Ministerio del Poder Popular para la Salud
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648

Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412

Scientific Note

Estimation of potential epidemic risk in the landslide zone based on physical factors in the Sillapa district

Estimación de riesgo potencial epidémico en la zona de deslave basado en factores físicos en el distrito Sillapa

https://doi.org/10.52808/bmsa.7e6.626.036

Luis Eduardo Oré Cierto [1]
https://orcid.org/0000-0003-2836-2436
Ahnel Karen Zelaya Moya[2]
https://orcid.org/0000-0002-7371-8908
Wendy Caroline Loarte Aliaga[3]
https://orcid.org/0000-0001-9489-046X
Jorge Rafael Diaz Dumont [1,*]
https://orcid.org/0000-0003-0921-338X
Gianmarco Garcia Curo[1]
https://orcid.org/0000-0001-6685-3207
Adiel Alvarez Ticllasuca [1]
https://orcid.org/0000-0002-5410-0571
José Torres Huamaní[1]
https://orcid.org/0000-0001-9764-0749
Andres Olivera Chura [4]
https://orcid.org/0000-0001-6509-8991
Atilio Rodolfo Buendía Giribaldi [5]
https://orcid.org/0000-0002-8613-3730

Recibido: 24/09/2022

Aceptado: 16/12/2022

ABSTRACT

Almost 17% of causes of death due to natural hazards are the product of landslides. Most of them occur in the most deprived places of less developed countries, coexisting a lethal combination of factors that point to this type of tragedies: the natural and the human factor. On the other hand, after a disaster, health care needs and priorities may change; in this sense, the food security of refugees, the supply of drinking water, the disposal of excreta and solid waste, the need for shelters, attention to personal hygiene needs, vector control, attention to injuries after the cleanup activities and the conduct of public health surveillance becomes a priority. To mitigate the disruption, public health authorities must act promptly to avert the adverse effects of the disaster, prevent further damage, and restore public service delivery as soon as possible. In this sense, public health surveillance, epidemiology, can identify local problems and establish priorities for decision-making in the health area. In this article, mention is made of one of the most alarming events that occurred in Sillapata, Peru, where a level 4 landslide affected the infrastructure of the population. Considering an established statistical model, it is possible to predict the zoning of higher risks, and thus establish the most appropriate territorial planning and epidemiological surveillance when similar events reach this population or other populations of the Peruvian State.

Keywords: landslide, landslide predictive models, epidemiology surveillance, epidemiological control.

RESUMEN

Casi el 17 % de causas de muerte por amenazas naturales es producto de los deslizamientos de masa. La mayoría de ellas ocurre en los sitios más deprimidos de los países menos desarrollados coexistiendo una combinación letal de factores que apuntan a este tipo de tragedias: el factor natural y el humano. Por otra parte, después de un desastre, las necesidades y prioridades de cuidado de salud pueden cambiar; en ese sentido, el aseguramiento alimenticio de los refugiados, el suministro de agua de potable, la disposición de excretas y desechos sólidos, la necesidad de albergues, la atención de las necesidades de higiene personal, el control de vectores, la atención de las lesiones después de las actividades de limpieza y la conducción de la vigilancia en salud pública se hace prioritarias. Para mitigar el trastorno, las autoridades de salud pública deben actuar con prontitud para evitar los efectos advesos del desastre, prevenir más daños y restaurar la prestación de servicios públicos lo más pronto posible. En ese sentido, la vigilancia en salud pública, la epidimiología, puede identificar los problemas del lugar y establecer prioridades para la toma de decisiones en el área de la salud. En este artículo, se hace mención a uno de los eventos más alarmante ocurrido en Sillapata, Perú, donde un deslizamiento nivel 4 afectó la infraestructura de la población. Tomando en cuenta, un modelo estadístico establecido es posible predecir la zonificación de mayores riesgos, y de esta manera establecer la planificación territorial y de vigilancia epidemiológica mas adecuada cuando eventos similares alcance a esta población o a otras poblaciones del Estado Peruano.

***Palabras clave**: deslizamiento de masa, modelos predictivos de deslaves, vigilancia epidemiología, control epidemiologico.*

[1] Universidad Nacional Autónoma de Tayacaja Daniel Hernández Morillo, Huancavelica, Perú.
[2] Universidad Nacional Agraria de la Selva, Tingo María, Perú.
[3] Consultor, Constructor & Auditor LEOC E.I.R.L, Perú.

Ministerio del Poder Popular para la Salud
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648
Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412
[4] Universidad Nacional del Altiplano, Puno, Perú.
[5] Universidad Interamericana para el Desarrollo, Jalisco, México.
*Autor de Correspondencia: jorge.diazdu@ciplima.org.pe

Introduction

Almost 17% of causes of death from natural hazards are the product of landslides. Most of them occur in the most depressed areas of less developed countries. In this case, there is a lethal combination of factors that point to this type of tragedies: the natural and human factors, which fail to equate the same situation for this type of event in developed countries; becoming a system with great economic and human losses (Kalsnes, B., Lacasse, F., & Nadim, F. (2010). Living with Landslide Risk. Geotechnical Engineering Journal of the SEAGS & AGSSEA, 41(4). Available at: https://www.researchgate.net/publication/267423148 (Accessed July 2022)). In this sense, Peru does not escape this situation, being nestled between the Pacific Ring of Fire and the Andes Mountains ranges that, united by the South Pacific Anticyclone, achieve a perfect, diverse and complex combination of its geography, hydrometeorology, and geology, among other physical characteristics, which impart the richness of biodiversity and landscapes, but which also expose it to various natural phenomena that have affected this country since time immemorial. According to the international disaster data agency EM-DAT (CRED/ UC Louvain (2020). EM-DAT Public, the International Disaster Database. Available at: https://public.emdat.be/mapping (Access July 2022), between 1,900 and 2021, Peru was ranked as the sixth country with the highest number of massive landslides but at the same time, with the highest number of fatalities (9,977 victims) and the highest estimated economic damage: $4,340,902,000 worldwide. According to the Civil Defense Institute (National Institute of Civil Defense-INDECI (2018). Guide and train for a prepared country. Available at: https://www.indeci.gob.pe/preparacion/peligros/ (Access July 2022)), during the first six months of each year, landslides are the second most frequent emergency, and recurring in jungle and mountainous areas.

On the other hand, after a disaster, health care needs and priorities may change after the emergency phase, in this sense, the food security of refugees, the supply of drinking water, the disposal of excreta and waste solids, the need for shelters, attention to personal hygiene needs, vector control, care for injuries after cleanup activities, and conducting public health surveillance are priorities (PAHO, Pan American Health Organization. (1983). Health services organization in the event of disaster. Washington, D.C. Available at: https://www.paho.org/en/health-emergencies (Accessed July 2022). In this sense, after an environmental disaster, such as a landslide, and where the quality of the water has decreased, boiling it is a very good recommendation to ensure its drinkability. The Environmental Protection Agency (EPA) recommends boiling water for at least one minute, but it must increase one more minute of boiling every 1000m above sea level (WHO, World Health Organization. (1993). Guidelines for drinking-water quality: recommendations. Volume 1, 2nd ed., 1-29. Available at: https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151_eng.pdf (Accessed July 2022)), which allows to inactivate bacterial pathogens such as: *V. cholerae*, *Yersinia enterocolitica*, enterotoxigenic *E. coli*, *Salmonella*, *Shigella sonnei*. *Campylobacter jejuni* and protozoa such as *Cryptosporidium parvum*, *Giardia lamblia* and *Entamoeba histolytica* (CDC, Centers for Disease Control and Prevention. (1994). Assessment of inadequately filtered public drinking water--Washington, D.C., December 1993. MMWR. Morbidity and mortality weekly report, 43(36), 661–669. Available at: https://pubmed.ncbi.nlm.nih.gov/8072479/ (Accessed July 2022)). In the case of large populations, the immediate availability of large sources such as rivers and lakes, purification should be done using some antibacterials such as iodine, potassium permanganate or chlorine in order to reduce microbial contamination (PAHO, 1983. *Op. Cit.*; Kozlicic, A., Hadzic, A., & Bevanda, H. (1994). Improvised purification methods for obtaining individual drinking water supply under war and extreme shortage conditions. Prehospital and disaster medicine, 9(2 Suppl 1), S25–S28. https://doi.org/10.1017/s1049023x00041145). However, in some cases, efficient filtration systems are necessary to eliminate chlorine-resistant parasites such as *Cryptosporidium parvum* (CDC, 1994. *Op. Cit.*).

In the case of improper disposal of feces, it can adversely affect public health, and can transmit diseases such as typhoid fever, cholera, bacillary and amoebic dysentery, hepatitis, poliomyelitis, schistosomiasis, various helminthiasis and common gastroenteritis (Noji, E. (2000). Impact of disasters on public health. Pan American Health Organization. Available at: https://iris.paho.org/bitstream/handle/10665.2/754/9275323321.pdf?sequence=1&isAllowed=y (Accessed July 2022); Feachem, R. G., Bradley, D. J., & Garelick. H. (1983). Sanitation and disease: health aspects of excreta and wastewater management. New York: John Wiley & Sons. Available at: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/704041468740420118/sanitation-and-disease-health-aspects-of-excreta-and-wastewater-management (Accessed July 2022)). Likewise, the proper disposal of urine can be an important public health consideration. Emergency projects dispose of human excreta by burying, burning, or composting (manure) (Feachem *et al.*, 1983. *Op. Cit.*); in undeveloped areas, the population will require burial of feces or use in dug latrines (UNICEF, United Nations Children's Fund. (1992). Assisting in emergencies: a resource handbook for UNICEF field staff. New York, 34-365. Available at: https://www.humanitarianlibrary.org/resource/emergency-field-handbook-guide-unicef-staff-0 (Accessed July 2022)). In the case of shelters with large populations, it is important to have an important source of drinking water and excreta management through environmental investigations. On the other hand, the risk of communicable diseases must be minimized due to overcrowding (Noji, 2000. *Op. Cit.*).

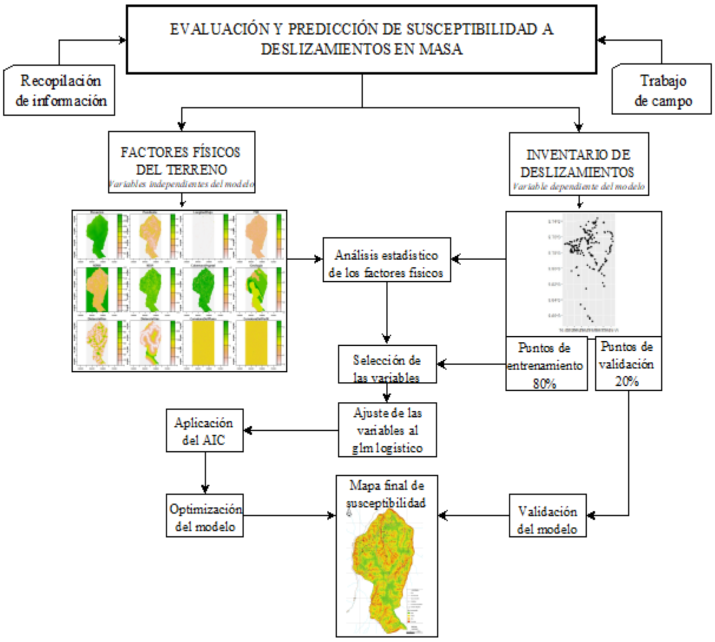
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648
Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412

However, in a disaster situation, such as landslides, the relationships between people and their environment are disrupted as well as the relationships between groups of individuals. To mitigate the disruption, public health authorities must act promptly to avert the adverse effects of the disaster, prevent further damage, and restore public service delivery as soon as possible. In this sense, public health surveillance, epidemiology, can identify local problems and establish priorities for decision-making, evaluating the effectiveness of these activities. Thus, disaster epidemiologists apply various descriptive and analytical techniques to study the natural phenomenon. The epidemiologist quickly defines the nature and extent of the health problems, identifying the population group at risk, thus optimizing the response, the effectiveness of the efforts, and the recommendation to reduce the consequences of the disaster (Noji, E. K. (1992). Disaster epidemiology: challenges for public health action. Journal of public health policy, 13(3), 332–340. Available at: https://pubmed.ncbi.nlm.nih.gov/1401051/ (Accessed July 2022); Glass, R. I., & Noji, E. K. (1992). Epidemiologic surveillance following disasters. In: Halperin W, Baker EL, editors. Public Health Surveillance. New York: Van Nostrand Reinhoid; 1992. p.195-220. Available at: http://cidbimena.desastres.hn/pdf/eng/doc2698/doc2698.htm (Accesses July 2022)).

In this article, mention is made of one of the most alarming events that occurred in the Huanuqueña Sierra, in Peru. In this event, a level 4 landslide occurred in the Sillapata district, affecting the infrastructure of the population. For their part, national authorities, such as INGEMMET (Geological, Mining and Metallurgical Institute-INGEMMET (1996). Geomorphological map of Peru at a scale of 1:100,000. Lima Peru; Geological, Mining and Metallurgical Institute-INGEMMET (2019). Technical Report No. A6915. Geological evaluation of the Sillapata sector - First report. Lima Peru, Available at: https://un-spider.org/geological-mining-and-metallurgical-institute-peru-ingemmet (Accessed July 2022)), recommended the relocation of the affected population. Despite the advances in the disaster risk management system, the systematic evaluation of hazard, vulnerability and risk is still limited, showing the inefficiency of the post-disaster stages. In response, rehabilitation and reconstruction leaves the population in a very precarious state. To date, the databases at the national level are very scarce, which makes it difficult to protect the population and invest in future disaster impacts due to natural phenomena. In this case, massive landslides, caused by internal geodynamic phenomena, are characterized by having a great impact on Andean environments such as the Peruvian highlands, many of them violent and recurring, so it is important to have enough studies to make a preliminary identification of the places with the greatest risks. Considering an objective point of view based on statistics, the zoning of higher risks is possible, useful to know the probable effects associated with landslides, and thus establish territorial planning. In this sense, the Sillapata district will be studied to determine the factors that could intervene to a greater degree in modeling its susceptibility to landslides. The results will lead to zoning metrics that are closer to reality. In addition to the above, the zoning would also allow real-time surveillance of epidemiological events that could affect the area in studies or other areas with similar problems.

Methodological development

The methodology followed is summarized in Figure 1.



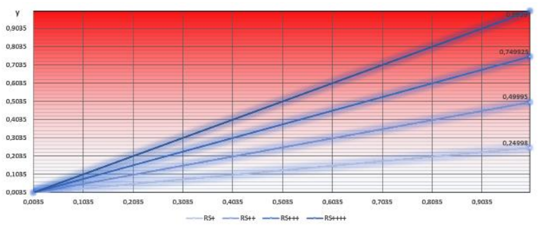
Estimation criteria:

Syndrome	Disease	Likely source of contamination	Risk potencial
acute diarrheal	Unspecified diarrhoea	water, food, overcrowding, high	+++
	salmonella/shigellosis	temperatures, poor sanitation	++ +/2
	cholera		++
	food poisoning		++ +/2
Respiratory	acute respiratory infections	activity, exposure to cold, shelter	+++
	tuberculosis	practice, low vaccination coverage	+
	whooping cough	interruption of control programs	++
	diphtheria		++
febrile	Malaria	Presence of mosquito breeding	++ +/2
	Dengue	ground, temperature increase,	++
	Plague	accumulation of useless, inadequate	+
	typhus	water storage, unhygienic conditions, inadequate control of rodents.	+
Acute Jaundice febrile	yellow fever	Increase in vectors, endemic areas,	++
	Hepatitis A	low immunization coverage,	++ +/2
	leptospirosis	migration of people, water/food contamination, no rodent center, inadequate sanitation	++

Ecuation:

$$EP = y \cdot CF$$

Probability plot:



Estimation of "y" Risk of massive movement probability plot

"EP" estimate of potential epidemic risk

Figure 1. Methodological development

The study area corresponded to the District of Sillapata, in the Provinces of Dos de Mayo/Huánuco, Peru, located in the Andean zone of the country and frequently hit by external geodynamic phenomena. The district limits, according to the INEI, are the following (Table 1).

Table 1. Geographical location according to the UTM System Zone 18S – Datum WGS84

Spot	East	South	Altitude
North limit point	306600	8923808	3068
West limit point	301960	8916457	3176
South limit point	306549	8908671	3572
East limit point	309771	8921176	3416
Lowest point	305910	8923109	3076
Highest point	307031	8916903	4073
Capital	305478	8920898	3423

Information gathering

The data collected was taken from different sources. The data on past landslides were provided by the Civil Defense Secretariat of the Sillapata District Municipality, the Dos de Mayo Provincial Highway Institute and the SINPAD of the National Civil Defense and Huánuco Regional Government. The geographic and satellite information of the scene was taken from NASA's Alos Palsar digital elevation model (Available at: https://search.asf.alaska.edu/#/ (Accessed July 2022)) completed with two quads of the satellite mosaic PlanetScope during the monthly period of February 2021 (QGIS Planet_Explorer. (2021). Google-Satellite satellite mosaic in high definition; WMS Geology Service Map 1:100'000–INGEMMET and NATIONAL PROVIAS, Available at: https://mappinggis.com/2021/09/how-to-download-google-satellite-images-with-qgis/ (Accessed July 2022)).

Points inventory

Due to the scant geographic information collected, field work and satellite photointerpretation were carried out to identify landslide and non-landing points. Field surveys were made from the three highest peaks to the lowest areas, taking the coordinates of previous and active landslide events. In the case of inaccessible sites, the data was taken at the foot of the landslide. Some of the "no slip" points were taken in the field and others were taken with the support of the Google-Satellite high-definition mosaic. All these points, 66 landslide points and 110 non-slide points, were registered in landslide inventory files under *csv format.

Mapping of physical factors of the terrain

The physical factors used in the statistical model were substitutes for physical parameters described in the literature that influence the materialization of massive landslide phenomena whose large-scale estimation is not feasible (Pourghasemi, H; Sadhasivam, N; Amiri, M; Eskandari, S & Santosh, M. (2021). Landslide susceptibility assessment and mapping using state-of-the-art machine learning techniques. Springer, Natural Hazards. Available at: https://doi.org/10.1007/s11069-021-04732-7 (Accessed July 2022) (Table 2).

Table 2. Source and spatial resolution of the variables used

Variable	Source	Resolution / scale
Elevation	MDE ALOS PALSAR	12.5m x 12.5m
Slope (degrees)	MDE ALOS PALSAR	12.5m x 12.5m
flow length	MDE ALOS PALSAR	12.5m x 12.5m
TWI	MDE ALOS PALSAR	12.5m x 12.5m
NDWI	PlanetScope Mosaic	4.68m x 4.68m
NDVI	PlanetScope Mosaic	4.68m x 4.68m
plant cover	PlanetScope Mosaic	4.68m x 4.68m
geology	Geological map - INGEMMET	1: 100,000
Distance to water flows	MDE ALOS PALSAR	12.5m x 12.5m
Distance to tracks	Road network - PROVIAS	Does not have
plane curvature	MDE ALOS PALSAR	12.5m x 12.5m
profile curvature	MDE ALOS PALSAR	12.5m x 12.5m

The distribution and frequency of each factor or variable is determined in Figure 2.

Ministerio del Poder Popular para la Salud
Boletín
de Malariología
y Salud Ambiental
ISSN: 1690 - 4648

Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412

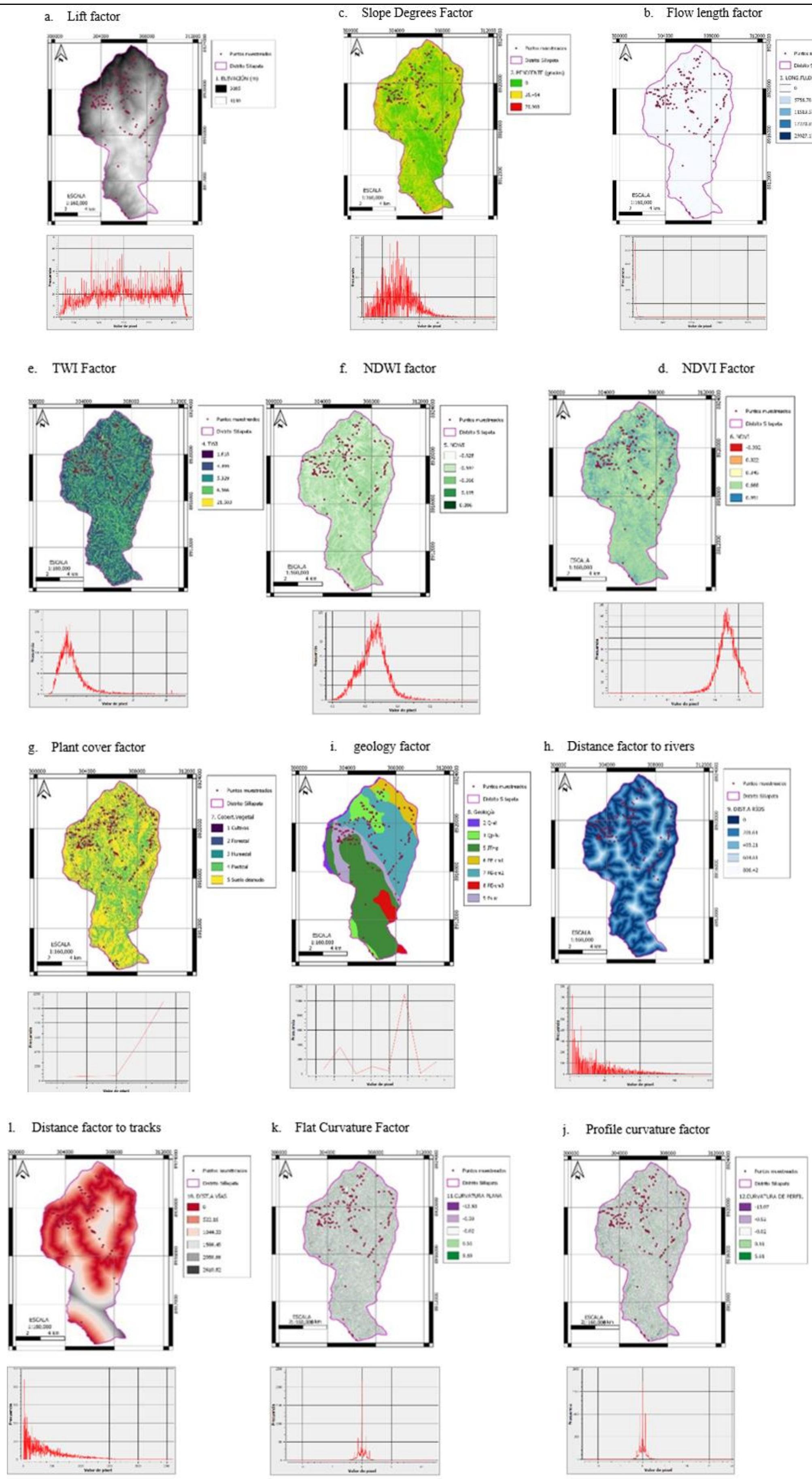


Figure 2. Spatial distribution and frequency of physical factors used in the present work

The geographic information collected was processed and georeferenced in UTM 18S projection and WGS84 datum, generating the 12 most important physical factors (Table 2). These physical factors were mapped in raster format, aligned in 4.68 m pixels according to the Sillapata district polygon. The results obtained are shown in individual maps (Figure 3). The variables: Elevation, Slope, Flow Length, TWI, Distance to water flows, Curvatures, NDWI and NDVI were generated by remote sensing of digital elevation models and satellite images with the support of GIS tools.

The Vegetation Cover variable, allowed to determine the type of vegetation located that grows on the ground based on the economic activities that take place in that place. Points were collected and interpreted using high-resolution Google-Satellite images, based on the Gdal and Scikit Learn libraries (Duchesnay, E., Gramfort, A., Grisel, O., Michel, V., Pedregosa, F., Thirion, B., & Varoquaux, G. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830. Available at: https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf (Accessed July 2022)). The vegetation cover area was grouped into five main categories: Crops, Forestry, Wetlands, Grasslands, and Bare Soil.

The Geology of the model was extracted from the National Geological Maps of INGEMMET (Geological, Mining and Metallurgical Institute-INGEMMET (1996). Geomorphological map of Peru at a scale of 1:100,000. Lima Peru. Geological, Mining and Metallurgical Institute-INGEMMET (2019). Technical Report No. A6915. Geological evaluation of the Sillapata sector-First report. Lima Peru, Available at: https://un-spider.org/geological-mining-and-metallurgical-institute-peru-ingemmet (Access July 2022)), identifying five geological units (Table 3). Finally, the variable Distance to highways was generated from the database generated by PROVIAS, updated according to high-resolution images from Google's satellite. Using GIS it was possible to determine the layer of Euclidean distances of roads in the entire district. All these variables were trained and validated according to the proposed model. Table 4 shows a summary of the variables considered in the new general database: the independent ones (physical variables), the dependent ones (slip susceptibility) and the location of each point.

Exploration of physical factors

In the optimization of the model, the variables were reviewed, refined and selected through a statistical analysis of the behavior of the 176 points and their data. The 10 quantitative independent variables were analyzed by correlation, while the 2 qualitative variables were analyzed using a mosaic graph; the process made it possible to discard duplicate variables due to correlation.

Table 3. Geological formations in the study area

Name	Training	lithology	Age
PE-cm	Cashew Complex	Schists and gneiss. Outcrop of brown to gray slates, schists, and phyllites	Neo-proterozoic
JTr-p	Pucara Group	Blue-grey limestones in medium to coarse banks with chert nodules	Triassic - Jurassic
Ps-m	Mitu Group	Purple andesite and lava flows, pyroclastic andesite, pebble conglomerate of andesite and red sandstone in varying proportions	Early Paleogene
Q-al	Alluvial Deposits	Accumulation of gravel, sand, silt and clays, with sub-angular to angular clasts of different composition	
Qp-lu	Formation The Union	Coarse polymictic conglomerates and semi-consolidated sandstones in a sandy matrix	Pleisto I dine

Source: INGEMMET, 1996.

Application of the generalized linear model – glm

The generalized linear model-glm was proposed using the independent variables (continuous and categorical) and the dependent variable (binomial variable), in this way the glm logistic function and the training database were used, resulting in the AIC value of the adjustment of said model (Table 4) (Aristizábal, E., & Ospina-Gutiérrez, J. (2021). Application of artificial intelligence and machine learning techniques for the evaluation of susceptibility to mass movements. Mexican Journal of Geological Sciences, 38(1), 43-54. http://dx.doi.org/10.22201/cgeo.20072902e.2021.1.1605).

Adjustment and simplification of the model

The model was adjusted according to the Akaike Information Criterion in order to simplify it and select the adjustment function to the field data. This model worked with the most significant variables.

Susceptibility map prediction and validation

With the final model, the entire district territory was predicted and classified by the Natural break method (Pourghasemi, H; Sadhasivam, N: Amiri, M; Eskandari, S. & Santosh, M. (2021). Landslide susceptibility Assessment and mapping using state-of-the-art machine learning techniques Springer, Natural Hazards. https://doi.org/10.1007/s11069-021-04732-7; Chen, W., Pourghasemi, H. R, Zhang, S., & Wang, J. (2019). A comparative study of functional Data Analysis and Generalized Linear Model Data–Mining Techniques for Landslide Spatial Modeling. Elsevier, 467-484. https://doi.org/10.1016/B978-0-12-815226-3.00021-1) in 4 levels of susceptibility

Ministerio del Poder Popular para la Salud
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648
Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412

to massive landslides: low, medium, high and very high. The model was validated with 20% of the points from the general database using the ROC curve and the AUC-COR to determine its discriminative capacity (Table 5).

Table 4. Characteristics of the variables (dependent and independent) used

floating rate		Variable	Unit of measurement	Study area	Guy	Occurrence range
SAW	1-	Elevation	meters	geomorphology	continuous quantitative	3085msnm to 4140msnm
SAW	2-	slope degrees	sexagesimal degrees	geomorphology	continuous quantitative	0° to 70.91°
SAW	3-	flow length	meters	hydrogeology	continuous quantitative	0m to 23027.13m
SAW	4-	TWI	Percentage	geomorphology	continuous cuantitative	1,618% to 21,503%
SAW	5-	NDWI	dimensionless	land use	continuous quantitative	-0.828 to -0.096
SAW	6-	NDVI	dimensionless	land use	continuous quantitative	-0.302 to 0.991
SAW	7-	plant cover	dimensionless	land use	Qualitative categorical	1(Crops) 2(Forest) 3(Wetland) 4(Grassland) 5(Bare soil)
SAW	8-	geology	dimensionless	geology	Qualitative categorical	2(Q-al) 3(Qp-lu) 5(JTr-p) 6(PE-cm1) 7(PE-cm2) 9(Ps-m)
SAW	9-	Water flow distance	meters	Hydrography	continuous quantitative	0m to 806.42m
SAW	10-	Track distance	meters	land use	continuous quantitative	0m to 2610.82m
SAW	11-	plane curvature	dimensionless	geomorphology	continuous quantitative	-12,504 to 9,695
SAW	12-	profile curvature	dimensionless	geomorphology	continuous quantitative	-13,073 to 5,805
YOU	landslide susceptibility		dimensionless	Risk management	Qualitative binomial	0 to 1,000

VI: Independent variable. DV: Dependent variable.

The correlation between the pairs Length of flow-TWI and Curvature of the plane-TWI did not exceed the value |0.900| and was retained for further analysis. The NDWI–NDVI pair presented a correlation close to 1,000. NDVI has a history as an influential variable (Chen *et al.*, 2019. *Op. Cit.*; Mohamed, A. & Pourghasemi, H. (2020). Landslide susceptibility mapping using machine learning algorithms and comparison of their performance at Abha Basin, Asir Region, Saudi Arabia. Geoscience Frontiers Magazine. Available at: https://doi.org/10.1016/j.gsf.2020.05.010), while NDWI an adaptation of the original equation (Gao, B. (1996). NDWI–A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space. Elsevier, 58, 257-266. Available at: https://doi.org/10.1016/S0034-4257(96)00067-3 (Accessed July 2022); McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing, 17(7), 1425-1432. https://doi.org/10.1080/01431169608948714) may be subject to variations in representation, so it was discarded.

Table 5. Correlation matrix of quantitative variables

	Elevation	pending	length Flow	TWI	NDWI	NDVI	Dist. Rivers	Distance Tracks	curve flat	curve profile
Elevation	1,000	-0.126	-0.035	-0.015	0.108	-0.145	0.405	0.207	0.028	-0.053
pending	-0.126	1,000	-0.050	-0.225	0.118	-0.097	-0.111	-0.047	-0.006	-0.001
length Flow	-0.035	-0.050	1,000	0.855	0.072	-0.073	-0.178	0.044	-0.368	0.011
TWI	-0.015	-0.225	0.855	1,000	0.029	-0.025	-0.165	0.048	-0.502	0.159
NDWI	0.108	0.118	0.072	0.029	1,000	-0.976	0.095	-0.061	0.091	-0.070
NDVI	-0.145	-0.097	-0.073	-0.025	-0.976	1,000	-0.119	0.095	-0.045	0.079
Dist. Rivers	0.405	-0.111	-0.178	-0.165	0.095	-0.119	1,000	-0.053	0.149	-0.186
Distance Tracks	0.207	-0.047	0.044	0.048	-0.061	0.095	-0.053	1,000	0.110	-0.051
curve flat	0.028	-0.006	-0.368	-0.502	0.091	-0.045	0.149	0.110	1,000	-0.476
curve profile	-0.053	-0.001	0.011	0.159	-0.070	0.079	-0.186	-0.051	-0.476	1,000

An exploratory analysis of the qualitative variables according to figure 3, shows that the vegetation cover factor indicates that cover 5(bare soils) is the one with the highest concentration and to a lesser degree in cover 4 (pastures). In all categories there is a greater number of stable points, so that in category 1(crops) and 3(wetlands) there are no sliding points, while in categories 2(forestry), 4 and 5 present points slippage increasingly distributed. As for geology, with six categories, class 7(PE-cm2) is the one that concentrates the largest number of points and the greatest equity between landslide and non-landscape points.

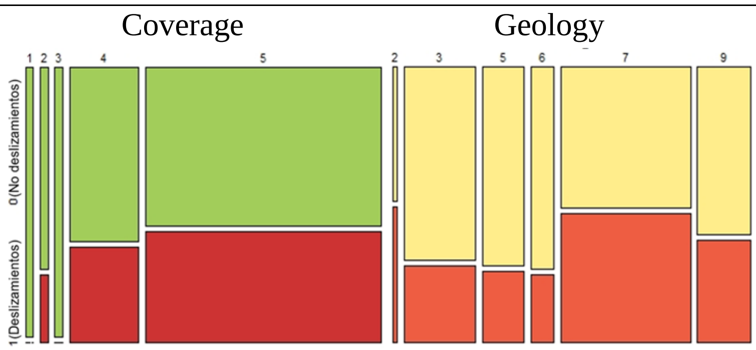


Figure 3. Mosaic chart of qualitative type variables

Table 6. Most influential physical factors to be used in the GLM.

Selected physical factor	Designation on the model*
Elevation	Elevation
Earring	Earring
flow length	FlowLength
TWI	TWI
NDVI	NDVI
plant cover	Plant Coverage
geology	Geology
Distance to water flows	DistanceRios
Distance to tracks	DistanceWays
plane curvature	CurvatureFlat
profile curvature	CurvatureProfile

Development of the landslide susceptibility model

The most appropriate generalized linear model to susceptibility according to the types of research variables is shown in table 7.

Table 7. Selection of the generalized linear model.

Response variable	Predictor variables	Main recommended method*	link function
Nominal / Binary	Categorical and continuous	Logistics	$\log\left(\frac{\mu}{n-\mu}\right)$

Selection of variables and model

Initially, the generalized linear model was processed including all the variables, but later it was adjusted to deviance analysis and the Akaike Information criterion. The variables selected for the final generalized linear model are shown below (Table 8).

Tabla 8. Summary of the variable selection process for the final model

Variable	AIC	final model	Coefficient in the model
All	194.31	Nope	-
plant cover	188.12	Nope	-
geology	185.47	Nope	-
plane curvature	183.49	Nope	-
Distance to tracks	181.63	Nope	-
flow length	179.82	Nope	-
Elevation	178.75	Nope	-
profile curvature	179.34	Yes	0.4325
Distance to water flows	180.57	Yes	-0.0029
Earring	181.62	Yes	0.0571
Normalized Differential Vegetation Index - NDVI	182.13	Yes	-4.8733
Topographic Moisture Index - TWI	183.61	Yes	0.2951
Independent	-	Yes	0.4199

The AIC value is a relative metric for finding the model that best fits the data via the minimum AIC. When discarding some variables from the models, the AIC values reached the lowest values, but in the models that exclude the variables Profile curvature, Distance to water flows, Slope, NDVI and TWI and the AIC values showed an increase and they were not excluded. Thus, the final model was made up of the five variables mentioned, as presented below:

Ministerio del Poder Popular para la Salud
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648

$$y = \frac{1}{1+e^{-(0.420 + 0.432X_1 - 0.003X_2 + 0.057X_3 - 4.873X_4 + 0.295X_5)}}$$

Where:

Y is the mass landslide susceptibility value, X1 are the profile curvature values, X2 are the distance values to water flows, X3 are the slope values, X4 are the NDVI values, and X5 are the values of TWI. The final variables and the model are a unique result for the study area, and should not be generalized to other places (Aristizábal & Ospina-Gutiérrez, (2021). *Op. Cit.*). Figure 4 shows the map of the entire district territory, whose prediction was based on the final model. The raster of the final model of susceptibility to massive landslides is made up of 3,325,477 pixels in the study area, whose values fluctuate between 0.0035 and 0.9999; using GIS. These values were categorized with the Natural breaks method (Figure 4), obtaining four terrain categories: low levels between 0.0035 to 0.2184, medium levels between 0.2184 to 0.3864, high levels between 0.3864-0.5936 and very high levels between 0.5936-0.9999. With the model of susceptibility to massive landslides applied to the Sillapata territory, it can be observed (Figure 5) that the largest proportion of said territory has a medium level of susceptibility (33.92%) and the smallest proportion has a very high level of susceptibility high (9.27%).

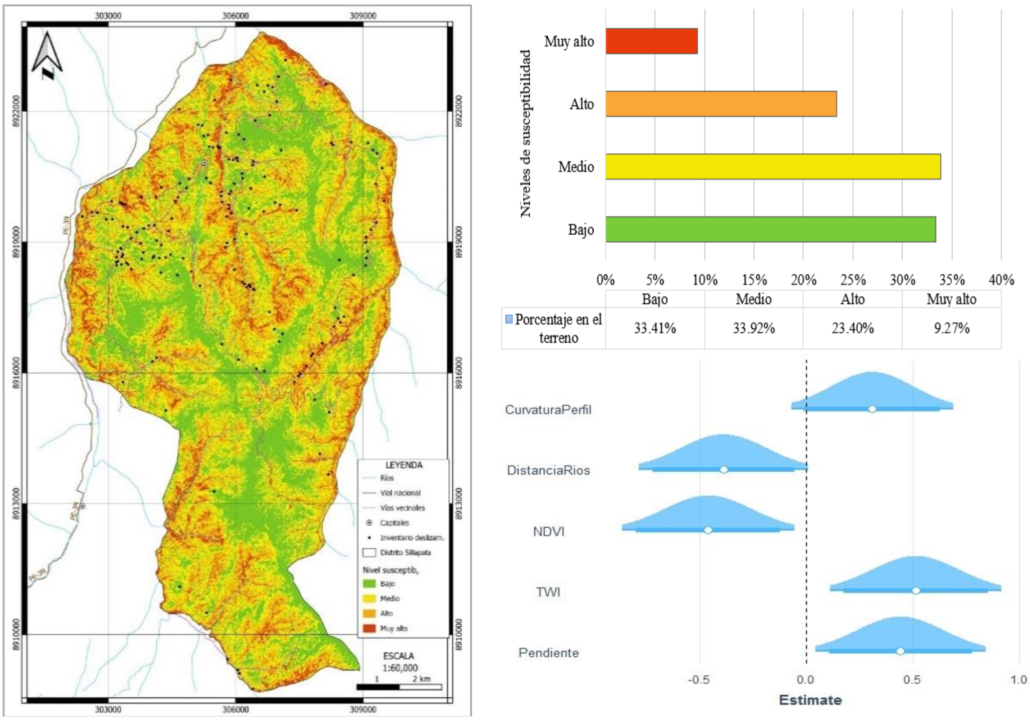


Figure 4. Mass landslide susceptibility map and percentage of the Sillapata district, using the final logistic generalized linear model

The analysis of the ROC curve of the model (figure 5) determined that the best threshold or Youden index is the point with a value of 0.422 (with a confidence interval of 84% to 80%), this means that the fraction below this value is The highest fraction of zones prone to stability, which, in turn, will be correctly identified with the model, and vice versa, above said threshold is the highest fraction of zones prone to mass landslides (with the presence of landslide susceptibility) determined as such by the model.

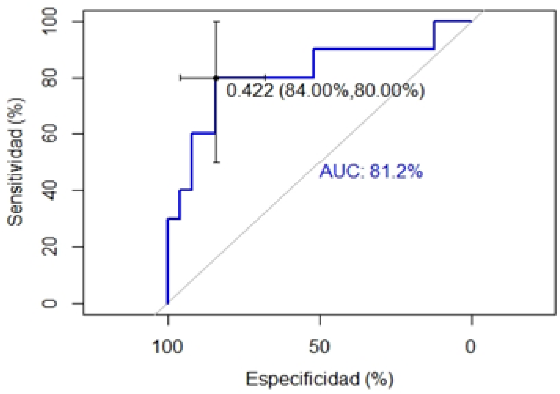


Figure 5. ROC curve for validation of the susceptibility model to landslides

The area under the ROC curve (AUC-ROC) of the final model determined a capacity of 81.2% to discriminate areas susceptible and not susceptible to landslides. Pourghasemi *et al.*, (2021) (*Op. Cit.*) obtained an AUC value of 90.8%, while Chen *et al.*, (2019) (*Op. Cit.*) obtained values with 71.8% in their model. Both authors used fewer training points (70%), which shows that the proportion of points per training and validation does not influence the generation of a more valid model. This result can be improved by refining certain processes such as determining the relationship between the physical factors and the model. Once the susceptibility model to massive landslides has been determined, it is important to know the type of relationship it has with the independent variables that were part of the final model, which will be useful as an approach for prospective risk management, in the process of decreasing susceptibility to landslides.

Figure 4 (Lower right grid) shows that the relationship between the response variable. NDVI and Distance to water flows showed a negative association with landslide susceptibility, which means that at lower NDVI values or at shorter distances from water bodies, there is a higher spatial probability of landslides occurring. In the case of the Profile Curvature, Slope and TWI variables, the association was positive, that is, for areas with higher values of these variables, the spatial probability of landslide occurrence is also higher. NDVI values in soil range from -0.302 to 0.991, negative values represent man-made structures (dwelling areas); values close to 0, bare soils, water surfaces and rocks; while values close to 1, areas with vegetation. The negative association exposed in this work proves that the surfaces with more vegetation correspond to areas with less susceptibility to massive landslides and vice versa. Suárez (Suarez, J. (2009). Landslides: Geotechnical Analysis. Edit. Industrial University of Santander UIS, 01, 582 pp. Available at: https://www.erosion.com.co/deslizamientos-tomo-i-analisis -geotechnical/ (Access July 2022)) mentions that the vegetation slows down the rate of infiltration and reinforces the stability of the soil with its roots and the organic matter that is generated; however, its influence is subject to variations due to other factors such as precipitation, slope, and depth of the slide. The highest values of moisture accumulation of the TWI are located in areas with drainage depressions and the lowest values of TWI coincide with the mountainous formations (greatest slope), being the highest part where the least amount of water accumulates. According to the positive association, those areas with greater moisture accumulation (depressions and riverbanks) showed greater susceptibility to landslides, contradicting the inverse association exposed by Meinhardt *et al.* (Meinhardt, M., Fink, M., & Tunschel, H. (2014). Landslide susceptibility analysis in central Vietnam base don an incomplete landslide inventory: Comparison of a new method to calculate weighting factors by means of bivariate statistics. Geomorphology Magazine. https://doi.org/10.1016/j.geomorph.2014.12.042), who also explains that the influence of climatic factors can produce a sharp increase in pore water pressure and destabilize slopes (lower TWI), which encourages the need to study the temporal component (variation of climatic parameters and of the TWI variable itself) to better understand the danger of massive landslides. In general, seasonality should be considered at the time of data collection, to avoid variations due to the effect of those variables whose range is subject to temporary variations. The final model discarded the categorical variables: geology and land cover, while Chen *et al*., (2019) (*Op. Cit.*) determined the high influence of variables such as land use and lithology, with similar sources of information to the variables of the present work, unlike the scale of acquisition of the geology used in the present and the lithology used in the mentioned investigations (1:100,000 and 1:1,000,000 respectively); Since the antecedent models depend on this type of categorical variables, the spatial resolution of the initial variables and the scale of analysis must be carefully considered (Mohamed & Pourghasemi, (2020) *Op. Cit.*) and they can be a source of error by omitting variables with potential bias.

The data obtained were compared to those reported by CENEPRED and published in SIGRID (Available at: https://sigrid.cenepred.gob.pe/sigridv3/map (Access July 2022)). In this map, areas without data were observed that covered 9% of the points with massive landslides confirmed by the inventory, while 83% of the points with landslides corresponded to areas of high and very high susceptibility. Regarding the visual comparison of both maps, the CENEPRED map differentiated two regions of susceptibilities, while the map of the present investigation, the susceptibility was more varied due to the scale of work and the change in the measurements of the variables. Finally, the diversity of relationships of the 'independent variable–model' type in models of susceptibility to massive landslides of different mountain scenarios in the world (Londono-Linares, J.P. (2017). Landslide susceptibility calculation by discriminant analysis. Application on a regional scale. Dyna Magazine, 84 (201), 278–289. https://doi.org/10.15446/dyna.v84n201.61385; Chen *et al.*, (2019). *Op. Cit.*; Mohamed & Pourghasemi, (2020). *Op. Cit.*; Aristizábal & Ospina-Gutiérrez, (2021) *Op. Cit.*; Pourghasemi et al., (2021) *Op. Cit.*) show that these are not always valid for planning and evaluation purposes in very local spaces, if not they are worked taking into account the uniformity of scales between variables and the territory (Hong, S., Jung, H., & Lee, S. (2017). A support vector machine for landslide susceptibility mapping in Gangwon Province, Korea. Sustainability, 9(48). https://doi.org10.3390/su9010048; Mohamed & Pourghasemi, (2020). *Op. Cit.*), so it will be necessary to choose independent variables based on field observations. According to the statistical analyses, five physical factors intervened in the final model of susceptibility to massive landslides in the Sillapata district, which were later evaluated.

The physical factors that best intervened in the final model of susceptibility to landslides in the study area were: degrees of slope, soil moisture index TWI, normalized differential vegetation index NDVI, distance to rivers, and curvature of the terrain profile.

Ministerio del Poder Popular para la Salud
Boletín de Malariología y Salud Ambiental
ISSN: 1690 - 4648
Noviembre-Diciembre 2022, Vol. LXII (6), 1401-1412

Potential epidemic risk estimation criteria

The geographical location of Peru locates it under an epidemiological profile under permanent risk of occurrences that of course can affect collective health at any moment in time and that is exceeded by the resolution capacity of basic health services, increasing the probability of getting sick and therefore die later. One of the causes that can increase the potential epidemic risk are natural disasters, including landslides or mass displacement. Under these circumstances, living conditions gets worse in a very short period of time, since the basic services of drinking and sewage water, food production and distribution, transportation, electricity, and sanitary services are interrupted, which in the best of cases can last a few months but in other cases, it can take years. In that same period, the population density increases, and they are located in shelters or unorganized housing system that increases the aforementioned basic necessities of life. This deterioration, more than acute, is a breeding ground for the rebirth of multiple diseases that overload the response capacity of health services.

By that time, the implementation of post-disaster Epidemiological Surveillance must present a surveillance plan that allows controlling and reducing the risk of illnesses or deaths, proposing an epidemiological pattern that controls disease outbreaks. The guidelines allow: evaluating the potential epidemic risk, implementing the surveillance system after health emergencies and implementing the health situation room in the event of a health emergency. Several factors are directly related to the potential epidemic risk: change in pre-existing morbidity, ecological changes resulting from the disaster, displacement of populations (migrations), changes in population density, disruption of public services, interruption of basic services. of public health. For the Evaluation of the Potential Epidemic Risk, the following procedures must be followed: Evaluation of the previous epidemic activity in the area: the evaluation must review the trends of the diseases reported during the disaster, and they must be compared with previous years, determining the morbidity and mortality in the last two weeks, Evaluation of the endemic level under surveillance: with previous epidemiological information, it is possible to monitor the endemic channels to determine the history of possible outbreaks, stratify them through epidemic maps, and then, with this information, determine what type of diseases could be considered as tracers of post-disaster epidemiology and Assessment of post-disaster living conditions: it is important to know how basic services are: drinking and sewage water, environmental sanitation, electricity. From these reports, a preliminary damage assessment and needs analysis can be established. With all the information obtained: epidemiological information and morbidity/mortality, an evaluation of potential epidemic risks will be possible (Manual for the implementation of epidemiological surveillance in Disasters, MINSA (2004, Available at: https://www.gob.pe/institucion/minsa/informes-publicaciones/353483-manual-for-the-implementation-of-epidemiological-surveillance-in-disasters (Accessed July 2022).

The table 9 shows the post-disaster communicable diseases with epidemic potential.

Table 9. Epidemic potential of post-disaster communicable diseases

Syndrome	Disease	Likely source of contamination	Risk potential
	unspecified diarrhea	Water, food, overcrowding, high temperatures, poor sanitation	++++
acute diarrheal	Salmonella/Shigellosis		+++
	Anger		++
	food poisoning		+++
	Acute respiratory infections	Overcrowding, Exposure to cold, shelter	++++
Respiratory	Tuberculosis	Overcrowding, low vaccination coverage	+
	Whooping cough	Interruption of control programs	++
	Diphtheria		++
Febrile	Malaria	Presence of mosquito breeding ground, increase in temperature,	+++
	Dengue	accumulation of useless, inadequate water storage, unhygienic	++
	Plague	conditions, inadequate control of rodents	+
	Typhus		+
Acute jaundiced febrile	Yellow fever	Increase in vectors, endemic areas, low immunization coverage,	++
	Hepatitis A	migration of people, water/food contamination, no rodent control,	+++
	Leptospirosis	inadequate sanitation	++

Source: MINSA, (2004). *Op. Cit.*

Taking into account the data collected in the mathematical model used (equation 1) and the established epidemiological data determined by MINSA (MINSA. (2004). Manual for the implementation of epidemiological surveillance in Disasters. Available at: https://www.gob.pe/institucion/minsa/informes-publicaciones/353483-manual-para-la-implementacion-de-la-vigilancia-epidemiologica-en-desastres (Access July 2022)) (Table 10) it is possible to establish the following formula corresponding to the estimation of potential epidemic risk (EP) (equation 2):

$$EP= y \cdot CF \quad (2)$$

Where:

EP= disease category (diarrhea, cholera, food poisoning, etc., (Table 10))

y = is the susceptibility value to mass landslide determined according to equation 1

CF= is a proportionality factor that is determined according to the probability of contracting a disease associated

With the contaminant present during the landslide or natural disaster. This value can be classified as low, medium, high and very high probability (Table 10); which for this mathematical model, was expressed quantitatively in 0.249 = +, 0.499 = ++, 0.749 = +++ and 0.999 = ++++.

Equation 2 is closely related to Figure 6, Estimation of potential epidemic risk. With this equation it is possible to establish the epidemic risk of a disease (EP value), knowing its risk potential (CF) and the mass landslide susceptibility value (y). The higher the landslide value, and the higher the risk value, there is a greater probability that a disease can occur in the natural disaster. Diseases such as tuberculosis, plague or typhus are less likely to represent an imminent danger when a landslide occurs, since it has a lower risk value (CF=0.249); while those diseases such as: unspecified diarrhea and acute respiratory infections, and to a lesser degree Salmonella, malaria and hepatitis (CF= 0.999 and 0.749) must be attended to and controlled immediately in order to prevent their spread, since they are one of the first diseases to appear after the event occurred. In this sense, the epidemiological actions that the government must take for its control become imperative. In order to avoid massive infections, the rescue and epidemiological teams must provide the population with a drinking water system, initiate a massive, extensive and immediate vaccination program, provide non-perishable and highly nutritious food, and provide service for the discharge of sewage, located away from rivers and lakes that may be in contact with established shelters.

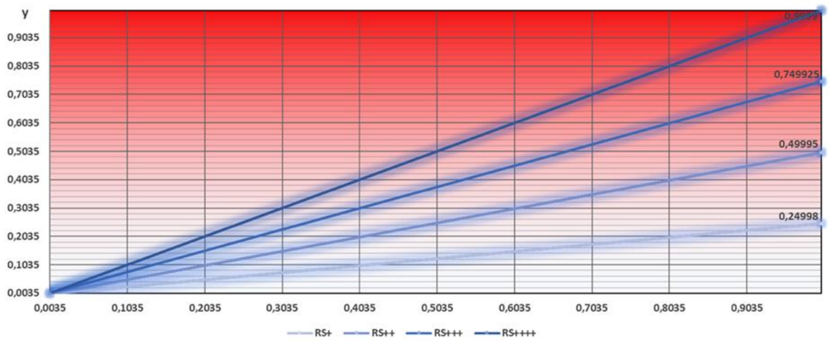


Figure 6. Estimation of potential epidemic risk

Final considerations

The most appropriate generalized linear model, according to the types of variables, was the logistic link function, achieving a satisfactory discrimination capacity with AUC-ROC of 81.2%, which represented 9.27% with very high susceptibility, with a high susceptibility 23.4%, with medium susceptibility 33.92% and with low susceptibility 33.41% of the territory studied.

The physical factors NDVI and distance to rivers presented a negative association with the susceptibility model, since the lower the values, the greater the susceptibility to massive landslides. On the other hand, the factors curvature of the profile, degrees of slope and TWI presented a positive association, since the higher the values, the greater the susceptibility to landslides.

On the other hand, after a disaster, such as landslides, health care needs and priorities may change after the emergency phase, in this sense, the food security of refugees, the supply of drinking water, the disposal of excreta and solid waste, the need for shelters, care for personal hygiene needs, vector control, care for injuries after clean-up activities, and conducting public health surveillance priority. From these reports, a preliminary damage assessment and needs analysis can be established, and epidemiological and morbidity/mortality information can be provided for the evaluation of potential epidemic risks. In this sense, it was possible to formulate an equation with high potential that manages to correlate the magnitude of the landslide with the possible occurrences of permissible diseases. In this way, in figure 6, the potential epidemic risk of diseases with less danger (CF=0.249) of those more aggressive diseases (CF= 0.999 and 0.749) can be estimated. The latter, if not attended to in time, are easily susceptible to being an epidemic regardless of the magnitude of the landslide that occurred. Therefore, immediate action by the authorities is imperative in order to control possible massive contagion, which makes actions such as: drinking water service, starting massive, extensive and immediate vaccination programs, providing non-perishable and highly nutritious, and provide service for wastewater discharges, locating them far from rivers and lakes in order to ensure health in established shelters.

Conflict of interests

No conflict of interest is reported.

Acknowledgment

To the collaborators and to our education and research centers.