Island: Rock Hazard Management: Circulation Management, Protections Works and New Route. In *International Symposium Rock Slope Stability*; Chambéry, France, 2018. Available online: https://c2rop.fr/wp-content/uploads/2023/01/RSS2018_proceedings.pdf (accessed on 25 April 2024).
37. Cruchet, M.; Lucas, E. *Projet Tram Train Régional: Expertise Géologique Complémentaire sur le Massif de La Montagne—Interprétation Géologique de Panoramas Photographiques de la Falaise du Littoral*; BRGM/RP-56178-FR; 2008; p. 20.
38. Courrioux, G.; Cruchet, M.; Chevalier, P.; Antemi, E.; Fontanet, C. *Projet Tram Train Régional: Etude Géologique et Hydrogéologique Préliminaire du Secteur Compris entre Saint-Denis et la Possession*; BRGM/RP-53744-FRb, BRGM REUNION; ANTEA REUNION; 2005.
39. Jaques, E.; Courrioux, G.; Cruchet, M.; Chevalier, P. *Projet Tram—Train Régional: Etude géologique du Secteur Compris Entre Saint-Denis et La Possession (Massif de La Montagne). Intégration des Données Complémentaires de Cartographie. Rapport N°2*; BRGM/RC-54335-FR; 2006.
40. Jaques, E.; Courrioux, G.; Antemi, E.; Piraud, J. *Projet Tram-Train Régional: Synthèse géologique le long de la variante de tracé par les Hauts (Massif de La Montagne). Intégration des Données Géophysiques et de Sondages, 70 p. 7 ann., 2 pht*; BRGM/RP-53744-FRb; BRGM REUNION; ANTEA REUNION; 2006.
41. Batista, D.; Binet, C.; Rat, M.; Alfonsi, P. *Etude Statistique de l'aléa Chute de Pierres Sur la RN1 à la Réunion, Géoline 2005: Géologie et ouvrages linéaires*; Arnould, Ledru, Eds.; BRGM éditions; Orléans, France, 2005.
42. Yang, B.; Nazari, R.; Elmo, D.; Stead, D.; Eberhardt, E. Data preparation for machine learning in rock engineering. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1124*, 012072. [CrossRef]
43. Cutler, A.; Cutler, D.R.; Stevens, J.R. Random Forests. In *Ensemble Machine Learning*; Zhang, C., Ma, Y., Eds.; Springer New York: New York, NY, USA, 2012; pp. 157–175, ISBN 978-1-4419-9325-0.
44. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Müller, A.; Nothman, J.; Louppe, G.; et al. Scikit-learn: Machine Learning in Python. *arXiv* **2012**, arXiv:1201.0490v4. [CrossRef]
45. Hermanns, R.; Oppikofer, T.; Anda, E.; Blikra, L.H.; Böhme, M.; Bunkholt, H.; Crosta, G.B.; Dahle, H.; Devoli, G.; Fisher, L.; et al. *Recommended Hazard and Risk Classification System for Large Unstable Rock Slopes in Norway*; NGU-rapport 2012.029; 2012. Available online: https://www.ngu.no/upload/Publikasjoner/Rapporter/2012/2012_029.pdf (accessed on 25 April 2024).
46. Stock, G.M.; Martel, S.J.; Collins, B.D.; Harp, E.L. Progressive failure of sheeted rock slopes: The 2009–2010 Rhombus Wall rock falls in Yosemite Valley, California, USA. *Earth Surf. Process. Landf.* **2012**, *37*, 546–561. [CrossRef]
47. Collins, B.D.; Stock, G.M. Rockfall triggering by cyclic thermal stressing of exfoliation fractures. *Nat. Geosci.* **2016**, *9*, 395–400. [CrossRef]
48. Gasc-Barbier, M.; Merrien-Soukatchoff, V.; Virely, D. The role of natural thermal cycles on a limestone cliff mechanical behaviour. *Eng. Geol.* **2021**, *293*, 106293. [CrossRef]
49. Intrieri, E.; Carlà, T.; Gigli, G. Forecasting the time of failure of landslides at slope-scale: A literature review. *Earth-Sci. Rev.* **2019**, *193*, 333–349. [CrossRef]
50. Fell, R.; Corominas, J.; Bonnard, C.; Cascini, L.; Leroi, E.; Savage, W.Z. Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Eng. Geol.* **2008**, *102*, 85–98. [CrossRef]