

**UNIVERSIDADE DE CAXIAS DO SUL
ÁREA DE CONHECIMENTO DE CIÊNCIAS
EXATAS E ENGENHARIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE
PROCESSOS E TECNOLOGIAS**

**IDENTIFICAÇÃO DE ÁREAS COM SUSCETIBILIDADE DE
OCORRÊNCIA DE *HUAICOS* NO DEPARTAMENTO DE LIMA-
PERU, COM A UTILIZAÇÃO DE *MACHINE LEARNING***

Geise Macedo dos Santos

Orientação: Prof. Dr. Matheus Poletto
Coorientação: Prof.^a Dr.^a Gisele Cemin
Colaboração: Prof.^a Dr.^a Vania Elisabete Schneider

Caxias do Sul, 2026

Geise Macedo Dos Santos

**IDENTIFICAÇÃO DE ÁREAS COM SUSCETIBILIDADE DE
OCORRÊNCIA DE *HUAICOS* NO DEPARTAMENTO DE LIMA-
PERU COM A UTILIZAÇÃO DE *MACHINE LEARNING***

Tese apresentada no Programa de Pós-graduação em Engenharia de Processos e Tecnologias da Universidade de Caxias do Sul, visando a obtenção do grau de Doutora em Engenharia de Processos e Tecnologias, orientado por Prof. Dr. Matheus Poletto e coorientado por Prof.^a Dr.^a Gisele Cemin, com a colaboração da Prof.^a Dr.^a Vania Elisabete Schneider.

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
Sistema de Bibliotecas UCS - Processamento Técnico

S237i Santos, Geise Macedo dos

Identificação de áreas com suscetibilidade de ocorrência de *huaicos* no departamento de Lima-Peru, com a utilização de *machine learning* [recurso eletrônico] / Geise Macedo dos Santos. – 2026.

Dados eletrônicos.

Tese (Doutorado) - Universidade de Caxias do Sul, Programa de Pós-Graduação em Engenharia de Processos e Tecnologias, 2026.

Orientação: Matheus Poletto.

Coorientação: Gisele Cemin.

Modo de acesso: World Wide Web

Disponível em: <https://repositorio.ucs.br>

1. Deslizamentos de terra - Peru. 2. Geomorfologia - Peru. 3. Catástrofes naturais - Peru. 4. Prevenção de desastres. 5. Sistemas de informação geográfica. 6. Inteligência artificial. 7. Aprendizado do computador. I. Poletto, Matheus, orient. II. Cemin, Gisele, coorient. III. Título.

CDU 2. ed.: 551.3(85):004.8

Catalogação na fonte elaborada pela(o) bibliotecária(o)
Ana Guimarães Pereira - CRB 10/1460

Caxias do Sul, 2026
Geise Macedo Dos Santos

IDENTIFICAÇÃO DE ÁREAS COM SUSCETIBILIDADE DE OCORRÊNCIA DE
HUAICOS NO DEPARTAMENTO DE LIMA-PERU COM A UTILIZAÇÃO DE
MACHINE LEARNING

Tese apresentada no Programa de Pós-graduação em Engenharia de Processos e Tecnologias da Universidade de Caxias do Sul, visando a obtenção do grau de Doutora em Engenharia de Processos e Tecnologias, orientado por Prof. Dr. Matheus Poletto e coorientado por Prof.^a Dr.^a Gisele Cemin, com a colaboração da Prof.^a Dr.^a Vania Elisabete Schneider.

TESE APROVADA EM 28 DE MAIO DE 2026

Orientador: Prof. Dr. Matheus Poletto/Universidade de Caxias do Sul

Co-Orientadora: Prof.^a Dr.^a Gisele Cemin/Universidade de Caxias do Sul

Colaboradora: Prof.^a Dr.^a Vania Elisabete Schneider/Universidade Federal de Sergipe

Banca Examinadora:

Prof. Dr. André Quintão de Almeida/Universidade Federal de Sergipe

Prof.^a Dr.^a Camila Baldasso/Universidade de Caxias do Sul

Prof. Dr. Leandro Luís Corso/Universidade de Caxias do Sul

Prof. Dr. Roger Dias Gonçalves/Universidade Federal de Sergipe

AGRADECIMENTOS

Agradeço à minha família, na figura dos meus pais e irmã pelo apoio na trajetória que precedeu esta etapa. Agradeço especialmente ao meu marido por me acompanhar e auxiliar ao longo deste período. Ter vocês ao meu lado, fez toda a diferença.

Ainda, agradeço aos meus orientadores pela confiança e perseverança. Ao Programa de Pós-graduação em Engenharia de Processos e Tecnologias da Universidade de Caxias do Sul, agradeço a oportunidade.

LISTA DE SIGLAS

IFRC	<i>International Federation of Red Cross and Red Crescent Societies</i>
INDECI	<i>Instituto Nacional de Defensa Civil</i>
INGEMMET	<i>Instituto Geológico, Minero y Metalúrgico</i>
MAXENT	Máxima Entropía
ODS	Objetivos de Desarrollo Sustentável
PCM	<i>Presidencia del Consejo de Ministros</i>
SENAMHI	<i>Servicio Nacional de Meteorología e Hidrografía del Perú</i>
SRTM	<i>Shuttle Radar Topography Mission</i>
TRI	<i>Terrain Ruggedness Index</i>
TWI	<i>Topographic Wetness Index</i>
UN	<i>United Nations</i>
UNDRR	<i>United Nations Office for Disaster Risk Reduction</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	OBJETIVO GERAL	14
1.2	OBJETIVOS ESPECÍFICOS	14
1.3	MOTIVAÇÃO, INOVAÇÃO E RELEVÂNCIA	15
2	REVISÃO BIBLIOGRÁFICA	17
2.1	DESASTRES NATURAIS	17
2.1.1	Políticas de redução de riscos naturais	21
2.2	FENÔMENO EL NIÑO	22
2.3	BACIAS HIDROGRÁFICAS DE RIOS EFÊMEROS	26
2.3.1	Huaicos	30
2.3.2	Ocorrências de <i>huaicos</i> no Peru	34
2.4	<i>MACHINE LEARNING</i>	35
2.4.1	Classificação de modelos	40
2.4.2	Tipos de algoritmos	41
2.4.3	Estado da arte	48
3	MATERIAIS E MÉTODOS	50
3.1	ÁREA DE ESTUDO	51
3.2	MAPEAMENTO TEMÁTICO	52
3.2.1	Dados históricos de ocorrência de <i>huaicos</i>	52
3.2.2	Variáveis ambientais	53
3.3	ALGORITMOS SELECIONADOS	56
3.3.1	MaxEnt	57
3.4	CONFIABILIDADE DOS RESULTADOS	58
4	RESULTADOS E DISCUSSÕES	59
4.1	VARIÁVEIS AMBIENTAIS	59
4.1.1	Altitude	60
4.1.2	Declividade	61
4.1.3	Aspecto	62
4.1.4	Terrain Ruggedness Index - TRI	64
4.1.5	Topographic Wetness Index - TWI	65
4.1.6	Geologia	66
4.1.7	Geomorfologia	68
4.1.8	Pedologia	70
4.1.9	Ecossistemas	72
4.1.10	Distância a partir de cursos hídricos	73
4.1.11	Distância a partir de áreas urbanizadas	75
4.1.12	Precipitação média histórica	76
4.1.13	Correlação	78
4.2	MODELAGENS	81
4.2.1	Floresta Aleatória	81
4.2.2	Máquina de Vetores de Suporte	92
4.2.3	MaxEnt	100
4.2.4	Redes Neurais Artificiais	110
4.2.5	Regressão Logística	118
4.2.6	Considerações a respeito das modelagens	127
5	CONCLUSÕES	131
	REFERÊNCIAS	133

APÊNDICES

Apêndice A – Identifying areas at risk of huaicos occurrence in the department of Lima, Peru

Apêndice B – Ocorrência de huaicos no Peru: Uma análise dos fatores desencadeantes

LISTA DE FIGURAS

Figura 1. Correntes marítimas na América. Fonte: Molina-Schiller, Rosales, Freitas (2005).	23
Figura 2. Fenômeno El Niño espacializado ao longo do globo. Fonte: SENAMHI (2020).	24
Figura 3. Precipitação nos fenômenos de El Niño em 1998, 2017, 2023 e 2024. Fonte: SENAMHI (2025).	25
Figura 4. Bacias hidrográficas do Peru. Fonte: PCM (2003).	27
Figura 5. Regimes de escoamentos de cursos hídricos. Fonte: XAVIER (2020).	29
Figura 6. Áreas secas ao redor do planeta. Fonte: UNEP-WCMC (2007).	30
Figura 7. Figura que exemplifica o funcionamento de um huaico em uma quebrada. Fonte: Passuni (2019).	31
Figura 8. Exemplificação de huaico. Fonte: Ojo Peru (2023).	32
Figura 9. Huaico atravessando uma área urbanizada e alcançando o Oceano em Trujillo. Fonte: La Mula (2017).	32
Figura 10. Tipos de movimentos de massa. Fonte: Lumen Learning (2021).	34
Figura 11. Inventário das ocorrências de huaicos no Peru. Fonte: INGEMMET (2023).	35
Figura 12. Fluxograma do desenvolvimento do estudo. Fonte: a autora (2026).	51
Figura 13. Inventário das ocorrências de huaicos no departamento de Lima. Fonte: INGEMMET (2023).	52
Figura 14. Fluxograma de elaboração dos modelos a partir dos algoritmos regressão logística, floresta aleatória, máquina de vetores de suporte e redes neurais artificiais no QGIS.	56
Figura 15. Fluxograma do processo de modelagem no MaxEnt	58
Figura 16. Altitude. Fonte: adaptado de SRTM (2000).	60
Figura 17. Declividade. Fonte: a autora (2026).	62
Figura 18. Aspecto. Fonte: a autora (2026).	63
Figura 19. TRI – Terrain ruggedness index. Fonte: a autora (2026).	64
Figura 20. TWI – Topographic wetness index. Fonte: a autora (2026).	66
Figura 21. Classificação geológica. Fonte: INGEMMET (2022).	67
Figura 22. Classificação geomorfológica. Fonte: INGEMMET (2016).	69
Figura 23. Classificação pedológica. Fonte: MINAM (2009).	71
Figura 24. Ecossistemas do Peru. Fonte: MINAM (2019).	73
Figura 25. Distância a partir dos cursos hídricos. Fonte: adaptado de NGI (2024).	74
Figura 26. Distância a partir das áreas urbanizadas. Fonte: adaptado de NGI (2024).	75
Figura 27. Precipitação média mensal histórica. Fonte: SENAMHI (2024).	77
Figura 28. Resultados de suscetibilidade considerando os modelos mensais de Floresta Aleatória. Fonte: a autora (2026).	84
Figura 29. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos Floresta Aleatória. Fonte: a autora (2026).	85
Figura 30. Contribuição das variáveis para cada modelo mensal de Floresta Aleatória, considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	88
Figura 31. Contribuição das variáveis para cada modelo mensal de Floresta Aleatória, considerando as variáveis ambientais indicadas pelo INGEMMET e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	90
Figura 32. Resultados de suscetibilidade considerando os modelos mensais de Máquina de Vetores de Suporte. Fonte: a autora (2026).	94
Figura 33. Áreas com e sem suscetibilidade de ocorrência de huaicos, segundo os modelos de Máquina de Vetores de Suporte. Fonte: a autora (2026).	95
Figura 34. Contribuição das variáveis para cada modelo mensal de Máquina de Vetores de	

Suporte considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).	97
Figura 35. Contribuição das variáveis para cada modelo mensal de Máquina de Vetores de Suporte, considerando as variáveis ambientais indicadas pelo INGEMMET e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	99
Figura 36. Resultados de suscetibilidade considerando os modelos mensais Maxent. Fonte: a autora (2026).	103
Figura 37. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos Maxent. Fonte: a autora (2026).	104
Figura 38. Contribuição das variáveis para cada modelo mensal do Maxent, considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	106
Figura 39. Contribuição das variáveis para cada modelo mensal do Maxent, considerando as variáveis ambientais indicadas pelo INGEMMET e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	108
Figura 40. Resultados de suscetibilidade considerando os modelos mensais Redes Neurais Artificiais. Fonte: a autora (2026).	112
Figura 41. Áreas com e sem suscetibilidade de ocorrência de huaicos pelos modelos Redes Neurais Artificiais. Fonte: a autora (2026).	113
Figura 42. Contribuição das variáveis para cada modelo mensal Redes Neurais Artificiais considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).	115
Figura 43. Contribuição das variáveis para cada modelo mensal Redes Neurais Artificiais considerando as variáveis ambientais indicadas pelo INGEMMET e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).	117
Figura 44. Resultados de suscetibilidade considerando os modelos mensais. Fonte: Santos et al. (2025).	120
Figura 45. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos de Regressão Logística. Fonte: a autora (2026).	121
Figura 46. Contribuição das variáveis para cada modelo mensal Regressão Logística considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).	124
Figura 47. Contribuição das variáveis para cada modelo mensal de Regressão Logística, considerando as variáveis ambientais indicadas pelo INGEMMET e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).	126

LISTA DE QUADROS

Quadro 1. Trabalhos científicos que utilizam o algoritmo floresta aleatória para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos	43
Quadro 2. Autores que utilizaram máquina de vetores de suporte para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos	44
Quadro 3. Autores que utilizaram o programa MaxEnt para áreas de deslizamentos e inundações	45
Quadro 4. Autores que utilizaram o algoritmo redes neurais artificiais para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos.....	47
Quadro 5. Artigos que utilizam o algoritmo regressão logística para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos	48
Quadro 6. Caracterização geológica.....	68
Quadro 7. Caracterização geomorfológica.....	70
Quadro 8. Caracterização pedológica.....	72
Quadro 9. Resumo dos resultados obtidos	129

LISTA DE TABELAS

Tabela 1. Danos históricos a pessoas decorrentes de desastres naturais	18
Tabela 2. Danos materiais históricos decorrentes de desastres naturais.....	19
Tabela 3. Ocorrências por tipo de perigo	19
Tabela 4. Eventos de El Niño e sua intensidade.....	26
Tabela 5. Registros por departamento de ocorrências de huaicos em 2023	34
Tabela 6. Quantificação da altitude	61
Tabela 7. Quantificação da declividade.....	62
Tabela 8. Quantificação do aspecto	63
Tabela 9. Quantificação do TRI.....	65
Tabela 10. Quantificação do TWI	66
Tabela 11. Quantificação dos ecossistemas.....	73
Tabela 12. Quantificação das distâncias a partir dos cursos hídricos.....	74
Tabela 13. Quantitativos com relação às distâncias a partir de áreas urbanizadas	73
Tabela 14. Matriz de correlação de Pearson	79
Tabela 15. Matriz de correlação de Spearman.....	80
Tabela 16. Qualidade dos modelos de Floresta Aleatória	82
Tabela 17. Coincidência entre modelos mensais Floresta Aleatória – Variáveis segundo o INGEMMET (%).....	85
Tabela 18. Coincidência entre modelos mensais Floresta Aleatória – Variáveis segundo a bibliografia (%).....	86
Tabela 19. Qualidade dos modelos de Máquina de Vetores de Suporte	92
Tabela 20. Coincidência entre modelos mensais Máquina de Vetores de Suporte – Variáveis segundo o INGEMMET (%)	95
Tabela 21. Coincidência entre modelos mensais Máquina de Vetores de Suporte – Variáveis segundo a bibliografia (%)	96
Tabela 22. Qualidade dos modelos Maxent.....	101
Tabela 23. Coincidência entre modelos mensais Maxent – Variáveis segundo o INGEMMET (%).....	104
Tabela 24. Coincidência entre modelos mensais Maxent – Variáveis segundo a bibliografia (%)	105
Tabela 25. Qualidade dos modelos Redes Neurais Artificiais	110
Tabela 26. Coincidência entre modelos mensais Redes Neurais Artificiais – Variáveis segundo o INGEMMET (%)	113
Tabela 27. Coincidência entre modelos mensais Redes Neurais Artificiais – Variáveis segundo a bibliografia (%)	114
Tabela 28. Qualidade dos modelos de Regressão Logística.....	119
Tabela 29. Coincidência entre modelos mensais Regressão Logística – Variáveis segundo o INGEMMET (%).....	121
Tabela 30. Coincidência entre modelos mensais Regressão Logística – Variáveis segundo a bibliografia (%).....	122

RESUMO

O fenômeno conhecido no Peru como *huaico* caracteriza-se pela formação de fluxo superficial em áreas de relevo acentuado, sem cobertura vegetal e com formação geológica e geomorfológica frágil. Disso decorre o carreamento de materiais de granulometria variável arrastados pela água, fenômeno recorrente no país. A ocorrência dos *huaicos* está associado à urbanização das áreas mais planas, localizadas nas zonas de deposição aluvial mais próximas ao litoral, resultando, muitas vezes, em perdas socioeconômicas significativas. A identificação das áreas com suscetibilidade de ocorrência pode contribuir para o desenvolvimento de políticas de redução de risco de desastres. Neste contexto, no presente trabalho, foram utilizados algoritmos de *machine learning* para estimar a suscetibilidade de ocorrência de *huaicos* na região do departamento de Lima, no Peru. A pesquisa foi conduzida por meio da identificação de fatores ambientais desencadeantes do fenômeno, aliada ao uso do inventário de ocorrências. A abordagem de modelos mensais para os algoritmos e a comparação entre variáveis desencadeantes apontadas na bibliografia e relacionadas ao fenômeno buscam identificar a melhor representação da ocorrência de *huaicos*. Objetivou-se, por esta via, gerar uma metodologia válida para a identificação das áreas com suscetibilidade de ocorrência de *huaicos*. Para tanto, foram utilizados os ambientes do QGIS, com o uso da ferramenta Scikit-Learn, envolvendo os algoritmos Floresta Aleatória, Máquina de Vetores de Suporte, Redes Neurais Artificiais e Regressão Logística, e o MaxEnt, com vistas a gerar mapas de suscetibilidade. Os resultados envolvem a caracterização ambiental da área de estudo, incluindo os registros históricos de ocorrências de *huaicos*, bem como as características de altitude, declividade, geologia, geomorfologia e precipitação apontadas como desencadeantes do fenômeno pelo INGEMMET (Peru). Informações de aspecto, TRI, TWI, pedologia, ecossistemas, distância a partir de cursos hídricos e áreas urbanizadas também foram utilizadas para a representação do fenômeno. Todos os modelos aplicados apresentaram resultados adequados quanto à qualidade do modelo e à validade, com base no índice Kappa. Os modelos que consideram as variáveis do INGEMMET e o algoritmo Maxent apresentaram os melhores índices de qualidade (média de 0,92 de 1,00) e o índice Kappa (média de 0,96 de 1,00). Nessa condição, cerca de 26% das áreas do departamento de Lima apresentam suscetibilidade de ocorrência de *huaicos*. Destaca-se, ainda, que, apesar da variação decorrente da precipitação média mensal, 95% das áreas apontadas com suscetibilidade de ocorrência de *huaicos* nos modelos do Maxent são coincidentes. Para esses modelos, as variáveis mais importantes foram a geologia, com participação média de 58%, seguida da altitude e da precipitação média mensal. Embora a precipitação tenha apresentado participações elevadas em todos os modelos, a presença dessa variável não se tornou clara nas mudanças nos resultados dos meses. Dessa forma, apesar dos resultados terem sido altamente satisfatórios, sugere-se uma avaliação futura sem essa variável. Destaca-se, ainda, que os Objetivos do Desenvolvimento Sustentável abrangidos por este trabalho são os de erradicação da pobreza (1), de saúde e bem-estar (3), de educação de qualidade (4), de água potável e saneamento (6), de cidades e comunidades sustentáveis (11) e de ação contra a mudança global do clima (13).

Palavras-chave: Floresta Aleatória, Máquina de Vetores de Suporte, MaxEnt, Redes Neurais Artificiais, Regressão Logística.

ABSTRACT

The phenomenon known in Peru as huaico is characterized by the formation of surface flow in areas of steep relief, lacking vegetation cover, and possessing fragile geological and geomorphological formations. This results in the transport of material with variable grain size swept away by water—a recurring phenomenon in the country. The occurrence of huaicos is associated with the urbanization of flatter areas located in alluvial deposition zones near the coast, often resulting in significant socioeconomic losses. Identifying areas with a probability of occurrence can contribute to the development of disaster risk reduction policies. In this context, the present study utilized machine learning algorithms to estimate the probability of huaico occurrence in the department of Lima, Peru. The research was conducted by identifying environmental factors that trigger the phenomenon, combined with the use of an occurrence inventory. A monthly modeling approach for the algorithms and a comparison between triggering variables identified in the literature sought to determine the best representation of huaico occurrences. The objective was to generate a valid methodology for identifying areas prone to these events. To this end, QGIS environments were utilized alongside the Scikit-Learn tool, involving the Random Forest, Support Vector Machine, Artificial Neural Networks, and Logistic Regression algorithms, as well as MaxEnt, to generate probability maps. The results involve the environmental characterization of the study area, including historical records of huaico occurrences, as well as characteristics of altitude, slope, geology, geomorphology, and precipitation identified as triggers by INGEMMET (Peru). Information on aspect, TRI, TWI, pedology, ecosystems, and distance from watercourses and urbanized areas was also used to represent the phenomenon. All applied models presented adequate results regarding model quality and validity based on the Kappa index. The models considering the INGEMMET variables and the MaxEnt algorithm showed the best quality indices (average of 0.92 out of 1.00) and Kappa indices (average of 0.96 out of 1.00). Under these conditions, approximately 26% of the areas in the department of Lima show a probability of huaico occurrence. Furthermore, despite the variation resulting from average monthly precipitation, 95% of the areas identified with a probability of huaico occurrence in the MaxEnt models are coincident. For these models, the most important variables were geology, with an average contribution of 58%, followed by altitude and average monthly precipitation. Although precipitation showed high contributions in all models, the presence of this variable did not clearly influence changes in the results across different months. Thus, while the results were highly satisfactory, a future evaluation without this variable is suggested. Finally, the Sustainable Development Goals covered by this work are Poverty Eradication (1), Good Health and Well-being (3), Quality Education (4), Clean Water and Sanitation (6), Sustainable Cities and Communities (11), and Climate Action (13).

Keywords: Random Forest, Support Vector Machine, MaxEnt, Artificial Neural Network, Logistic Regression.

1 INTRODUÇÃO

O aumento da temperatura ocasiona não apenas o aquecimento térmico dos ambientes, mas também se reflete em alterações no ciclo hidrológico. O aumento da temperatura, em particular nos oceanos, pode desencadear fenômenos como o El Niño no Oceano Pacífico, escalonando assim, os níveis de precipitação decorrentes disso.

Em regiões costeiras próximas à linha do Equador, como no Peru, caracterizadas pela aridez, eventos de El Niño podem resultar em aumento dos níveis de precipitação. O que poderia se configurar como um fenômeno positivo acaba, porém, por se tornar um risco para a população que habita a região.

Altas precipitações, associadas à formação geomorfológica das bacias de corpos hídricos efêmeros, características da região costeira peruana, pode resultar em fluxos compostos por água, material rochoso, material particulado, vegetação e resíduos diversos resultantes de edificações e atividades antrópicas. Os fluxos de lama ou corridas de lama são também chamados de *huaicos*¹ e ocorrem de forma recorrente no Peru, em decorrência da precipitação que ativa o escoamento superficial dos rios efêmeros, particularmente nas zonas denominadas regionalmente de *quebradas*, carreando materiais em direção aos exutórios das bacias, onde, em alguns casos, situam-se áreas urbanizadas, causando prejuízos tanto de natureza ambiental quanto econômicos e sociais.

Apesar da recorrência desses desastres, a urbanização na região costeira do Peru vem se desenvolvendo de forma desordenada, sem atentar aos riscos de novos impactos decorrentes dessas corridas de lama. No entanto, já existem, a nível mundial, federal e provincial, políticas e planos para a gestão e prevenção de riscos de desastres, sendo necessária a compatibilização com o fator local. Nesse contexto, torna-se importante mapear as zonas e perigos relacionados às bacias das *quebradas*, tanto para ações específicas para essas áreas, quanto para direcionar a urbanização futura da região, com vistas a minimizar e mitigar os impactos dos *huaicos*

Algoritmos de *machine learning*, como floresta aleatória (*Random Forest*), máxima entropia (MaxEnt), máquina de vetores de suporte (*Support Vector Machine*), Naïve Bayes e redes neurais artificiais (*Artificial Neural Network*), têm sido aplicados com o viés de identificação de áreas com suscetibilidade de ocorrência dos mais variados tipos de desastres

¹ Termo de origem peruana, caracteriza a massa de lama e pedras desprendidas das montanhas em decorrências de chuvas torrenciais que ao alcançar os cursos d'água causam inundações (Real Academia Española, 2022a);

²Termo que caracteriza as fendas na montanha (Real Academia Española, 2022b).

naturais, como deslizamentos, erosão, inundações, queimadas naturais e terremotos. O uso dessa ferramenta é justificado pela possibilidade de avaliar grandes áreas, além da potencialidade de identificar relações e padrões entre dados, como as características ambientais e antrópicas de uma área e as ocorrências de desastres.

Existe um número muito pequeno de bibliografia fazendo uso de algoritmos como regressão logística, redes neurais artificiais, máquina de vetores de suporte e floresta aleatória para fluxos de detritos. E não foram identificados trabalhos envolvendo Maxent para este mesmo fenômeno. Ainda estes algoritmos não possuem registro de aplicação nem de metodologia desenvolvida para os *huaicos* ocorrentes no Peru, muito menos consideram distintos conjuntos de variáveis para comparação, nem possuem uma abordagem mensal da precipitação.

A área de estudo do departamento de Lima foi escolhida como área amostra para a execução do trabalho, devido ao fato de que, historicamente, cerca de 25% dos huaicos ocorridos no país estiveram ali concentrados. Ademais, a aplicação não ocorreu em todo o país devido ao alto poder de processamento exigido para essa atividade, o que inviabilizou sua aplicação em toda essa extensa área.

1.1 OBJETIVO GERAL

Desenvolver e validar um modelo de *machine learning* para identificar áreas com suscetibilidade de huaicos no departamento de Lima, Peru.

1.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos estão identificados a seguir:

- a) Identificar e mapear os fatores ambientais e climáticos que contribuem para a ocorrência de *huaicos*;
- b) Elaborar uma metodologia de mapeamento de identificação de áreas com suscetibilidade de ocorrência de *huaicos* com os algoritmos identificados pela utilização de machine learning;
- c) Validar a metodologia de identificação de áreas com suscetibilidade de ocorrência de *huaicos* com os algoritmos identificados pela utilização do Índice Kappa e índices de qualidade dos modelos de *machine learning*;
- d) Avaliar os resultados obtidos em termos de área suscetível à ocorrência de *huaicos* no departamento de Lima.

1.3 MOTIVAÇÃO, INOVAÇÃO E RELEVÂNCIA

A abordagem adotada para este trabalho, envolveu a criação de modelos mensais para a reflexão da importância da precipitação quase nula em determinados meses do ano. Esperava-se como resposta uma grande influência da precipitação, como sendo um dos elementos que desencadeia a ocorrência de huaicos, não sendo esta encontrada na literatura. Além disso, são poucos os trabalhos onde são identificadas as suscetibilidades de ocorrência de fenômenos catastróficos que abordam fluxo de detritos por meio de *machine learning*. Especificamente, em relação ao algoritmo Maxent, não há aplicações para huaicos, com exceção do artigo “*Identifying areas at risk of huaicos occurrence in the department of Lima, Peru*” (Santos, et al., 2025). Além das particularidades destacadas, este trabalho introduz outro aspecto diferenciado a abordagem do tema: a comparação entre os fatores apontados pela bibliografia em geral como desencadeantes dos fluxos de detritos e os fatores apontados pelo INGEMMET (Instituto Geológico, Minero y Metalúrgico) do Peru como preponderantes para a formação de *huaicos*.

Considera-se que a temática do uso de *machine learning* para a identificação de áreas com suscetibilidade a ocorrência de desastres naturais é um tema de crescente interesse. Nesse contexto, uma busca na base de periódicos da Capes, os termos “*machine learning*”, “suscetibilidade” e os desastres naturais de “deslizamentos”, “erosão”, “incêndios”, “inundações” e “terremotos”, resultou na identificação de um crescimento exponencial de publicações entre 2010 e 2024. Somente nos últimos 5 anos, o aumento do número de publicações, alcançou cerca de 1.100 artigos contendo esses temas. Isso demonstra, tanto o aumento do interesse pelo assunto, quanto sua aplicabilidade, relacionada a crescente importância da prevenção do risco de desastres, que deve ser iniciada pelo conhecimento dos perigos.

O Peru é um país atingido por diversos fenômenos naturais catastróficos, entre eles os *huaicos*. Só em 2023 foram registrados 235 *huaicos* no País (INDECI, 2024). Embora sejam recorrentes e existam registros do INDECI e um inventário do INGEMMET (2023), não se identifica um mapeamento de áreas com suscetibilidade de ocorrência de *huaicos* ao longo do território peruano. O próprio inventário possui registros que retrocedem somente até 1998, sendo que as condições de ocorrência de huaicos são bastante específicas e podem existir áreas onde ainda não foram atendidas essas condições para a formação de um huaico, ou simplesmente não houve registro dessa informação. Sendo assim, um mapeamento completo dos departamentos onde há altas densidades de ocorrências de *huaicos* auxiliaria especialmente no ordenamento territorial, com vistas a prevenir a ocupação de áreas com suscetibilidade de ocorrência desse fenômeno.

Alinhando-se aos Objetivos do Desenvolvimento Sustentável (ODSs) (UN, 2015) este trabalho envolve o ODS 1 - erradicação da pobreza, pela meta 1.5 de “até 2030, construir a resiliência dos pobres e daqueles em situação de vulnerabilidade, e reduzir a exposição e vulnerabilidade destes a eventos extremos relacionados com o clima e outros choques e desastres econômicos, sociais e ambientais”; ODS 3 - de saúde e bem-estar, com a meta 3.d de “reforçar a capacidade de todos os países, particularmente os países em desenvolvimento, para o alerta precoce, redução de riscos e gerenciamento de riscos nacionais e globais de saúde” e, ODS 4 - de educação de qualidade, pela meta 4.7 de

até 2030, garantir que todos os alunos adquiram conhecimentos e habilidades necessárias para promover o desenvolvimento sustentável, inclusive, entre outros, por meio da educação para o desenvolvimento sustentável e estilos de vida sustentáveis, direitos humanos, igualdade de gênero, promoção de uma cultura de paz e não violência, cidadania global e valorização da diversidade cultural e da contribuição da cultura para o desenvolvimento sustentável.

Também permeiam este trabalho o objetivo 6 de água potável e saneamento, considerando a meta 6.1 de “até 2030, alcançar o acesso universal e equitativo a água potável e segura para todos” e meta 6.2 de “até 2030, alcançar o acesso a saneamento e higiene adequados e equitativos para todos, e acabar com a defecação a céu aberto, com especial atenção para as necessidades das mulheres e meninas e daqueles em situação de vulnerabilidade”. O trabalho alinha-se ainda ao objetivo 13 de ação contra a mudança global do clima, através da meta 13.1 de “reforçar a resiliência e a capacidade de adaptação a riscos relacionados ao clima e às catástrofes naturais em todos os países”, meta 13.2 “integrar medidas da mudança do clima nas políticas, estratégias e planejamentos nacionais” e a meta 13.3 melhorar a educação, aumentar a conscientização e a capacidade humana e institucional sobre mitigação, adaptação, redução de impacto e alerta precoce da mudança do clima.

No ODS 11 de cidades e comunidades sustentáveis são destacadas as metas de

- 11.1 Até 2030, garantir o acesso de todos à habitação segura, adequada e a preço acessível, e aos serviços básicos e urbanizar as favelas;
- 11.3 Até 2030, aumentar a urbanização inclusiva e sustentável, e as capacidades para o planejamento e gestão de assentamentos humanos participativos, integrados e sustentáveis, em todos os países;
- 11.4 Fortalecer esforços para proteger e salvaguardar o patrimônio cultural e natural do mundo;
- 11.5 Até 2030, reduzir significativamente o número de mortes e o número de pessoas afetadas por catástrofes e substancialmente diminuir as perdas econômicas diretas causadas por elas em relação ao produto interno bruto global, incluindo os desastres relacionados à água, com o foco em proteger os pobres e as pessoas em situação de vulnerabilidade;
- 11.6 Até 2030, reduzir o impacto ambiental negativo per capita das cidades, inclusive prestando especial atenção à qualidade do ar, gestão de resíduos municipais.

2 REVISÃO BIBLIOGRÁFICA

Para alcançar os objetivos deste trabalho, neste capítulo serão abordados, inicialmente, os temas de desastres naturais, políticas de redução de riscos naturais, o fenômeno El Niño, as bacias hidrográficas de rios efêmeros, os huaicos e sua ocorrência no Peru, e o machine learning, com classificação geral de modelos e tipos de algoritmos.

2.1 DESASTRES NATURAIS

Os desastres naturais são caracterizados, conforme a *International Federation of Red Cross and Red Crescent Societies* (IFRC, 2022) como graves interrupções no funcionamento de uma comunidade que excedem sua capacidade de resolver com seus próprios recursos, decorrentes da exposição à perigos naturais. Prasad e Franciscutti (2017) acrescentam ainda que os desastres estão associados a esgotamento de recursos, mortalidade e morbidade. Esses desastres podem ser originados por diferentes eventos desencadeantes, sendo que os desastres naturais podem ser ocasionados por eventos de ordem geofísica, hidrológica, climatológica, meteorológica ou biológica (IFRC, 2022).

Só no Peru, no ano de 2017, cerca de 295.000 pessoas precisaram deixar suas residências em decorrência dos fenômenos naturais (INTERNAL DISPLACEMENT MONITORING CENTRE, 2020). Em 2017, a ocorrência do evento de *El Niño* costeiro surpreendeu o país, com precipitações anômalas e subsequente geração de *huaicos* que atingiram áreas urbanizadas.

Anualmente no Peru, são elevados os números de pessoas afetadas das mais diversas formas por desastres naturais. No informativo mais recente da Defesa Civil peruana (INDECI – Instituto Nacional de Defensa Civil) de fevereiro de 2024 (INDECI, 2024), ocorreu no ano de 2017, o evento de El Niño costeiro, em que mais de 2 milhões de pessoas foram atingidas. Embora desde 2003, alguns anos tenham menos de 300.000 pessoas afetadas por desastres naturais, nesse período, a média foi de 950.000 pessoas por ano. Em 7 dos 20 anos com dados disponíveis, foram mais de 1 milhão de habitantes do Peru atingidos por desastres naturais, chegando nos anos de 2016 e 2017 a ultrapassar 2 milhões de pessoas, totalizando entre 2003 e 2023 mais de 20 milhões de habitantes atingidos (Tabela 1).

Além disso, destacam-se os números de desaparecidos, lesionados e óbitos, que totalizaram, ao longo de 20 anos, respectivamente: 419, 14.239 e 3.302 casos. Em 25% dos 20

anos de registros, houve mais de 1.000 pessoas lesionadas em decorrência dos desastres naturais e em 85% dos anos do período analisado, houve mais de 100 óbitos, sendo que só no ano de 2007, foram registrados 628 óbitos.

Tabela 1. Danos históricos a pessoas decorrentes de desastres naturais

Ano	Danos a pessoas			
	Afetados	Desaparecidos	Lesionados	Óbitos
Total	20.116.550	419	14.239	3.302
2003	249.791	30	392	215
2004	920.027	28	284	142
2005	876.847	17	370	121
2006	999.364	6	166	87
2007	1.664.976	11	2.436	628
2008	1.368.056	34	273	165
2009	942.279	39	196	109
2010	834.320	53	2.489	130
2011	899.216	19	252	121
2012	866.751	31	377	180
2013	727.189	18	168	137
2014	309.984	17	123	73
2015	1.427.719	18	418	162
2016	2.236.473	5	335	103
2017	2.089.624	23	683	225
2018	1.114.113	6	228	162
2019	166.995	13	1.636	79
2020	355.544	14	1.391	110
2021	408.925	6	269	87
2022	338.888	16	320	102
2023	1.319.469	15	1.433	164

Fonte: INDECI (2024).

Com relação aos danos materiais decorrentes dos desastres naturais descritos na Tabela 2, o ano de 2017 destaca-se novamente, apresentando um elevado número de habitações afetadas, e/ou destruídas. No entanto, observa-se que esse fenômeno ocorrido em 2017 não foi o único evento responsável por perdas, com capacidade de afetar tantas pessoas e estruturas. Os próprios números de 2023 são extremamente elevados, tanto com relação as pessoas afetadas, quanto aos danos materiais.

Além disso, o INDECI destaca os tipos de desastres naturais que ocorrem no país, conforme dados apresentados na Tabela 3. Em especial, os *huaicos*, foram registrados 1.549 eventos nos últimos 10 anos. Embora as ocorrências de *huaicos* representem cerca de 3% dos registros anuais de desastres naturais, seus impactos, geralmente associados as áreas urbanizadas não podem ser ignorados.

Tabela 2. Danos materiais históricos decorrentes de desastres naturais

Ano	Danos a habitações		Danos a centros de saúde		Danos em áreas de cultivos	
	Afetados	Destruídos	Afetados	Destruídos	Afetados	Destruídos
Total	2.057.244	351.796	5.066	409	2.807.355	956.891
2003	35.122	8.564	82	8	36.771	14.056
2004	31.987	6.110	78	1	253.058	91.274
2005	30.275	9.059	57	12	192.558	14.804
2006	53.922	5.518	31	27	75.652	8.718
2007	65.239	96.377	87	10	32.310	5.719
2008	151.793	15.542	217	14	103.588	18.098
2009	62.461	6.609	29	8	6.375	276
2010	44.556	12.255	69	8	31.228	5.636
2011	101.712	20.899	343	5	70.274	24.562
2012	167.245	12.942	277	13	48.031	45.516
2013	59.267	4.430	90	0	35.491	12.960
2014	33.421	2.624	54	0	214.611	19.692
2015	148.854	5.869	311	12	191.106	25.311
2016	65.996	3.967	143	3	384.093	95.957
2017	393.675	38.602	1.180	33	136.532	55.411
2018	29.068	2.651	111	2	95.034	25.862
2019	67.775	9.676	197	15	47.062	33.016
2020	59.065	6.183	149	12	164.151	63.274
2021	84.928	13.075	295	55	155.486	99.664
2022	43.864	11.478	103	37	124.219	74.539
2023	328.019	59.366	1.163	134	409.665	223.546

Fonte: INDECI (2024).

Tabela 3. Ocorrências por tipo de perigo

Perigo	Ano									
	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Total	2.618	3.420	4.235	6.675	4.363	3.801	4.350	6.342	5.708	10.336
Atividade vulcânica	12	3	0	1	1	8	3	0	0	10
Avalanche	4	10	8	14	3	1	1	0	0	3
Baixas temperaturas	512	911	1.201	705	1.298	651	989	1.087	863	1.390
Derrame	2	9	16	5	7	2	4	1	11	6
Derrubada em colinas	20	19	9	122	53	92	56	141	159	161
Deslizamento	185	227	64	321	226	253	189	428	386	448
Erosão	45	86	51	63	58	4	29	99	130	354
<i>Huaico</i>	51	96	83	570	55	200	73	110	76	235
Incêndio florestal	47	73	280	143	248	72	311	165	382	342
Inundação	156	283	125	407	168	335	267	321	250	302
Chuva intensa	1.011	1.130	757	3.644	1.263	1.607	1.328	2.324	1.902	4.807
Maré	7	16	7	9	3	1	0	8	3	15
Seca	27	28	848	54	355	1	93	2	157	333
Sismo	36	28	56	47	66	156	15	280	93	39
Tempestade com raios	13	20	25	55	25	18	23	55	27	81
Ventos fortes	490	481	705	515	534	400	969	1.321	1.269	1.810

Fonte: adaptado de INDECI (2024).

O risco associado aos huaicos está na localização das áreas de ocorrência e áreas atingidas, a região costeira do Peru, pois essa região concentra cerca de 65% da população do país (French et al., 2020).

Somente entre 2003 e o primeiro trimestre de 2021 foram registradas 2421

ocorrências de huaicos no país (INDECI, 2021) atingindo aproximadamente 450.000 pessoas e danificando aproximadamente 37.000 habitações, além de 141 óbitos. Embora sejam eventos recorrentes a localização destes é bastante variável ao longo do tempo. Somente no primeiro trimestre de 2021, o departamento com maior quantidade de huaicos registradas foi o departamento de Cusco, com 41 registros. No primeiro trimestre de 2019, foram registrados 83 huaicos no departamento de Lima, sendo esta, a região com maior recorrência nesse ano (INDECI, 2019; 2021). Apesar da imprevisibilidade espacial de ocorrência, o departamento do Peru mais atingido por huaicos é o departamento de Lima, onde se localiza a capital, com 24% das ocorrências entre 2003 e 2021, foram 584 huaicos afetando 38.100 pessoas (INDECI, 2021).

No entanto, o IFRC (2022) acrescenta ainda que enquanto os perigos são inevitáveis, os desastres não são. Neste ponto é que entram as políticas de redução de riscos, tanto para preparação quanto para resposta a desastres naturais.

Segundo a UNDRR (United Nations Office for Disaster Risk Reduction, 2022) risco de desastre pode ser entendido como

“...A perda potencial de vidas, lesões ou bens destruídos ou danificados que podem ocorrer a um sistema, sociedade ou comunidade, em um período de tempo específico, determinado probabilisticamente em função do perigo, exposição, vulnerabilidade e capacidade.”

Segundo mesmo autor, os componentes do risco são definidos como:

- Capacidade: combinação dos pontos fortes, atributos e recursos disponíveis dentro de uma organização, comunidade ou sociedade para gerenciar e reduzir os riscos de desastres e fortalecer a resiliência;
- Exposição: localização de pessoas, infraestrutura, habitação, capacidades de produção e outros bens humanos tangíveis localizados em áreas propensas a riscos;
- Perigo: processo, fenômeno ou atividade que pode causar perda de vidas, ferimentos ou outros impactos à saúde, danos à propriedade, perturbação social e econômica ou degradação ambiental;
- Vulnerabilidade: as condições determinadas por fatores ou processos físicos, sociais, econômicos e ambientais que aumentam a suscetibilidade de um indivíduo, uma comunidade, ativos ou sistemas aos impactos dos perigos.

Destaca-se que este trabalho está focado no potencial risco associado a ocorrência, com intuito de identificar as áreas com maior suscetibilidade de ocorrência do fenômeno. Ao final, as áreas com suscetibilidade de ocorrência serão cruzadas com as áreas urbanizadas e as áreas modeladas de aumento de urbanização, identificando assim a exposição ao fenômeno.

2.1.1 Políticas de redução de riscos naturais

A Estrutura Sendai para Redução do Risco de Desastres 2015–2030 foi adotada na Terceira Conferência Mundial das Nações Unidas sobre Redução do Risco de Desastres, realizada em 2015 em Sendai, Japão (UN, 2015). A Conferência representou uma oportunidade para os países participantes adotarem uma estrutura pós 2015 concisa, focada, voltada para o futuro com vistas à ação para redução de riscos de desastres; avaliar a experiência adquirida através das estratégias nacionais e regionais para redução do risco de desastres e identificar modalidades de cooperação para implementar posterior a 2015 (UN, 2015).

As ações guiam para alcançar a minimização dos riscos de desastres segundo UN (2015) passam por:

- I. Compreender o risco de desastres;
- II. Fortalecer a governança do risco de desastres para gerenciar o risco de desastres;
- III. Investir na redução de desastres para resiliência;
- IV. Melhorar a preparação para desastres para uma resposta eficaz e “reconstruir melhor” a recuperação, reabilitação e reconstrução.

No âmbito nacional no Peru, existem outros instrumentos reguladores sobre a gestão de risco: A Lei n° 29.664 (PCM, 2011) instituiu o Sistema Nacional de Riscos de Desastres e o validado pelo Decreto Supremo n° 048/2011 (PCM, 2011) que aprova os regramentos da Lei n° 29.664. Nesses documentos são apontadas as entidades responsáveis pela gestão de risco e elaboração do plano e desempenho de atividades, dentre outras disposições.

Destaca-se ainda, o Plano Nacional de Gestão de Riscos, o qual contém uma estimativa do risco, a prevenção, a redução, a preparação, a resposta, a reabilitação e reconstrução associadas aos riscos existentes no País. Além do plano nacional geral, ocorrem planos nos demais níveis governamentais, dentre os quais se destacam: Planos de prevenção e redução de risco de desastres, de preparação, de operação de emergência, de educação comunitária, de reabilitação e de contingência. A última versão do Plano Nacional data do período de 2014 a 2021, conforme o Decreto Supremo n° 034 de 2014 (PCM, 2014).

Uma avaliação feita sobre a gestão de desastres no Peru, com relação aos eventos de *El Niño* apontou os fatores que contribuem para agravar os impactos desses fenômenos, além das causas naturais de ocorrência. Segundo French et al., (2020), ocorreu uma massiva preparação para o evento de *El Niño* do ciclo 2015-2016, com a aplicação de 1,5 bilhões de

dólares em obras de manutenção de canais, abertura de novos canais, reforço de estruturas em canais, manutenção e reforço de estruturas abastecimento de água, esgoto, estradas, pontes, escolas, hospitais e estruturas geodinâmicas. No entanto, o evento de 2015-2016 não ocorreu na intensidade esperada e, dessa forma, não foi possível avaliar a efetividade das medidas de preparação tomadas. Porém, em 2017, um inesperado evento de *El Niño* atingiu a região costeira do Peru, causando 3,1 bilhões de dólares em perdas, além de deixar 138 feridos. Após o ocorrido, houve a necessidade de entender o motivo desse evento ter sido tão catastrófico no país. French et al. (2020) apontaram que conhecer o risco, assim como recomendado na Estrutura Sendai, ainda se configura um dos fatores agravantes dos impactos dos eventos de *El Niño* no Peru. Mesmo sendo possível estimar os ciclos de ocorrência da fase *El Niño* do sistema ENOS, o início, a intensidade e a duração do fenômeno ainda são imprevisíveis.

Áreas atingidas por *huaicos* no evento de 2017, como Trujillo iniciaram movimentos de preparação para novos eventos. Obras de contenção nas quebradas e túneis para desvio da água são as soluções em desenvolvimento adotadas pela província para minimizar os impactos da ocorrência de *huaicos* (Gobierno Regional La Libertad, 2024). No entanto, destaca-se novamente a importância da não ocupação de áreas com perigo como a maior prevenção possível.

2.2 FENÔMENO EL NIÑO

Na região próxima ao Equador, encontra-se a chamada zona de convergência intertropical, que é caracterizada por um sistema de baixa pressão, com chuvas em níveis elevados, como resultado da convergência dos alísios dos hemisférios Norte e Sul. Já na América do Sul, atuam massas equatoriais (Equatorial Continental e Equatorial Pacífica), tropicais (Tropical Atlântica, Tropical Pacífica e Tropical Continental) e extratropicais (Polar Pacífica e Polar Atlântica). As características desses sistemas são diferentes e podem causar graus variados de risco a tempestades à medida que esses se desenvolvem e interagem (NUNES, 2015). Esses elementos têm ações atmosféricas, impactam no ciclo hidrológico e estão intimamente ligados as demais variáveis do sistema, como as correntes marítimas.

O fenômeno *El Niño* - Oscilação Sul (ENOS), é caracterizado por duas fases: a de resfriamento do pacífico, na região equatorial, chamada *La Niña* e, *El Niño*, quando ocorre o aquecimento das águas frias trazidas pela Corrente de Humboldt, em que a temperatura do Oceano Pacífico apresenta aumento, especialmente na porção equatorial (RUESTA, OHISHI, OLIVEIRA, 2011). Essa segue em direção as regiões equatoriais, onde encontram correntes

de maior temperatura, percorrendo áreas costeiras da América do Sul e do Oceano Pacífico, conforme a Figura 1. Em consequência do aquecimento anormal dessas águas, a temperatura e a umidade das regiões tropicais são alteradas, provocando mudanças dos regimes de chuva, particularmente nas latitudes médias, conforme a Figura 2 (CPTEC/INPE, 2022).

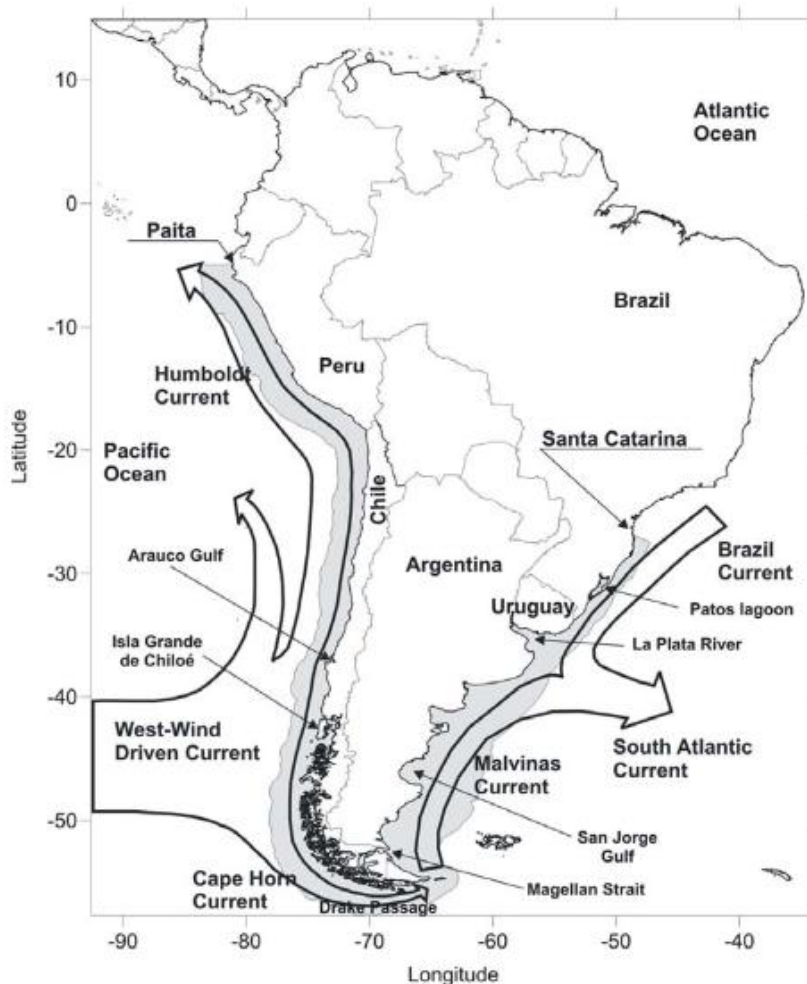


Figura 1. Correntes marítimas na América. Fonte: Molina-Schiller, Rosales, Freitas (2005).

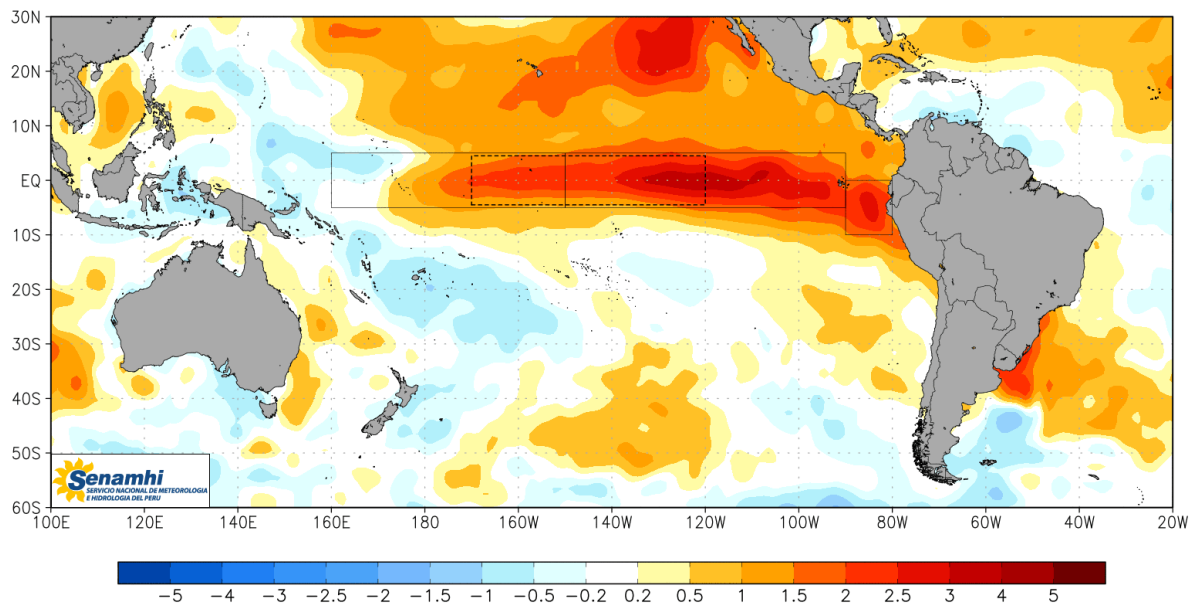


Figura 2. Fenômeno El Niño espacializado ao longo do globo. Fonte: SENAMHI (2020).

No Peru, devido a localização do país com relação ao fenômeno *El Niño*, conforme SENAMHI (2025), em 1998, o fenômeno foi classificado como extraordinário em nível de chuvas, com precipitações acima de 2.500 mm no período de janeiro, fevereiro e março. Já o fenômeno de 2017, foi classificado pelo SENAMHI (2025) como forte, onde ocorreram precipitações totais de até 1.000 mm no mesmo período, conforme apresentado na Figura 3. Ainda, na mesma figura, observam-se os eventos ocorridos em 2023, classificados como de intensidade forte, e o evento ocorrido no final de 2023, início de 2024, de intensidade moderada. Nesses dois eventos, que impactaram parte das regiões costeiras do Peru, os níveis de precipitação chegaram a alcançar até 800 mm no intervalo de três meses.

A intensidade do fenômeno *El Niño* reflete diretamente na ocorrência de *huaicos*, sendo que estes, normalmente ocorrem em anos de *El Niño*. Lembrando que o Peru possui boa parte de seu território considerado como áreas áridas e semiáridas, ou seja, com baixos níveis de precipitação anual. A alteração desses níveis, no caso, o aumento desses, desencadeia problemas que atingem tanto a área urbanizada quanto as áreas de uso agrícola ou de cobertura natural. No entanto, nas áreas urbanizadas, a concentração de pessoas e edificações potencializa os impactos desses eventos.

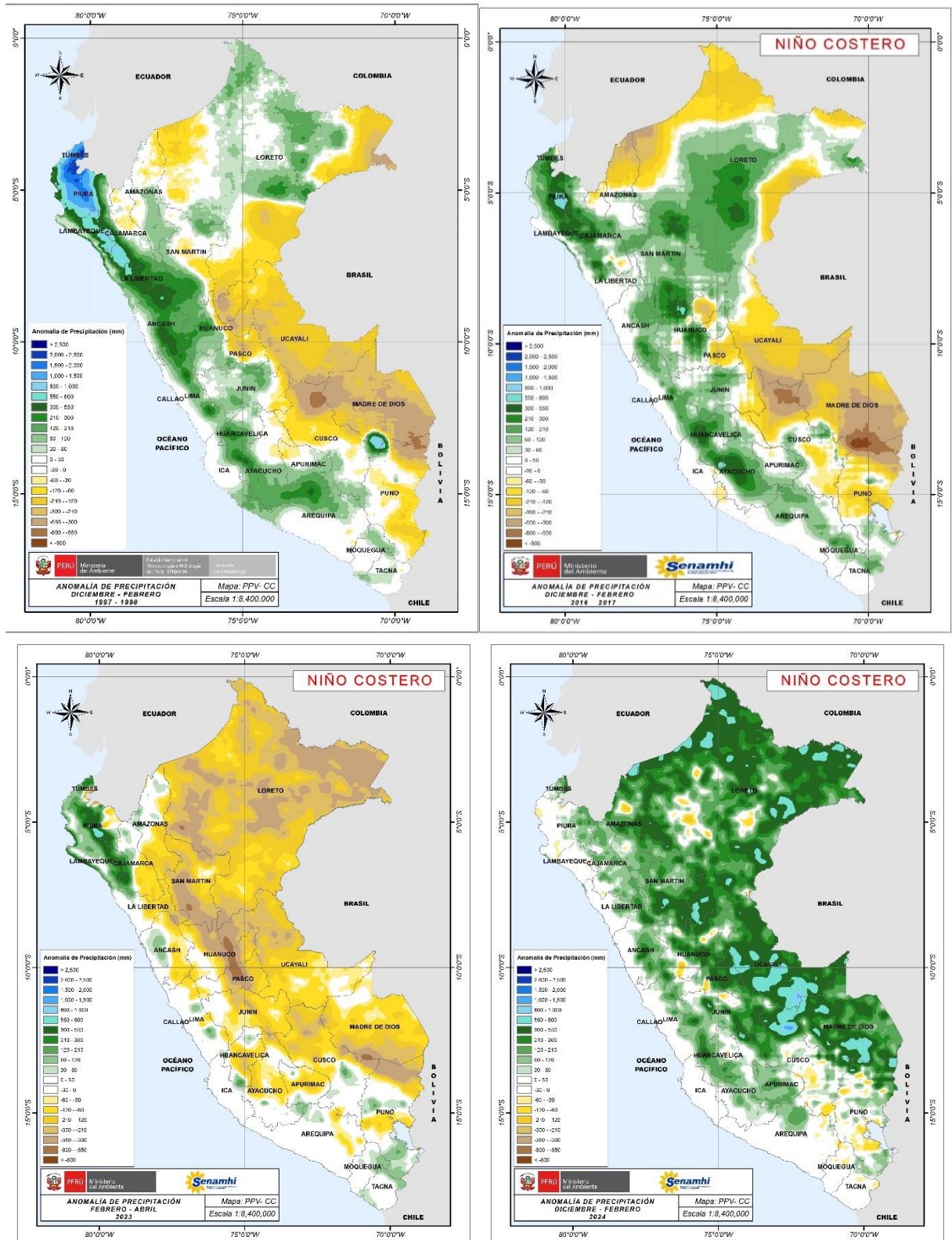


Figura 3. Precipitação nos fenômenos de El Niño em 1998, 2017, 2023 e 2024. Fonte: SENAMHI (2025).

Chen, Zhao e Zhi (2024) identificaram a recorrência dos eventos de *El Niño* entre 1875 e 2023. Além das características de cada evento, os autores classificaram as intensidades desses eventos, conforme dados apresentados na Tabela 4, destacando-se os ciclos de 1892-

1895, 1908-1911, 1916-1918, 1954-1956 e 1998-2001, considerados extremamente fortes. Embora identifique-se uma recorrência de eventos com duração de cerca de 1 ano, também espaçados cerca de 1 ano entre si, há uma imprevisibilidade com relação ao início do ciclo, duração e intensidade, tornando as práticas de prevenção de risco de desastres naturais extremamente importantes para serem adotadas e cultivadas nas áreas que são impactadas por essas condições climáticas.

Tabela 4. Eventos de El Niño e sua intensidade

Evento	Intensidade	Evento	Intensidade	Evento	Intensidade
1875-1876	Moderado	1924-1925	Moderado	1984-1985	Fraco
1878-1879	Fraco	1926-1927	Fraco	1988-1989	Forte
1879-1880	Moderado	1933-1934	Moderado	1995-1996	Fraco
1886-1887	Moderado	1938-1939	Moderado	1998-2001	Extremamente forte
1889-1890	Forte	1942-1943	Moderado	2005-2006	Fraco
1892-1895	Extremamente forte	1944-1945	Extremamente fraco	2007-2008	Moderado
1898-1899	Fraco	1946-1946	Extremamente fraco	2008-2009	Fraco
1903-1904	Moderado	1948-1948	Extremamente fraco	2010-2011	Forte
1906-1907	Fraco	1949-1950	Fraco	2011-2012	Moderado
1908-1911	Extremamente forte	1954-1956	Extremamente forte	2016-2016	Extremamente fraco
1915-1916	Fraco	1964-1965	Fraco	2017-2018	Fraco
1916-1918	Extremamente forte	1970-1972	Forte	2020-2021	Moderado
1921-1921	Fraco	1973-1974	Forte	2021-2023	Forte
1922-1923	Fraco	1974-1976	Forte		

Fonte: Chen, Zhao e Zhi (2024).

2.3 BACIAS HIDROGRÁFICAS DE RIOS EFÊMEROS

A hidrografia do Peru é subdividida em 3 grandes bacias hidrográficas: a do Lago Titicaca, a do Atlântico e a do Pacífico, conforme se observa na Figura 4. Especialmente, nas sub-bacias do Pacífico, são encontradas características de regimes de escoamento efêmero, associadas diretamente ao clima árido e superárido da região. É nessa última que se encontra também o departamento de Lima.

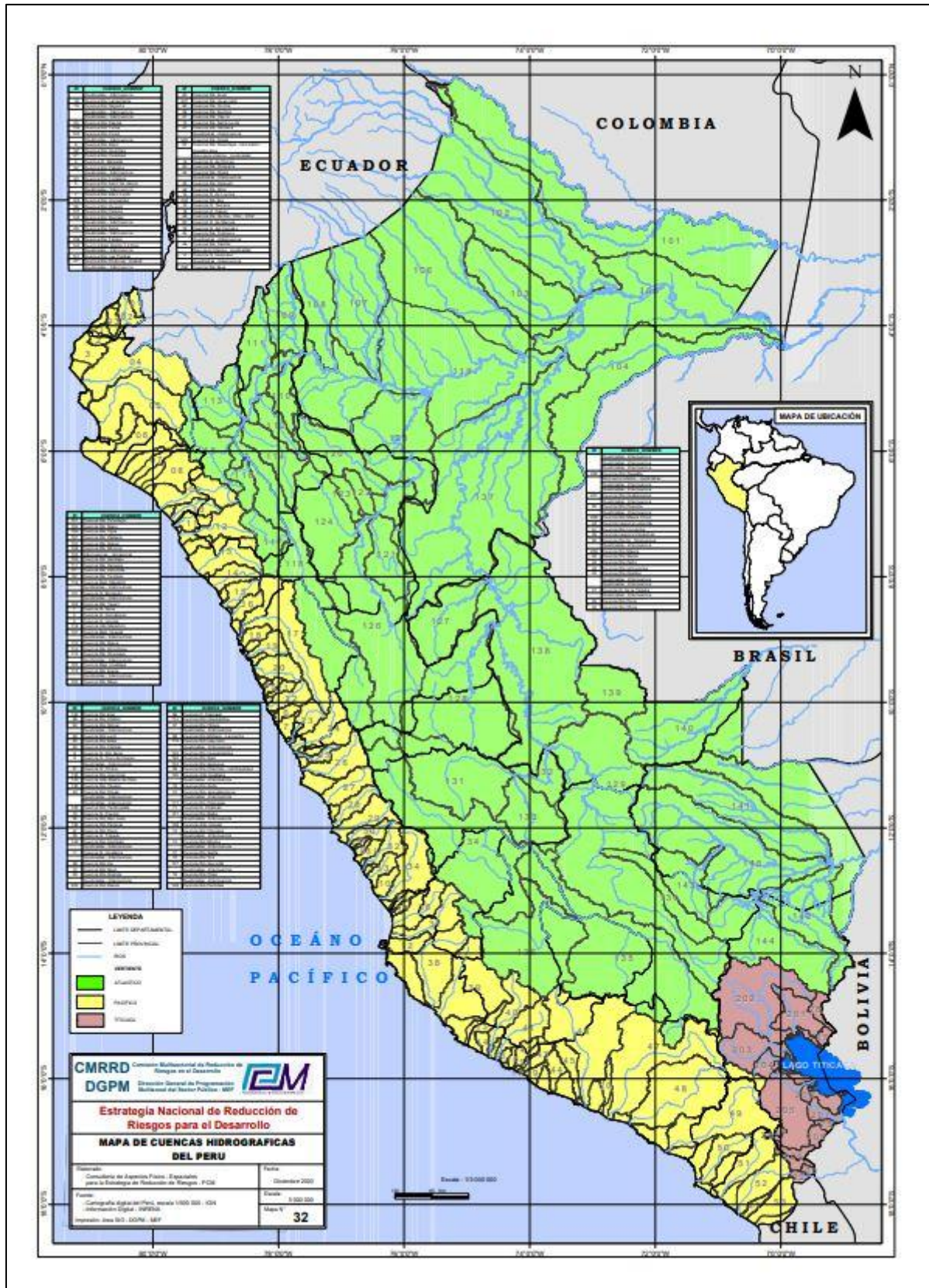


Figura 4. Bacias hidrográficas do Peru. Fonte: PCM (2003).

Dentre as classificações e características atribuídas aos cursos hídricos de uma bacia hidrográfica, encontra-se o regime de escoamento, que pode ser perene, intermitente e

efêmero, condição essa, relacionada com a permanência de fluxo no leito do rio (ASWM, 2014). De maneira simplificada, os cursos hídricos perenes apresentam fluxo contínuo ao longo do ano todo, sendo alimentados pelo fluxo basal em períodos sem precipitação, o que contribui para o aumento da vazão superficial. Os cursos hídricos intermitentes são corpos d'água que sazonalmente deixam de apresentar fluxo superficial, cujas alterações de regime podem estar associadas à ausência de fluxo proveniente de água subterrânea de vertentes ou ao derretimento de geleiras (EPA, 2008; ASWM, 2014).

Os cursos hídricos de regime intermitente possuem fluxo contínuo somente em certos períodos do ano, conforme exemplificado na Figura 5. Esses períodos sazonais podem estar associados a fluxo provenientes de água subterrânea de vertentes, ou derretimento de geleiras (EPA, 2008). Durante o período seco podem ocorrer trechos intercalados de leito seco e leito com vazão (EPA, 2008). Da mesma forma, em cursos hídricos perenes, o fluxo ocasionado por precipitações é uma vazão complementar, não configurando o fluxo de base (ASWM, 2014).

Os cursos hídricos efêmeros, por sua vez, só possuem vazão, ou só existem, por um pequeno período após eventos de precipitação, em resposta direta a estes (EPA, 2008). Nesses cursos, não há alimentação por águas subterrâneas, estando essas áreas mais elevadas do que o ponto de alimentação, conforme indicado na Figura 5 (ASWM, 2014). Essa condição pode ocorrer em pequenos trechos de uma bacia ou em toda a extensão da bacia.

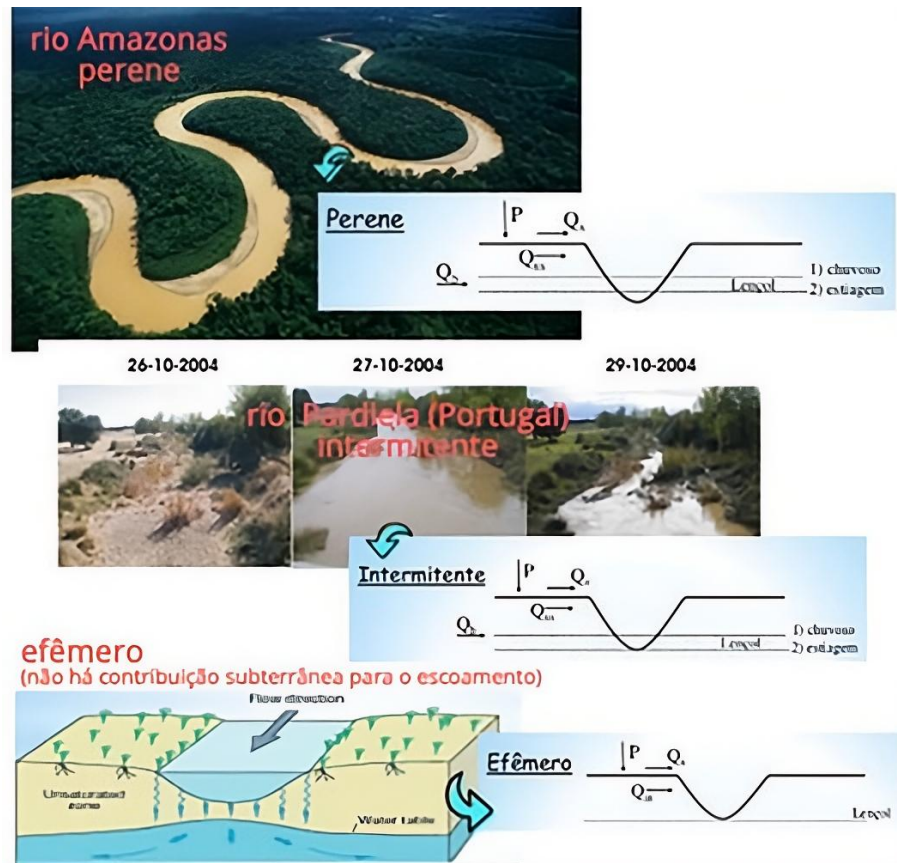


Figura 5. Regimes de escoamentos de cursos hídricos. Fonte: XAVIER (2020).

Cursos hídricos efêmeros e intermitentes, são encontrados em regiões áridas e semiáridas, chamadas de “terras secas”, do inglês *dryland* (EPA, 2008). Essas áreas, de acordo com a FAO (2018), correspondem a áreas onde evapotranspiração supera a precipitação. A classificação dessas áreas é feita com base no índice de aridez (IA), que é a razão entre a precipitação anual e a evapotranspiração potencial média. Ainda, conforme a FAO (2018), 41% das terras do planeta configuram-se como terras secas, que são áreas classificadas como hiperáridas, áridas, semiáridas e subúmidas secas, todas as regiões em que os níveis de precipitação são inferiores à evapotranspiração (Figura 6). Apesar disso, cerca de 2 bilhões de pessoas vivem em áreas de regiões áridas, além de estas regiões sustentarem 50% da pecuária mundial. As terras secas possuem, 35% das áreas de *hotspot* de biodiversidade global.

Particularmente, na América do Sul, podem ser encontradas zonas áridas e super áridas, como no Chile e no Peru, e no Nordeste brasileiro, especialmente nas áreas costeiras desses países, ambos localizados na porção oeste do continente americano. Parte dessas áreas, estão sob a ação direta do fenômeno *El Niño-Oscilação Sul* (ENOS) (FAO, 2018).

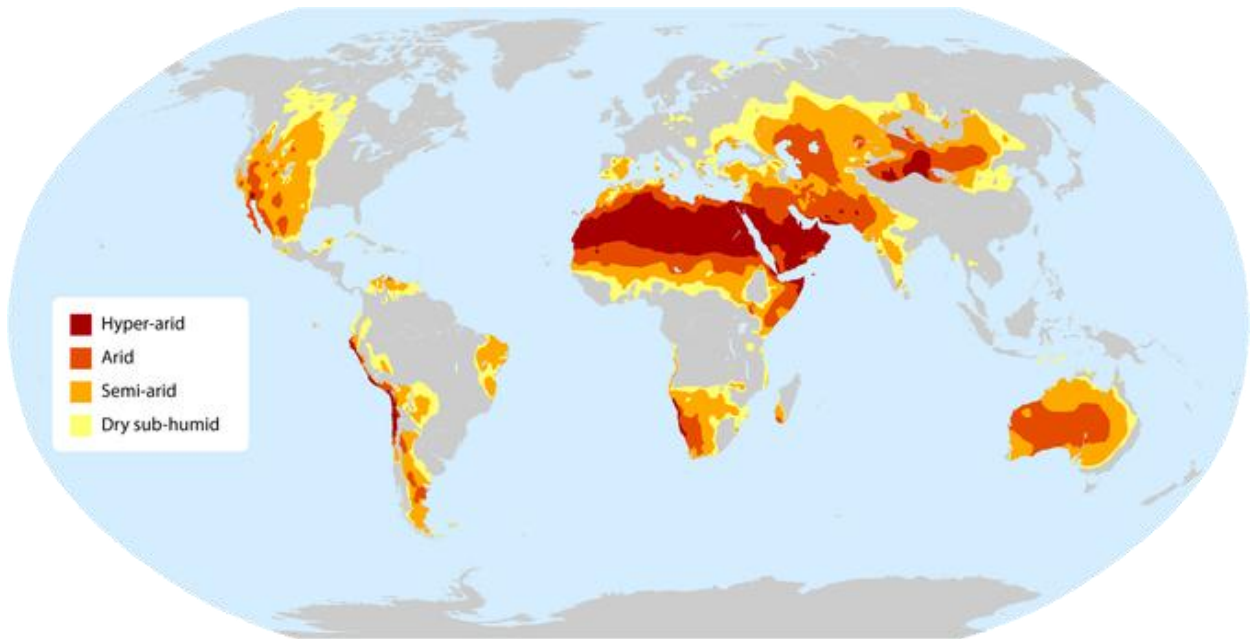


Figura 6. Áreas secas ao redor do planeta. Fonte: UNEP-WCMC (2007).

2.3.1 Huaicos

No Peru, utiliza-se o termo *huaico* para descrever fluxos aluviais, os quais se configuram como fluxos de escoamento superficial ocasionados por precipitação, normalmente em áreas de declividade acentuada, acompanhados do arraste de partículas sólidas, provenientes de solos pouco coesivos ou quedas e deslizamentos de rochas (Rivera; Vilchez; Vela, 2018). Portanto, as características do fenômeno dificultam sua classificação formal nos fenômenos apresentados acima.

A ocorrência dos *huaicos* está associada às bacias montanhosas de rios efêmeros, ativadas em eventos de precipitação elevada. Os *huaicos* por sua vez, descem por gravidade até as áreas mais planas, inundando essas áreas, as quais usualmente abrigam regiões urbanizadas. Na Figura 7 é explicado como se comporta o fluxo gerado na bacia de cursos hídricos efêmeros e os riscos ocasionados para as zonas de *quebrada*.



Figura 7. Figura que exemplifica o funcionamento de um *huaico* em uma *quebrada*. Fonte: Passuni (2019).

Na Figura 8 observa-se a exemplificação de *huaicos*, ainda na região montanhosa de acúmulo de fluxo. Observa-se o arraste tanto de água, quanto de materiais de pequenas e grandes granulometrias. Nas regiões de vale, a água já escoou pelas áreas de declividade mais acentuada e carregou os materiais superficiais e, são nesses locais onde ocorre o acúmulo do fluxo nas áreas no leito do rio efêmero. A Figura 9 exemplifica um *huaico* atravessando uma área urbanizada. Quanto mais longo o trajeto do fluxo e mais plano, menos material de granulometria grosseira é carregado, chegando ao exutório da bacia, que por vezes é o oceano, um fluxo menos viscoso. Ainda assim, ao longo de toda a passagem, ocorre o carregamento de muito material para as áreas urbanizadas, causando danos desde alagamentos até danificação das residências por materiais grosseiros. Além do óbvio impacto de perda de bens materiais de

residências, comércios, indústrias, infraestrutura pública, problemas de deslocamento, acesso a saneamento, alimentos e medicamentos, até veiculação de agentes contaminantes.



Figura 8. Exemplificação de huaico. Fonte: Ojo Peru (2023).



Figura 9. Huaico atravessando uma área urbanizada e alcançando o Oceano em Trujillo. Fonte: La Mula (2017).

O termo movimento de massa, de que se tratam os *huaicos* refere-se as variedades de processos em que uma grande porção de materiais da superfície da Terra se move, por gravidade, de um ponto a outro. Normalmente os processos de movimentação de massas envolvem a presença de água como elementos desencadeante. Esse conjunto contribui significativamente para a alteração da paisagem (Wilson, 1998).

Embora os movimentos de massa sejam condicionados e desencadeados pela geomorfologia, pedologia, geologia e fatores climáticos, a ação antrópica altera a condição natural de intensidade e frequência desse fenômeno. Assim, a ocupação inadequada de áreas de

com risco de ocorrência de movimentos de massa, tanto altera a dinâmica local, quanto potencializa o perigo sobre a população e infraestruturas (Louzeiro et al., 2019; Perdomo; Ladeira, 2010).

Não existe, no entanto, uma classificação unificada desses movimentos de massas (Lopes; Arruda Junior, 2015), mas alguns processos possuem características e nomenclatura recorrentes. As quedas são geralmente associadas a material rochoso; os rastejos caracterizam movimentos constantes, sazonais ou intermitentes; já os escorregamentos, que podem ser planares, circulares ou em cunha, e as corridas movimentam solo, rochas, detritos e água (Augusto Filho, 1992).

Os escorregamentos de terra são configurados pela movimentação de solo onde identifica-se uma separação entre o material instável e o material base mais estável, que pode ser tanto rotacional quanto translacional. O movimento rotacional consiste em uma superfície de ruptura côncava com movimento paralelo a superfície e transversal ao longo do deslizamento (Figura 10). O movimento translacional ocorre de forma praticamente planar com relação ao material base. Atenta-se ao fato que os escorregamentos de terra são mais propícios em relevos acentuados e em climas secos ou semiáridos que, mediante a chuvas abundantes criam as condições ideais para a ocorrência desse fenômeno, possuindo elevada força destrutiva (Perdomo; Ladeira, 2010; Lumen Learnig, 2021).

As quedas são movimentos abruptos de rochas ou pedregulhos que se desprendem de regiões de relevo acentuado e são ocasionados pela gravidade, pelo intemperismo e a presença de água intersticial (Figura 10). A separação ocorre em descontinuidades como fraturas e junções e o movimento pode ser de queda livre, ressalto ou rolamento (Lumen Learnig, 2021).

Os fluxos de lama, de detritos ou de terra correm em encostas com grande presença de materiais argilosos e pouca vegetação que possa reter a camada de solo, sob ação da gravidade e dependente da presença de água (Figura 10). Esse tipo de movimentação normalmente se incorpora aos cursos d'água, comportando-se como um fluido altamente viscoso, atingindo pontes, moradias e o que mais estiver em seu caminho (Perdomo; Ladeira, 2010; Lumen Learnig, 2021).

Existem ainda os movimentos lentos, que são as quedas de grandes rochas através de uma superfície curva, que ocorrem especialmente quando já há uma fragilidade na porção inferior do maciço, e os rastejos (Figura 10), causados pela tensão de cisalhamento, que podem ser sazonais, desencadeados pela umidade e temperatura do solo, ou contínuos, quando a tensão de cisalhamento é sempre superior a resistência do material (Lumen Learnig, 2021).

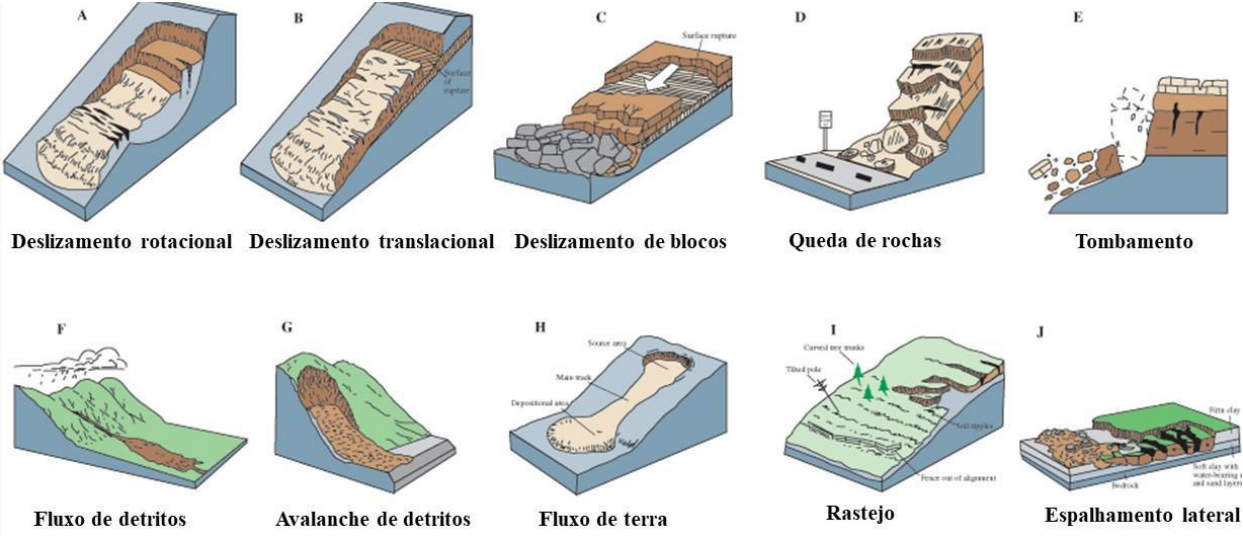


Figura 10. Tipos de movimentos de massa. Fonte: Lumen Learning (2021).

2.3.2 Ocorrências de *huaicos* no Peru

Somente em 2023 foram registrados 253 *huaicos* ao longo do território do Peru (Tabela 5). Os departamentos de Ancach e Lima, onde se localiza a capital, registraram a maior quantidade de *huaicos*. Os departamentos de Cajamarca, Ayacucho e Junin também possuem um número expressivo de ocorrências. Ainda, parte dos departamentos não possuía registros de ocorrências de *huaicos* (INDECI, 2024).

Tabela 5. Registros por departamento de ocorrências de *huaicos* em 2023

Departamento	Ocorrência de <i>huaicos</i>	Departamento	Ocorrência de <i>huaicos</i>	Departamento	Ocorrência de <i>huaicos</i>
Amazonas	12	Huancavelica	18	Madre de Dios	-
Ancash	34	Huánuco	3	Moquegua	6
Apuríamc	16	Ica	9	Pasco	7
Arequipa	12	Junin	18	Puno	3
Ayacucho	20	La Libertad	6	San Martin	1
Cajamarca	25	Lambayeque	-	Tacna	-
Callao	-	Lima	34	Tumbes	-
Cusco	11	Loreto	-	Ucayali	-

Fonte: INDECI (2024).

O Instituto Geológico, Minero y Metalúrgico (INGEMMET) possui no seu catálogo de metadados o inventário de riscos geológicos no Peru, constando registros históricos de quedas, deslizamentos, fluxos, movimentos complexos, tombamentos e outros. Neste inventário, os *huaicos* estão apontados como fluxo de detritos, embora a classificação não abarque totalmente a explicação do fenômeno.

Na Figura 11 são apontados 8.539 registros no período de 1998 até 2018. Observa-se que praticamente não existem registros de ocorrências de huaicos nos departamentos de Loreto (nordeste de Lima), Madre de Dios (sudeste de Lima) e Ucayali (leste de Lima), situados a oeste do país.

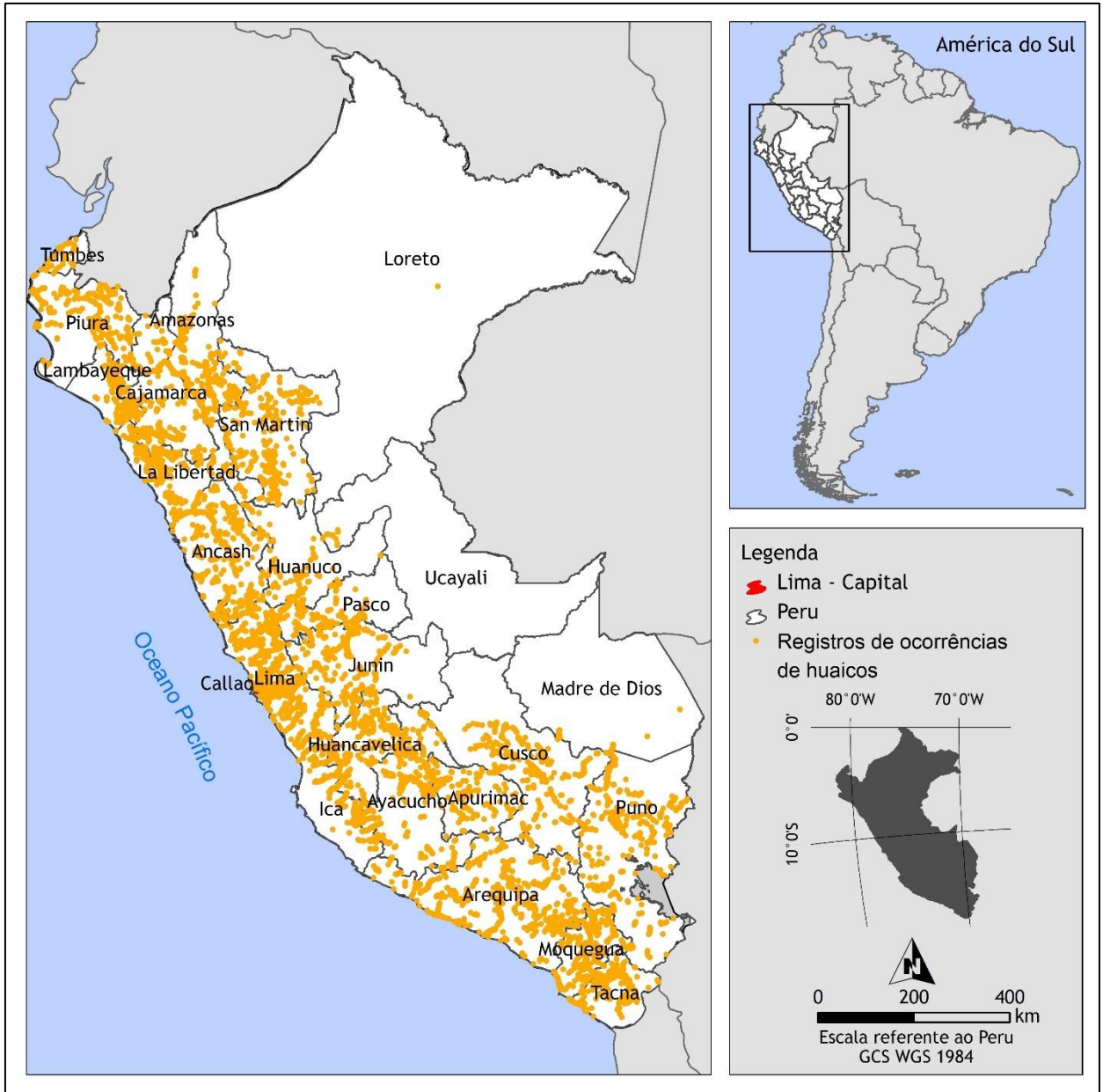


Figura 11. Inventário das ocorrências de huaicos no Peru. Fonte: INGEMMET (2023).

2.4 MACHINE LEARNING

Machine learning, ou Aprendizado de máquina, é um campo da inteligência artificial (IA) que se concentra em permitir que os sistemas aprendam com dados e melhorem seu desempenho sem programação explícita. Envolve o desenvolvimento de algoritmos estatísticos

que podem aprender com dados, identificar padrões e fazer previsões ou tomar decisões com intervenção humana mínima.

O algoritmo consiste em um conjunto de tarefas a serem executadas em sequência para transformar entradas em saídas. Os modelos, por sua vez, possuem o algoritmo como uma das suas partes, sendo a outra um banco de dados, ao qual é aplicado o algoritmo. Nesse sentido, um mesmo algoritmo, quando aplicado a diferentes dados, devolve um modelo diferente (Sicsú et al., 2023).

O aprendizado de máquina, do inglês *machine learning*, é uma coleção de metodologias, ferramentas e algoritmos computacionais com funcionalidade de ensinar as máquinas a analisarem, compreenderem e descobrirem padrões, por vezes ocultos nos dados, com o intuito de fazer previsões (Alkhatib et al., 2023). Liu et al. (2021) descrevem que o método *machine learning* simula a mente humana para decidir afiliações de dados com base nos dados conhecidos para modular relações entre os dados de treinamento e os hiperparâmetros inibindo e ativando dados.

Os métodos de *machine learning* geram um valor real ou uma previsão de suscetibilidade para uma amostra de teste e, posteriormente, comparam os valores previstos com uma classificação de corte, onde uma resposta positiva é atribuída se o valor de previsão for maior que o valor de corte e, caso contrário, a resposta é classificada como negativa (Liu et al., 2021).

Akel et al. (2021) acrescenta que as técnicas de *machine learning* fornecem significado aos dados brutos e resolvem problemas em que a quantidade de dados e/ou a complexidade dos padrões estatísticos implicam em que os humanos não seriam capazes de resolvê-los através de técnicas tradicionais de maneira confiável e eficiente. Além de que os algoritmos de *machine learning* aprendem com exemplos e, portanto, são treinados para encontrar padrões que podem ajudar na tomada de decisões e previsões baseadas em informações novas e não vistas.

Busca-se finalmente utilizar dados para autoaprendizagem, eliminando a necessidade de treinar computadores explicitamente, pois as máquinas treinadas podem aplicar padrões aprendidos para novos dados e aprimorar as previsões (Alkhatib et al., 2023).

O processo de *machine learning* compreende diversas etapas, destacando-se a coleta de dados e escolha de características apropriadas, a construção de máquinas de aprendizagem e avaliação dos modelos. Para os modelos de *machine learning* é imprescindível a existência de dados. Além disso, de maneira genérica, a qualidade do modelo é proporcional a quantidade de dados. Porém, não se deve desconsiderar também a qualidade dos dados. Um conjunto de dados de treinamento e caracterização com baixa qualidade ou insuficientes quantitativamente tendem

a resultar em respostas distorcidas. Dessa forma, a etapa inicial para a elaboração de um modelo confiável é a criação de um conjunto de dados também confiáveis (Alkhatib et al., 2023).

Normalmente a preparação dos dados envolve reformatar dados, fazer alterações e integrar conjuntos de dados para melhorá-los ou compatibilizá-los para servirem como amostras para treinamento do modelo (Alkhatib et al., 2023).

A etapa de modelagem envolve a seleção do algoritmo, otimização de hiperparâmetros e treinamento (Alkhatib et al., 2023). Diferencia-se também os parâmetros dos hiperparâmetros, sendo o primeiro a variável a qual busca-se a estimar e o segundo, os dados definidos antes de executar o algoritmo. Destaca-se ainda que, essa definição ocorre por tentativa e erro, através da configuração que alcança os melhores resultados para o modelo (Sicsú et al., 2023).

O ajuste de hiper parâmetros depende de vários fatores: tamanho da amostra, classificador, regressão modelos usados e tipo de modelo (Audebert; Le Saux; Lefèvre, 2019; Yang; Shami, 2020). É um passo adicional para melhorar a precisão e desempenho do modelo (Pannakkong et al., 2022). Por exemplo, seleção dos melhores recursos polinomiais em modelos de regressão, número de árvores em uma floresta aleatória, número de camadas e neurônios em uma rede neural, profundidade máxima em árvores de decisão e taxa de aprendizagem para descida gradiente (Pannakkong et al., 2022). Algumas técnicas comuns de ajuste de hiper parâmetros são pesquisa em grade, pesquisa aleatória, otimização bayesiana, otimização baseada em modelo sequencial, e algoritmos genéticos (Yang; Shami, 2020).

Dessa forma, numerosas iterações completam o problema, com modificações adicionais no conjunto de dados. Os ciclos de feedback entre os dados preparação e modelagem são elementos críticos para representar essas iterações (Alkhatib et al., 2023).

Os algoritmos de *machine learning* tem sido frequentemente utilizados no mapeamento e avaliação de numerosos processos naturais extremos. As abordagens baseadas em aprendizado de máquina são mais capazes de determinar relações precisas entre os dados espaciais e temporais além de prever condições futuras. Ao longo das duas últimas décadas, as abordagens de aprendizado de máquina evoluíram rapidamente. Já a combinação de aprendizado de máquina e sistemas de informação geográfica (SIG) produz resultados mais precisos (Arabameri et al., 2020).

Embora a quantidade de dados configure um fator relevante para a modelagem, as técnicas de aprendizado de máquina podem lidar com a escassez de dados e evitar o *overfitting* uma vez ajustada de forma adequada e correta e fornece uma previsão de alto desempenho. Os modelos de aprendizado de máquina atendem a quase todos os subcampos, particularmente na predição espacial de diferentes perigos naturais, exemplos dos quais pode ser encontrado na

literatura. A análise de perigos incide sobre a suscetibilidade espacial e temporal, na qual a predição espacial tem sido considerada há muito tempo uma parte essencial em caso de emergência e uma parte substancial da a complexa cadeia computacional de análise de risco (Satarzadeh et al., 2022).

Um indicativo importante para um modelo de *machine learning* é o viés (bias), o qual é dado pela diferença entre os valores observados (variável supervisora) e os resultados calculados pelo algoritmo. Quanto melhor o ajuste do modelo aos dados observados, menor é o viés. Um subajuste do modelo, quando o viés é alto, pode ser causado por um modelo simples demais para compreender a variável alvo. No entanto, dependendo o comportamento da variável alvo, embora um modelo mais complexo possa representar fielmente o conjunto de dados, pode representar bem inclusive os ruídos contidos na amostra. Além disso, esse pode se tornar sensível as variações amostrais, se tornando então a variância outro problema representativo nos modelos (Sicsú et al., 2023).

O *overfitting* ocorre quando o modelo aprende o ruído e os detalhes indesejados nodados de aprendizagem, o que impacta negativamente a previsão dos novos dados (Gu et al., 2016). *Underfitting* refere-se a um modelo que não modela os dados de treinamento nem generaliza a novos dados (Guyon; Yao, 1999). Ao comparar classificadores, enfatizamos que mais do que apenas a precisão global deve ser considerada; as precisões do usuário e do produtor para aulas individuais também deve ser considerado (Alnaim; Sun; Tong, 2022). Isto é particularmente verdadeiro se o mapeamento se concentrar em classes raras (isto é, classes de extensão limitada em os dados da imagem). Classes raras tendem a ter pouco efeito na precisão geral, mas pode, no entanto, ser vital para determinar a utilidade da classificação (Bogner et al., 2018).

Aplicações de inteligência artificial/aprendizado de máquina também estão sendo usadas para apoiar a análise de riscos e a redução de riscos sistêmicos. Novas arquiteturas de máquinas estão surgindo rapidamente para lidar com os requisitos especiais de redes neurais de convolução (redes neurais artificiais mais comumente aplicadas para analisar imagens visuais), aprendizagem profunda, treinamento e inferência. No entanto, como medir e comparar o desempenho da máquina para estes fins ainda é uma questão de muito debate (UNDRR, 2022). São diversas as abordagens possíveis para a utilização de *machine learning* que consideram o uso de aprendizado de máquina para avaliação de riscos naturais em diversos países do mundo.

Algoritmos de *machine learning*, como máquina de vetores de suporte, árvore de decisão, sistemas neuro-fuzzy, floresta aleatória, decisão árvore, regressão logística do kernel, redes neurais artificiais e máxima entropia são alguns dos modelos mais usuais para a avaliação

de riscos (Javidan et al., 2021; Shahabi et al., 2023).

De toda forma, Chang et al. (2019) destacam que apesar das diferentes estatísticas envolvidas, suas terminologias e capacidade computacional são similares, e que esses algoritmos se baseiam em grande parte nos pressupostos de que o passado é a chave para o futuro, além de fatores que envolvem os deslizamentos estão espacialmente ligados e, portanto, pode ser usado na previsão funções e, eventos futuros, os quais, provavelmente, acontecerão em condições semelhantes. Isso também se aplica as demais tipologias de fenômenos naturais.

Destaca-se, ainda, que uma base de dados robusta é essencial para o desempenho eficaz de algoritmos de aprendizado de máquina. Estudos mostram que a qualidade e a diversidade dos dados utilizados no treinamento dos modelos impactam diretamente a capacidade dos algoritmos em generalizar e evitar vieses. Dados faltantes, por exemplo, podem prejudicar significativamente o desempenho dos modelos, afetando tanto tarefas de classificação quanto de regressão (Kotsiantis et al., 2006). Além disso, a estruturação adequada dos dados, com a organização correta em subespaços de alta e baixa dimensão, pode aumentar a resiliência do modelo, ajudando a evitar a sobrecarga e a sobreajuste (Zhou, 2017). A ausência de uma base de dados bem estruturada e representativa pode comprometer os resultados dos algoritmos, resultando em modelos tendenciosos ou com desempenho desigual, como demonstrado por diversos estudos sobre o impacto de dados desequilibrados e viesados na performance dos sistemas de IA (Mehrabi et al., 2019). Assim, é evidente que a construção de bases de dados representativas, com diversidade e sem viés, é crucial para o sucesso e a confiabilidade dos modelos de aprendizado de máquina, garantindo que os resultados obtidos sejam precisos e justos.

A utilização de técnicas de aprendizado de máquina para a previsão e identificação de áreas suscetíveis a desastres naturais tem se tornado uma abordagem cada vez mais eficaz em diversos estudos. Essas técnicas permitem a análise de grandes volumes de dados geoespaciais e ambientais, resultando em modelos de previsão mais precisos e eficientes para deslizamentos de terra, inundações, fluxo de detritos e outros fenômenos. A seguir são destacados alguns exemplos de estudos que aplicam aprendizado de máquina para a identificação de áreas com suscetibilidade de ocorrência desses fenômenos.

Elbes, Alzu'bi e Kanan (2023) aplicaram deep learning para prever terremotos, Já Liu et al. (2025) utilizaram *machine learning* para a predição de incêndios florestais. Alcântara et al (2024) identificaram áreas com suscetibilidade de deslizamentos na região sudeste do Brasil. De maneira similar, no entanto mais robusta, tanto com relação as variáveis utilizadas quanto com relação aos algoritmos utilizados, Lokesh et al (2025) também identificaram áreas com

suscetibilidade de deslizamentos em Wayanad, na Índia. Ganjirad e Delavar (2023) exploraram o uso do algoritmo floresta aleatória e máquina de vetores de suporte para mapear a suscetibilidade a inundações na Lousiana nos Estados Unidos da América. Zhu, Guo e Huang (2024) estudaram o perigo de inundação em áreas urbanas na Tianjin na China com o uso de *machine learning*. Já trabalhos como o de Chauhan, Gupta e Dixit (2025) realizaram uma análise integrada para múltiplos riscos. Os autores avaliaram a suscetibilidade a sismos, deslizamentos e inundações relacionados ainda a vulnerabilidade socioeconômica no estado indiano de Uttarakhand.

2.4.1 Classificação de modelos

Huyen (2024) descreve os modelos de *machine learning* com viés de classificação e de regressão. Sendo que os modelos de classificação agrupam as entradas em categorias distintas. Já os modelos de regressão retornam como resultado um valor contínuo. Dessa forma, para modelos de regressão está associado um limiar de decisão.

Sicsú et al. (2023) classifica os modelos em dois grandes grupos, o dos algoritmos supervisionados e não supervisionados. Na aprendizagem supervisionada as máquinas são ensinadas a resolver problemas com a ajuda de humanos, que coletam e identificam dados para alimentar os sistemas, recaindo sobre a máquina a tarefa de examinar os dados para detectar padrões, classificar itens e avaliar se sua previsão está correta ou incorreta (Alkhatib et al., 2023). Em outras palavras, os modelos supervisionados envolvem uma base de dados em que a cada observação corresponde um conjunto de hiperparâmetros, geralmente denominados previsores e uma variável supervisora ou variável alvo, que indica a resposta esperada para o algoritmo. Assim o algoritmo identifica a relação entre as variáveis previsoras e a variável supervisora, que pode retornar resultados de classificação ou de regressão (Sicsú et al., 2023).

Já a aprendizagem não supervisionada envolve métodos principalmente para o agrupamento de dados não categorizados. Nesta categoria de aprendizagem, as máquinas aprendem a identificar padrões e tendências em dados não rotulados sem a supervisão de humanos (Alkhatib et al., 2023). Ou seja, não existe uma variável alvo para esta modalidade. A ausência de uma variável alvo torna o algoritmo mais complexo em relação a sua resolução e a sua interpretação também (Sicsú et al., 2023).

Pode-se, portanto, caracterizar os modelos de machine learning com relação a forma dos resultados. Os modelos do tipo output classifiers, ou hard classifiers, retornam diretamente a categoria à qual a resposta pertence. Já os *soft classifiers* ou *probability classifiers* apresentam

a estimativa da probabilidade que uma resposta pertença a uma categoria. Dessa forma, para se obter uma resposta binária (presença e ausência), deve-se adotar um valor de corte (Sicsú et al., 2023).

2.4.2 Tipos de algoritmos

Modelos de regressão linear ou múltipla preveem o valor de uma variável quantitativa dependente em função de outras variáveis independentes (ou preditoras). Um modelo de regressão linear possui apenas uma variável preditora, enquanto um modelo de regressão múltipla possui mais de uma variável preditora. Os parâmetros descritivos dos modelos de regressão podem ser obtidos através dos mínimos quadrados (Sicsú et al., 2023).

Os modelos podem ser de regressão linear, que são modelos mais populares em sistemas de informação geográfica (SIG) e sensoriamento remoto (Sudalaimuthu; Sudalayandi, 2019). A regressão linear ajuda a identificar e avaliar a relação entre dois ou mais fatores/covariáveis quando aproveitamos o poder do espaço em nossa análise usando as características de distância, por exemplo, a influência da distância até a água na seleção de habitat (Mansouri, et al., 2018). Este modelo ajuda a responder questões como por exemplo: se há uma “relação linear entre o diâmetro à altura do peito e o diâmetro da coroa de árvores”; “quais fatores demográficos contribuem para a utilização de taxas elevadas de transporte público?”; “quais fatores contribuem para a elevada propagação da COVID nas regiões geográficas?”; “qual a relação entre as alterações climáticas e a migração?”, entre outras.

Existem 3 tipos de regressão linear comumente usados em projetos baseados em sistemas de informação geográfica e sensoriamento remoto: distribuição Contínua (Gaussiana), Logística e Poisson. A distribuição deve ser normal (Susiluoto, et al., 2020) para a regressão linear da distribuição gaussiana, porém pode ser também chamado de contínuo porque a variável dependente pode assumir uma ampla gama de valores, como temperatura, precipitação e diâmetro (Avand; Moradi; Lasboyee, 2021). Se a variável dependente não tiver distribuição normal, pode-se reclassificá-la para valores binários (Shi; Li; Zhao, 2020).

Os modelos binários também conhecidos como modelos de regressão logística, são modelos com apenas dois resultados, presença/ausência (Ghosh et al., 2022). Já os modelos de regressão de contagem (Poisson) são recomendados quando as variáveis dependentes são as quantidades de ocorrências de um evento (Graff et al., 2020).

Nos modelos de regressão não linear, a relação entre as variáveis é justamente não linear (Liang et al., 2022). Como a maioria dos relacionamentos em O SIG e o sensoriamento

remoto são não lineares, esses modelos são amplamente utilizados neste setor (Adsuara et al., 2019).

Devido à sua flexibilidade uma grande variedade de modelos pode ser construída usando esta metodologia (Hsieh, 2020) como por exemplo: estudar as culturas e os processos do solo, estudar o preço dos imóveis e a imigração relação, estudar a relação entre diâmetro e cobertura do dossel (Hsieh, 2020).

2.4.2.1 Floresta aleatória

O *Random Forest*, floresta aleatória, é um classificador de conjunto eficaz, que constrói múltiplas decisões na forma de árvores para classificação utilizando um subconjunto de variáveis selecionadas aleatoriamente (Sicsú et al., 2023). Trata-se de um modelo de classificação supervisionada que pode ser aplicado à classificação e regressão de dados (Belgiu; Drăguț, 2016). É uma coleção de árvores de decisão que prevê os resultados com base em múltiplos modelos/submodelos (Pal, 2005).

O *Random Forest* é uma técnica de aprendizado de máquina que tem sido utilizada para resolver muitos problemas do mundo real, como monitoramento da cobertura da terra e previsão de riscos e de doenças (Chang et al., 2021). Na previsão de deslizamentos, a árvore aleatória também tem sido aplicada. Por se tratar de um método popular e de alto desempenho possui diversas vantagens como: ser um método não paramétrico baseado na natureza; ser capaz de determinar a importância de variáveis utilizadas, fornecer um algoritmo para estimar os valores faltantes e ser flexível para a análise de classificação, regressão e aprendizagem não supervisionada (Chang et al., 2019).

Na floresta aleatória, um subconjunto de variáveis preditoras é utilizado para construir cada árvore. E o número de árvores e o número de preditores usados para construir cada árvore podem ser diferentes, dependendo do conjunto de dados (Chang et al., 2019).

O algoritmo trabalha com base no princípio do ensacamento ao fazer modelos, o que significa que faz diferentes modelos baseados no subconjunto de dados de amostra de treinamento, e o resultado é baseado na maioria/média dos submodelos (Berhane et al., 2018). Alguns exemplos comuns de projetos que podem ser resolvidos usando o algoritmo floresta aleatória incluem: classificação da cobertura do solo e do uso da terra, extração de características como estradas, assentamentos, detecção de objetos como veículos, identificação de espécies de fauna e flora e modelagem de projetos relacionados, como projetos propensos a inundações/secas, classificação de tipos de solo, doenças e clima (Upreti, 2022).

Os modelos de árvores de classificação ou de regressão particionam gradativamente em subconjuntos homogêneos ou puros, que seguem até atenderem a um critério de parada definido inicialmente. O desafio da construção de uma árvore está na seleção das variáveis que gerarão as melhores partições. Destaca-se que os algoritmos de árvores buscam as melhores partições que geram os nós mais puros. No entanto, os resultados gerados por árvores tendem a ser inferiores aos obtidos com os demais classificadores aqui apresentados (Sicsú et al., 2023).

O Quadro 1 apresenta exemplos de aplicação do algoritmo floresta aleatória para a identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos.

Quadro 1. Trabalhos científicos que utilizam o algoritmo floresta aleatória para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos

Título	Referências
Debris-Flow Susceptibility Assessment in China: A Comparison between Traditional Statistical and Machine Learning Methods	Huang et al. (2022)
Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China	Xiong et al. (2020)
Debris Flow Susceptibility Assessment Using the Integrated Random Forest Based Steady-State Infinite Slope Method: A Case Study in Changbai Mountain, China	Si et al. (2020)

Fonte: a autora (2026).

Enquanto Xiong et al. (2020) focaram na avaliação comparativa de diferentes algoritmos puros de *machine learning* para capturar as não-linearidades de fatores topográficos e ambientais em terrenos acidentados de bacias montanhosas, Si et al. (2020) utilizaram uma abordagem híbrida, integrando o algoritmo *Random Forest* ao modelo físico determinístico de encosta infinita, obtendo assim um modelo com compreensão da mecânica de ruptura do solo. Já Huang et al. (2022) realizaram uma análise comparativa entre os métodos estatísticos convencionais e as abordagens de *machine learning*. Os autores ainda refletem sobre a avaliação custo computacional e a capacidade de generalização requeridos pelas avaliações de *machine learning* para a aplicação.

2.4.2.2 Máquina de vetores de suporte

O *Support Vector Machine* (SVM), ou máquina de vetores de suporte, é um conhecido sistema de aprendizado de máquina não supervisionado. O conceito principal desse algoritmo é aplicar um modelo linear para realizar cálculos não lineares, estabelecendo limites de classe por meio do mapeamento não linear dos vetores de entrada para um novo espaço de alta dimensão. Nesse espaço, constrói-se o hiperplano de separação ideal, que visa separar as classes de saída na classificação. De forma mais detalhada, esse hiperplano ideal é o de margem

máxima, oferecendo a separação mais ampla entre as classes de saída e as amostras de treinamento mais próximas, conhecidas como vetores de suporte (Chang et al., 2019).

No Quadro 2 são apresentados trabalhos que utilizam máquina de vetores de suporte para a identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos.

Quadro 2. Autores que utilizaram máquina de vetores de suporte para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos

Título	Referências
Application of Machine Learning to Debris Flow Susceptibility Mapping along the China–Pakistan Karakoram Highway	Qing et al. (2020)
Debris Flow Scale Prediction Based on Correlation Analysis and Improved Support Vector Machine	Li et al. (2023)
A Research on Susceptibility Mapping of Multiple Geological Hazards in Yanzi River Basin, China	Gao et al. (2021)
Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China	Xiong et al. (2020)
Spatial Predictions of Debris Flow Susceptibility Mapping Using Convolutional Neural Networks in Jilin Province, China	Chen et al. (2020)
Comparison of Machine Learning and Traditional Statistical Methods in Debris Flow Susceptibility Assessment: A Case Study of Changping District, Beijing	Gu et al. (2023)

Fonte: a autora (2026).

Qing et al. (2020) e Gao et al. (2021) demonstraram a eficácia dos modelos de *machine learning*, especialmente máquina de vetores de suporte, no mapeamento da probabilidade de ocorrência de desastres ao longo de infraestruturas críticas e bacias hidrográficas complexas, mesmo com inventários de dados limitados. Gu et al. (2023) validaram o seu alto desempenho em zonas urbanas, resultando em modelos superiores aos modelos estatísticos clássicos. Li et al. (2023) propuseram a integração de análises de correlação a um modelo de máquina de vetores de suporte permitindo prever com precisão não somente a suscetibilidade de ocorrência, mas também a escala e o volume de material sólido dos fluxos de detritos.

2.4.2.3 MaxEnt

Embora o software MaxEnt delineado pelo Centro de Biodiversidade e Conservação do Museu Americano de História Natural, tenha sido inicialmente delineado para a modelagem de distribuições de espécies, aplicando a técnica de *learning machine* chamada modelagem de entropia máxima (Phillips, 2017), o programa tem sido utilizado para mapeamento de riscos naturais.

Segundo Phillips (2017), o programa calcula a probabilidade de ocorrência de um evento considerando os dados de entrada de caracterização da área e os registros de ocorrência

a partir da delimitação de áreas que possuem características compatíveis com as que já exibiram registros de ocorrência.

No mapeamento de riscos, o MaxEnt funciona determinando a distribuição de riscos encontrando a distribuição de maior grau de aleatoriedade dentro das restrições atribuídas definido pela relação entre circunstâncias de risco na entrada e suas covariáveis ambientais (Rusk, 2022).

Com relação ao uso do programa MaxEnt para o mapeamento de risco de desastres naturais, foram identificados 5 trabalhos sobre incêndios em áreas naturais, 5 sobre erosão, 12 sobre deslizamentos e 6 sobre inundações. No Quadro 3 são apresentados os trabalhos que abordam desastres relacionados a ocorrência de *huaicos*.

Quadro 3. Autores que utilizaram o programa MaxEnt para áreas de deslizamentos e inundações

Área de estudo	Título	Referências
Deslizamentos	Landslide susceptibility assessment using maximum entropy model with two different data sampling methods	Kornejady, Ownegh e Bahremand (2017)
	Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques	Chen et al. (2017)
	Spatial prediction of landslide susceptibility in Taleghan basin, Iran	Mokhtari e Abedian (2019)
	Application of the coupled TOPSIS–Mahalanobis distance for multi-hazard-based management of the target districts of the Golestan Province, Iran	Vahedberdi, Aiding e Majid (2019)
	Combining Participatory Mapping, Cloud Computing and Machine Learning for Mapping Climate Induced Landslide Susceptibility in Lembeh Island, North Sulawesi	Yusri et al. (2019)
	Submarine landslide; mapping the susceptibility in European seas	Innocenti et al. (2020)
	Evaluation of multi-hazard map produced using MaxEnt machine learning technique	Narges et al. (2021)
	Landslide susceptibility analysis of photovoltaic power stations in Gangwon-do, Republic of Korea	Kim e Park (2021)
	A comparative evaluation of machine learning algorithms and an improved optimal model for landslide susceptibility: a case study	Liu et al. (2021)
	Chinese High Resolution Satellite Data and GIS-Based Assessment of Landslide Susceptibility along Highway G30 in Guozigou Valley Using Logistic Regression and MaxEnt Model	Liu et al. (2022)
	A Comparative Assessment of Machine Learning Models for Landslide Susceptibility Mapping in the Rugged Terrain of Northern Pakistan	Shahzad, Ding e Abbas (2022)
	Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya	Rusk et al. (2022)
	Regional assessment of coastal landslide susceptibility in Liguria, Northern Italy, using MaxEnt	Orefice, Innocenti (2024)
	Effectiveness of landslide susceptibility mapping using the maximum entropy model and weights of evidence modelling in the Kuningan Regency, West Java, Indonesia	Suhermat et al. (2024)
	Landslide susceptibility prediction based on landform predisposing indexes – An example from the Beiluo River Basin	Liu et al. (2024)
	Geospatial technologies for landslide monitoring: a case study of Sighetu Marmăției, Romania	Kalmar et al. (2024)

Inundações	Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques	Darabi et al. (2019)
	Application of the coupled TOPSIS–Mahalanobis distance for multi-hazard-based management of the target districts of the Golestan Province, Iran	Vahedberdi, Aiding e Majid (2019)
	Urban flood hazard mapping using machine learning models: GARP, RF, MaxEnt and NB	Norollahi e Seyed Kaboli (2021)
	Evaluation of multi-hazard map produced using MaxEnt machine learning technique	Narges et al. (2021)
	Exploring the sample size and replications scenarios effect on spatial prediction of food, using MARS and MaxEnt methods case study: saliantape catchment, Golestan, Iran Ahmad	Rajabi et al. (2021)
	Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya	Rusk et al. (2022)
	Flood susceptibility assessment using machine learning approach in the Mohana-Khutiya River of Nepal	Maharjan et al. (2024)
	Urban flood risk analysis using presence-only machine learning approach: an integrated MaxEnt-cloud model framework in Harbin, China	Hu, Pang, Deng (2025)

Fonte: organizado pela autora (2026).

Dos trabalhos encontrados para avaliação de inundações, não há registros de sua aplicação para áreas de estudo de bacias de cursos hídricos efêmeros, deixando em aberto o uso deste programa para esse tipo de área.

2.4.2.4 Redes neurais artificiais

As redes neurais artificiais são modelos complexos inspirados no funcionamento do cérebro humano, formados por unidades básicas de processamento interligadas. Dessa forma, é possível analisar uma grande quantidade de dados e detectar padrões entre eles. Esses algoritmos possuem normalmente a camada de entrada, camadas intermediárias ou ocultas, e a camada de saída. Destaca-se que cada unidade básica de processamento recebe, como entrada, as saídas das unidades básicas de processamento anteriores. O funcionamento da rede depende do treinamento desta, do qual decorre a quantificação dos pesos que permitem prever a resposta final com o menor erro possível (Sicsú et al., 2023).

No Quadro 4, são apresentados trabalhos que utilizam algoritmos de redes neurais artificiais na identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos.

Pei e Qiu (2023) realizam uma abordagem composta de modelos baseados na física, como o modelo de encosta infinita, eficazes na identificação de zonas de início dos fluxos de detritos e um modelo de redes neurais artificiais, utilizados para identificar áreas de propagação dos fluxos de detritos. Chen et al. (2020) utilizam redes neurais convolucionais bidimensionais (2D-CNN) com técnicas de *SAME padding* (S-CNN) e *VALID padding* (V-CNN), comparadas aos modelos de Máquina de Vetores de Suporte (SVM) e Redes Neurais Artificiais (ANN)

convencionais. Ullah et al. (2022), identificou na área de estudo no Paquistão, três diferentes perigos geológicos, além dos fluxos de detritos, os autores elaboraram o mapa de suscetibilidade de ocorrência de deslizamentos e inundações. Hien et al. (2022) validaram uma abordagem híbrida utilizando algoritmos de *machine learning*, dentre os quais redes neurais artificiais para apresentar a suscetibilidade de formação dos fluxos de detritos e posteriormente integrando os resultados ao modelo cinemático FlowR para calcular a rota e a propagação do material.

Quadro 4. Autores que utilizaram o algoritmo redes neurais artificiais para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos

Título	Referências
Spatial Predictions of Debris Flow Susceptibility Mapping Using Convolutional Neural Networks in Jilin Province, China	Chen et al. (2020)
Debris flow susceptibility mapping in Colorado Front Range, USA: A comparison of physics-based and data-driven approaches	Pei e Qiu (2023)
Multi-hazard susceptibility mapping based on Convolutional Neural Networks	Ullah et al. (2022)
Large-scale Mapping of Landslide and Debris Flow using Flowr Model with Statistical and Machine Learning Methods	Hien et al. (2022)

Fonte: organizado pela a autora (2026).

2.4.2.5 Regressão logística

A regressão logística estima a probabilidade de uma observação pertencer a uma categoria predeterminada. Uma das técnicas utilizadas para estimativa dos parâmetros é a máxima verossimilhança (Sicsú et al., 2023). Trata-se de um método popular de modelagem estatística amplamente aplicado em diversos problemas, como a seleção de genes na classificação do câncer e na análise de crimes. Na suscetibilidade a deslizamentos de terra, a análise e a regressão logística também têm sido utilizadas popularmente em muitas áreas de ocorrências desses eventos. Na regressão logística, o principal conceito matemático é usar o logaritmo natural de uma razão de chances.

Embora a regressão logística seja um algoritmo paramétrico de aprendizado de máquina, ou seja, o modelo de aprendizado que resume dados com um conjunto de parâmetros de tamanho fixo, não importando quantos dados são inseridos em um modelo paramétrico, o modelo não mudará a essência do resultado. Tanto o algoritmo máquina de vetores de suporte, quanto a floresta aleatória, são modelos não paramétricos, onde os algoritmos não fazem suposições fortes sobre a forma da função de mapeamento.

No Quadro 5 constam exemplos de aplicação direta do algoritmo regressão logística para a identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos.

Quadro 5. Artigos que utilizam o algoritmo regressão logística para identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos

Título	Referências
Debris-Flow Susceptibility Assessment in China: A Comparison between Traditional Statistical and Machine Learning Methods	Huang et al. (2022)
Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China	Xiong et al. (2020)
Comparison of Machine Learning and Traditional Statistical Methods in Debris Flow Susceptibility Assessment: A Case Study of Changping District, Beijing	Gu et al. (2023)
Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models	Ahmad et al. (2021)

Fonte: organizado pela autora (2026).

Ahmad et al. (2021) demonstraram a eficácia da aplicação de modelos estatísticos e de *machine learning* para mapear múltiplos riscos simultâneos, em áreas de característica física complexa. Gu et al. (2023) identificaram que algoritmos como a Máquina de Vetores de Suporte (SVM), são superiores as abordagens estatísticas clássicas, pois são altamente adaptáveis frente a uma baixa existência de dados de localização dos fenômenos e uma alta variabilidade espacial dos fatores condicionantes. Xiong et al. (2020) destacam que em terrenos montanhosos instáveis, a precisão dos modelos de *machine learning* depende de como cada modelo processa as não-linearidades das variáveis ambientais.

2.4.3 Estado da arte: Abordagens

As abordagens utilizadas na literatura quando do uso de algoritmos para identificação de suscetibilidade de ocorrência de fenômenos naturais, usualmente avaliam o comportamento de mais de um algoritmo, são os casos de Liu et al. (2021) que para deslizamentos fizeram o uso dos algoritmos Maxent, máquina de vetores de suporte, floresta aleatória e redes neurais artificiais. Para o fluxo de detritos, Huang et al. (2022) utilizaram regressão logística e floresta aleatória, com área de estudo a China.

Outros autores ainda avaliam mais de um tipo de fenômeno, a exemplo de Rusk et al. (2022), que consideraram inundações, deslizamentos e incêndios, e Chauhan, Gupta e Dixit (2025) que realizaram uma análise que integrou os perigos de sismos, deslizamentos e inundações. Poucos trabalhos avaliam a suscetibilidade de ocorrência de *deslizamentos*, como Gu et al. (2023), que utilizam regressão logística, Xiong et al. (2020), que fazem uso de floresta aleatória, Gao et al. (2021), com o algoritmo de vetores de suporte, e Ullah et al. (2022), com

redes neurais artificiais. No entanto, não foram identificados trabalhos que utilizem o algoritmo Maxent no fluxo de detritos. Além disso, na literatura, utiliza-se apenas um conjunto de variáveis ambientais.

Diante do exposto, considerando a reduzida quantidade de referências, optou-se pela avaliação de dois conjuntos distintos de variáveis: um segundo a literatura geral e outro considerando os fatores desencadeantes de huaicos do Instituto Geológico, Minero y Metalúrgico (INGEMMET) do Peru, que possui bibliografia específica sobre *huaicos*. Por outro lado, os estudos encontrados na literatura são elaborados com base em um único modelo para cada perigo ou algoritmo. Neste trabalho, devido às condições especiais do regime pluviométrico do Peru, com meses de precipitação quase nula e à influência do El Niño na ocorrência de huaicos, elegeu-se a elaboração de modelos mensais, para cada um dos algoritmos apontados, para a avaliação da mudança da resposta às diferentes precipitações. Para além disso, destaca-se que foi escolhida como área de estudo o departamento de Lima devido à alta incidência de ocorrência de huaicos na região e a dificuldade de elaboração do mapeamento para toda a região costeira do Peru, ocasionada pelo tamanho dos conjuntos de dados e o tempo que demandado para geração dos resultados.

2.5 VARIÁVEIS AMBIENTAIS

Nos trabalhos científicos que abordam a suscetibilidade de ocorrência de desastres naturais utilizando os algoritmos Floresta Aleatória, Máquina de Vetores de Suporte, MaxEnt, Redes Neurais Artificiais e Regressão Logística, é possível identificar variáveis ambientais que normalmente são consideradas relevantes para diversos fenômenos, como fluxo de detritos, deslizamentos, erosão e inundações.

Os autores Kornejady, Ownegh e Bahremand (2017); Chen et al.(2017); Mokhtari e Abedian (2019); Sheikh, Kornejady e Ownegh (2019); Yusri et al. (2019); Innocenti et al. (2020); Javidan et al. (2021); Kim e Park(2021); Liu et al. (2021); Liu et al. (2022); Shahzad, Ding e Abbas (2022); Rusk et al. (2022) utilizaram a altitude, aspecto, declividade, distância a partir de cursos hídricos permanentes, distância a partir de estradas, geologia, perfil de curvatura, plano de curvatura, precipitação, uso e cobertura do solo, índice de umidade topográfica (TWI) e índice de rugosidade do terreno (TRI) para a determinação de áreas com suscetibilidade de ocorrência de deslizamentos pelo MaxEnt.

Ao abordar inundações, Darabi et al. (2019); Sheikh, Kornejady e Ownegh (2019); Norallahi e Seyed Kaboli (2021); Javidan et al. (2021); Rajabi et al. (2021); Rusk et al. (2022)

consideram a altitude, aspecto, declividade, distância a partir de cursos hídricos perenes, precipitação, uso e cobertura do solo, como os demais trabalhos, mas utilizam ainda um coeficiente de escoamento.

Os trabalhos que tratam de identificação de áreas com suscetibilidade de ocorrência de fluxo de detritos, utilizam basicamente as mesmas variáveis citadas anteriormente. Independentemente do algoritmo utilizado, seja ele regressão logística, floresta aleatória, máquina de vetores de suporte ou rede neural artificial. De forma geral, as variáveis relacionadas ao relevo, ao uso e cobertura do solo, a precipitação e a formação da área, já citadas anteriormente foram utilizadas por Chen et al. (2020), Qing et al. (2020), Quin et al. (2020), Si et al. (2020), Xiong et al. (2020), Ahmad et al. (2021), Gao et al. (2021), Hien et al. (2022), Huang et al. (2022), Ullah et al. (2022), Gu et al. (2023), Li et al. (2023), Pei e Qiu (2023).

Kornejady, Ownegh e Bahremand (2017) tratam a altitude como a característica mais direta do relevo. Em geral, as características topográficas contribuem para a formação de ambientes propensos a desastres. Esses fatores topográficos se desdobram além da altitude em si, em declive, aspecto e índices compostos como o índice de umidade topográfica (TWI), que quantifica o controle da topografia sobre os processos hidrológicos e é um índice relacionado a deslizamentos de terra (2022) e o índice de rugosidade do terreno (TRI) é, outro índice associado a deslizamentos de terra.

Sheikh, Kornejady e Ownegh (2019) utilizaram a precipitação média anual para avaliar o risco de deslizamentos de terra e inundações. Arabameri et al. (2020) também utilizaram a precipitação média anual para avaliar o risco de erosão. Para a área deste estudo, entende-se que o uso da precipitação mensal, em vez da precipitação anual, é mais válido, dada a conhecida distinção entre meses chuvosos (especialmente em anos de El Niño) e meses secos.

3 MATERIAIS E MÉTODOS

O desenvolvimento do estudo seguiu o fluxograma apresentado na Figura 12 e descrito nos itens sequenciais do capítulo.

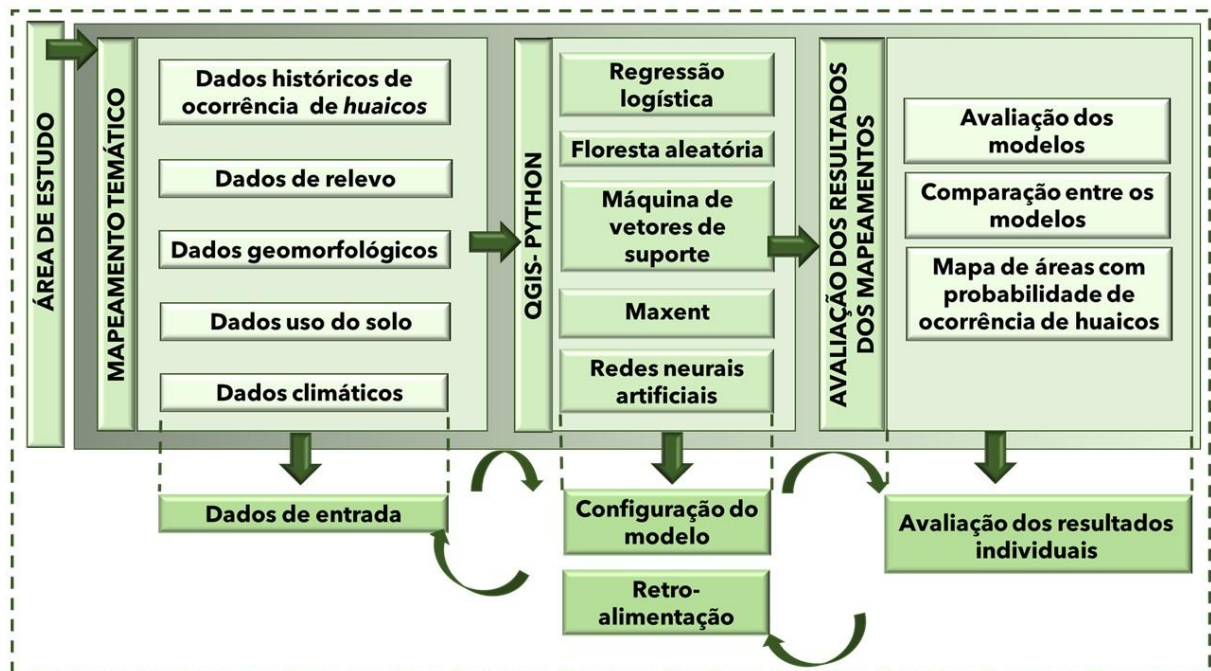


Figura 12. Fluxograma do desenvolvimento do estudo. Fonte: a autora (2026).

3.1 ÁREA DE ESTUDO

A área de estudo deste trabalho é o departamento de Lima, no qual se insere a capital do Peru, Lima, na coordenada central de latitude $12^{\circ} 3'14.45''\text{S}$ e longitude $77^{\circ} 2'38.86''\text{O}$, conforme apresentado na Figura 13. O departamento de Lima possui uma área de 35.011,04 km². Geopoliticamente, a área abrange 10 províncias: Barranca, Catajambo, Cañete, Canta, Huaral, Huarochirí, Huaura, Oyón, Yauyos e Lima, onde está localizada a capital do país. Nesta área, foram registrados 1.650 huaicos entre 2003 e 2021 (Figura 13) no inventário de perigos geológicos no Peru (INGEMMET, 2023).

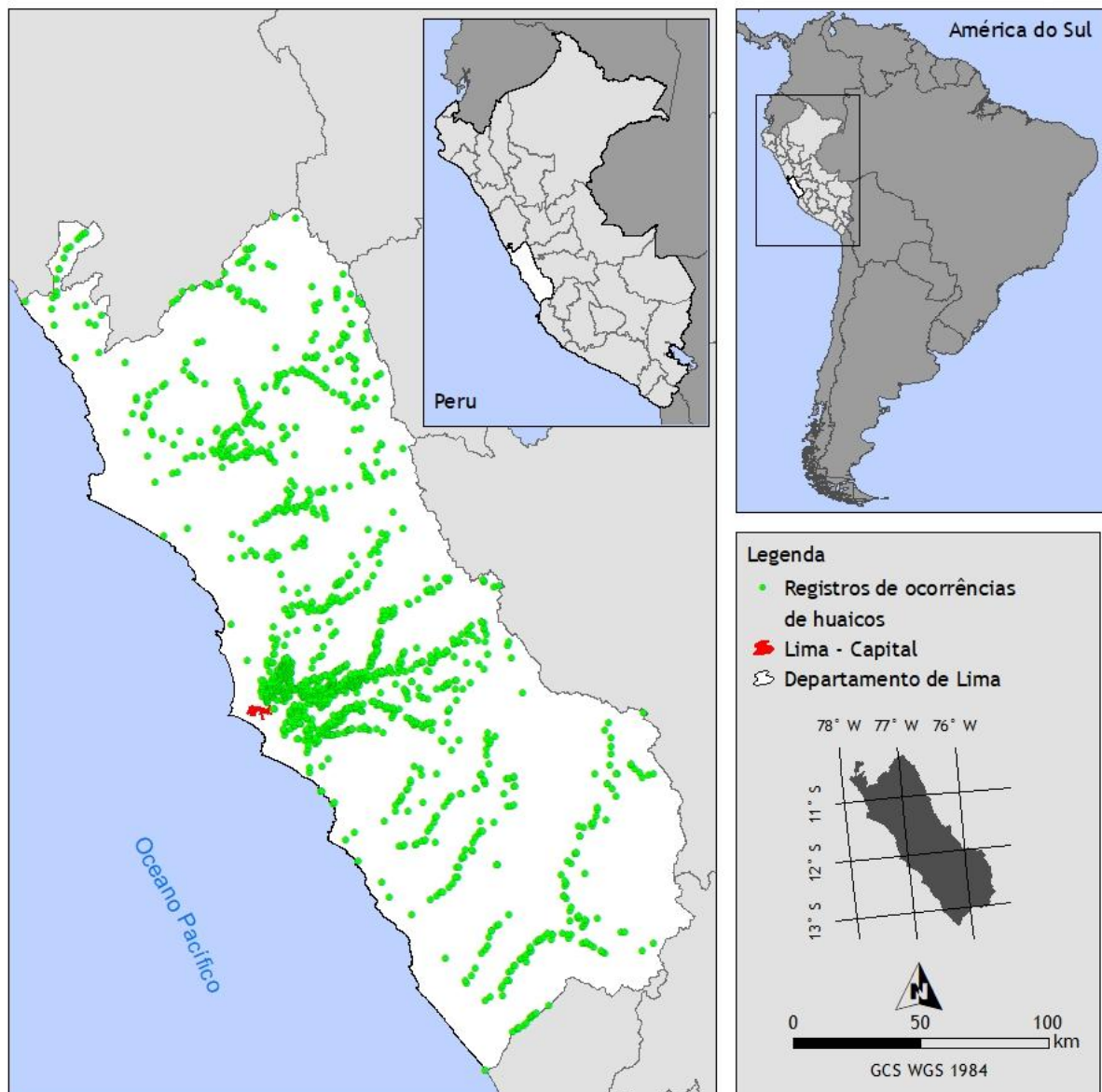


Figura 13. Inventário das ocorrências de *huaicos* no departamento de Lima. Fonte: INGEMMET (2023).

3.2 MAPEAMENTO TEMÁTICO

O mapeamento temático elaborado nesta etapa objetivou compreender a dinâmica local das bacias de cursos hídricos efêmeros, além de servir de dados de entrada aos algoritmos avaliados.

3.2.1 Dados históricos de ocorrência de *huaicos*

O Instituto Geológico, Minero-Metalúrgico (INGEMMET) possui, em seu catálogo de metadados, registros de perigos geológicos no Peru, incluindo inventários de quedas,

deslizamentos, fluxos, movimentos complexos, tombamentos e outros. Este inventário identifica os *huaicos* como fluxos de detritos (INGEMMET, 2023). O arquivo de perigos geológicos foi recortado no QGIS para abranger o departamento de Lima. Em seguida, utilizando o DATUM WGS 1984, as coordenadas geográficas das ocorrências foram extraídas e convertidas para o formato CSV (valores separados por vírgula) para serem inseridas como entrada de dados no MaxEnt. Para os demais algoritmos, utilizou-se o arquivo recortado inicialmente.

Ressalta-se que 30% dos registros foram selecionados e retirados dos arquivos de registros de ocorrências para uso posterior no cálculo do índice Kappa. Assim sendo, 70% dos dados alimentaram os algoritmos (Arabameri et al., 2020; Javidan et al., 2021).

3.2.2 Variáveis ambientais

A caracterização de áreas com ocorrência de *huaicos* no Peru, realizada pelo INGEMMET, destacou como fatores determinantes para o desencadeamento desses desastres a precipitação em áreas onde se formam cursos d'água efêmeros, a própria formação geológica e geomorfológica, além do relevo da região, tanto quanto à conformação altitudinal do terreno, quanto à sua declividade (INGEMMET, 2005; 2021). Espera-se que essas características sejam as mais proeminentes nos modelos de identificação de áreas com risco de ocorrência de *huaicos*. No entanto, outras variáveis ambientais encontradas na literatura já citada também foram adicionadas à modelagem, conforme descrito nos itens a seguir.

Também é essencial destacar a necessidade de avaliar a correlação entre as variáveis ambientais antes de aplicá-las aos modelos. Essa avaliação foi realizada no ambiente QGIS considerando as correlações de Pearson (Ullah, et al., 2022) e Spearman (Li et al., 2023).

3.2.2.1 Modelo digital de elevação

Para este trabalho, foi utilizado o modelo digital de elevação (MDE) Shuttle Radar Topography Mission (STRM), com resolução de 30 metros por pixel, datado de 2000, obtido do United States Geological Survey (USGS). Embora os dados do satélite Alos Palsar, que também fornece informações de altimetria, apresentem resolução espacial de 12,5 m, dada a grande extensão da área de estudo, a capacidade computacional para o processamento de dados mais detalhados seria extrapolada.

O modelo digital de elevação, recortado para a área de estudo no QGIS, serviu como dado de entrada para a modelagem. A partir disso, também foram gerados dados de declividade, em graus, no QGIS, utilizando a ferramenta de cálculo de declividade, com o modelo digital de elevação como dado de entrada. Além deste, foi utilizado o aspecto que mede a orientação de cada célula em relação ao norte, também criado no QGIS com o uso da ferramenta aspecto. O índice de rugosidade do terreno (TRI) também possui uma ferramenta que gera diretamente esta informação no QGIS. O índice de umidade topográfica (TWI), que indica a tendência de o terreno acumular água, foi gerado a partir da calculadora raster, calculando-se o logaritmo natural da razão entre a área do declive drenada em cada pixel e a colina, tendo como dados de entrada a altitude e a declividade.

3.2.2.2 Geologia, geomorfologia e pedologia

A literatura frequentemente descreve que a formação geológica da área influencia notavelmente a estabilidade da formação rochosa ao avaliar o risco de deslizamentos e erosão. No Peru, o INGEMMET disponibiliza dados geológicos na escala de 1:50.000, a qual foi atualizada em 2022.

Embora a formação geomorfológica esteja relacionada à ocorrência de *huaicos*, dados geomorfológicos e hidrogeológicos, não se observou o uso destes na literatura. Dados geomorfológicos na escala de 1:250.000, datados de 2016, e dados hidrogeológicos na escala de 1:2.000.000, também de 2016, foram obtidos do INGEMMET

A tipologia do solo foi obtida do Ministério do Meio Ambiente (MINAM) em escala 1:2.000.000, em 2009. Propriedades do solo, como resistência ao cisalhamento e permeabilidade, estão relacionadas à propensão ao deslizamento e ao movimento da água na porção superficial (2022).

3.2.2.3 Dados de uso e cobertura do solo

Foram elaborados mapas com elementos relacionados ao uso e cobertura do solo da área de estudo, contendo informações tanto naturais como antrópicas da região de estudo, a saber: distâncias a partir de áreas urbanizadas, distâncias a partir de cursos de água e classificação de ecossistemas presentes do Peru. Os cursos hídricos foram obtidos na Plataforma Nacional de Dados Georreferenciados Geo Peru, na escala 1:100.000, elaborada pelo *National Geographic Institute* (NGI). As áreas urbanizadas foram obtidas no Instituto

Geográfico Nacional, em escala 1:100.000, com data de 2024.

A distância linear de cada célula em relação ao ponto de interesse, neste caso, as áreas de interesse são as zonas urbanizadas e os cursos hídricos, foi gerada no QGIS. Embora os *huaicos* ocorram a partir da geração de fluxos superficiais efêmeros e em ambientes naturais, as distâncias a partir áreas urbanizadas e hidrografia buscam avaliar uma possível relação indireta entre esses fatores.

As informações de ecossistemas do Peru disponibilizadas pelo Ministério de Meio Ambiente do Peru (MINAM) (2019) foram usadas na escala de 1:2.200.000. Embora, à primeira vista, não estejam diretamente relacionadas com a ocorrência de *huaicos*, o uso dessas informações pode nos permitir avaliar a existência de uma correlação com a ocorrência de *huaico*.

3.2.2.4 Dados climáticos

O SENAMHI possui o cálculo das normais climatológicas para 182 estações do país com dados entre 1991-2020, só no departamento de Lima são 52 estações. São dados de temperatura máxima e mínima e precipitação média. O emprego de médias dos mínimos de precipitação não repercutiria em resultados positivos de relação com a ocorrência de *huaicos*, pois mesmo os valores médios das máximas no país apresentam meses de precipitação nula (Santos, 2021; Santos et al., 2025). Entretanto, Santos et al. (2025) utilizaram a média dos níveis máximos de precipitação para o departamento de Lima para alimentar o algoritmo *MaxEnt* na identificação de suscetibilidade de ocorrência de *huaicos*, também não retornando uma relação preponderante entre a precipitação e a ocorrência de *huaicos*. Constatou-se no artigo que, devido às características áridas da região, não são necessários altos níveis de precipitação para desencadear *huaicos*, assim, valores médios de precipitação podem ser mais adequados que valores máximos de precipitação para a identificação de relação com os *huaicos*.

Posteriormente, os dados de precipitação foram distribuídos espacialmente por interpolação (Kriging) em um ambiente SIG, gerando, para cada mês, um conjunto de polígonos de cobertura de precipitação para cada estação. Para inserção dos dados nos modelos, todos os dados de entrada, com exceção das ocorrências, precisam possuir formato matricial, com as mesmas características de número de colunas e linhas, resolução espacial e sistema de projeção espacial.

3.3 ALGORITMOS SELECIONADOS

Os algoritmos descritos nos itens a seguir foram selecionados a partir do capítulo de revisão bibliográfica, item 2.5. Foram utilizados os algoritmos de regressão logística, floresta aleatória, máquina de vetores de suporte, redes neurais artificiais e o MaxEnt, sendo o MaxEnt o de maior uso na avaliação de áreas com suscetibilidade de ocorrência de desastres naturais.

Com exceção do algoritmo MaxEnt, que possui execução direta pelo Java®, os demais algoritmos podem ser executados a partir do plugin QGIS Scikit-Learn, delineado para o QGIS. Sendo assim apresenta-se na Figura 14 o fluxograma de criação de modelos com os algoritmos regressão logística, floresta aleatória, máquina de vetores de suporte e redes neurais artificiais.

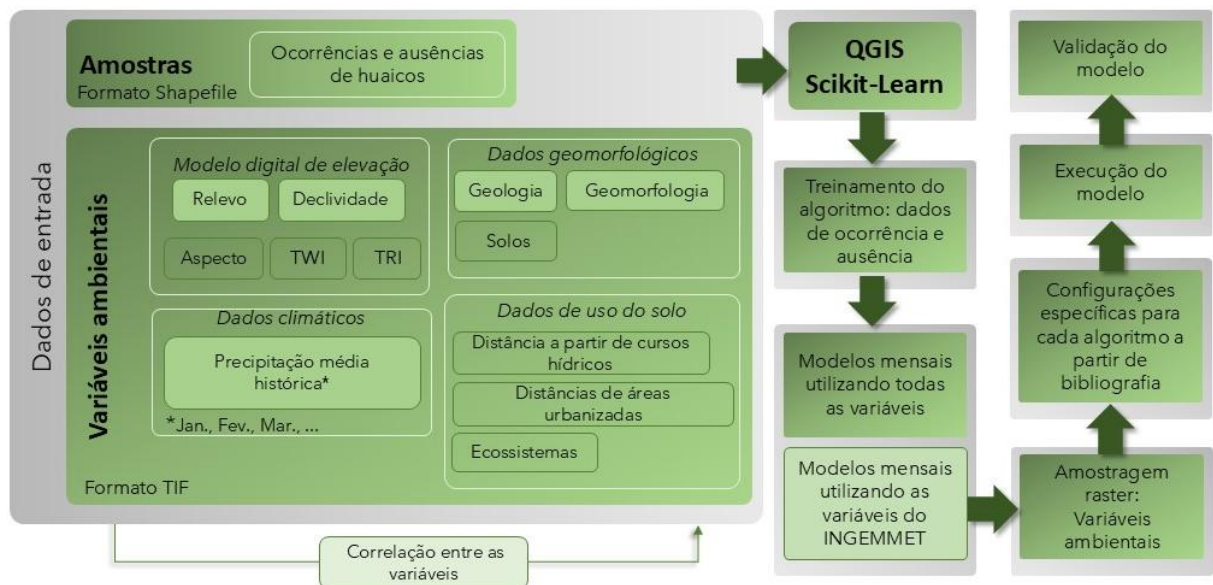


Figura 14. Fluxograma de elaboração dos modelos a partir dos algoritmos regressão logística, floresta aleatória, máquina de vetores de suporte e redes neurais artificiais no QGIS

Para cada um dos algoritmos foi necessário acrescentar pontos de treinamento de ausência do fenômeno. Além disso, foi necessário extrair valores das variáveis preditoras desses pontos com a ferramenta "Amostragem de Raster" (QGIS, 2025).

Por meio do plugin QGIS Scikit-Learn foi importado o arquivo de treinamento do algoritmo e anexou-se as variáveis preditoras, validando-se a condição de presença e ausência de *huaicos*. Na sequência, configurou-se os hiper parâmetros específicos de cada algoritmo com base na bibliografia e teste no algoritmo (QGIS, 2025).

Após a execução do treinamento do modelo, avaliou-se a acurácia desse, usando como métricas a AUC-ROC (área abaixo da curva) e índice Kappa (QGIS, 2025).

3.3.1 MaxEnt

Para inserir os dados no MaxEnt, todas as variáveis ambientais foram transformadas em um formato de matriz (Geotiff), garantindo que todos os dados tivessem o mesmo tamanho, número de pixels e posicionamento. O processo de compatibilidade dos dados ocorreu no software QGIS, onde os dados inicialmente em formato vetorial, foram transformados em dados matriciais. Os dados matriciais passaram por reprojeção do sistema de coordenadas e Datum (WGS UTM) e foram recortados para exibir apenas os departamentos de interesse.

Dos registros de ocorrência de *huaicos* no departamento de Lima, 70% foram selecionados aleatoriamente, reservando os 30% restantes para validação do modelo, de acordo com a metodologia proposta por Arabameri et al. (2020) e Javidan et al. (2021).

O MaxEnt tem uma interface amigável, simples para inserir dados e configurar. As amostras (70% dos registros), que são os dados sobre a ocorrência de *huaicos* no departamento de Lima, foram inseridas no formato CSV.

Após a inserção das variáveis ambientais no formato ASCII, cada variável foi considerada um dado contínuo ou categórico. As informações restantes correspondem à configuração do modelo, realizada de acordo com as diretrizes estabelecidas por Dalapicolla (2016), destacando a adoção de uma porcentagem de teste aleatória de 30%, a execução de réplicas do modelo bootstrap (15) e um número máximo de 2.000 iterações.

O processo de inserção de dados no MaxEnt seguiu os passos mostrados na Figura 15, sendo inicialmente organizado um conjunto de modelos mensais, ou seja, um modelo para cada mês, considerando todas as variáveis ambientais (altitude, declividade, aspecto, TRI, TWI, geologia, geomorfologia, solos, distância a partir de cursos hídricos, distância a partir de áreas urbanizadas, ecossistemas e precipitação média mensal) e outro conjunto de modelos mensais (um modelo para cada mês) apenas para as variáveis descritas pelo INGEMMET (altitude, declividade, geologia, geomorfologia e precipitação média mensal) como condutoras no processo de formação dos *huaicos*.

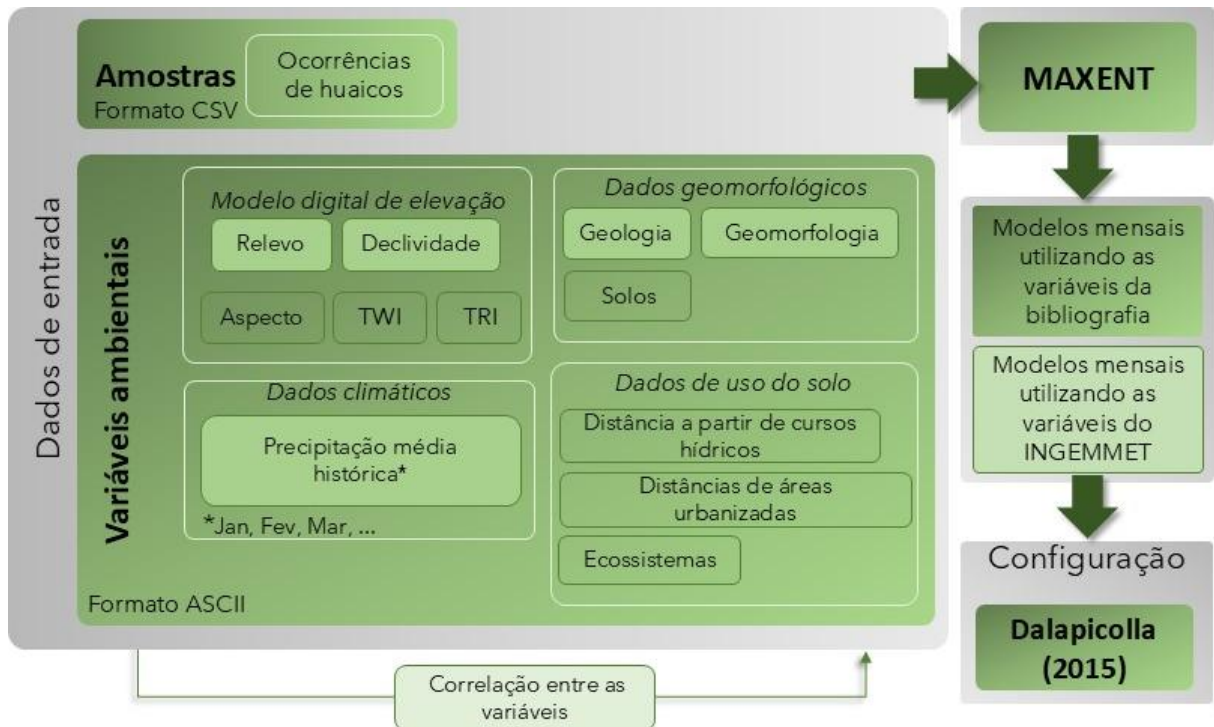


Figura 15. Fluxograma do processo de modelagem no MaxEnt

O modelo de probabilidade contínua (de 0 a 1) foi reclassificado com base no limiar logístico de presença de treinamento de 10 percentis. Após, será avaliada a sobreposição de cada área com suscetibilidade de ocorrência de cada modelo mensal sobre os outros meses.

Vale ressaltar que o algoritmo MaxEnt funciona apenas com a presença do evento, não sendo necessário apontar áreas onde o evento está ausente. Entretanto, é essencial observar que o modelo precisa considerar alguns pixels como uma ausência prática para calcular as estatísticas do modelo (Chen et al., 2017; Norollahi; Kaboli, 2021).

MaxEnt é descrito como dependente de uma resposta de aprendizado de máquina que faz previsões a partir de dados incompletos, resultando em uma saída MaxEnt no formato ASCII, como uma previsão contínua de presença específica variando de 0 a 1 (Javidan et al. 2021; Chen et al., (2017).

3.4 CONFIABILIDADE DOS RESULTADOS

Adotou-se o índice Kappa (Landis; Koch, 1977) com método para verificação da concordância dos resultados. A partir do uso de 30% dos registros de ocorrências de *huaicos*, selecionados aleatoriamente e não utilizados nos modelos, calculou-se o índice Kappa de cada algoritmo, por meio do cruzamento dessas informações no QGIS. O índice Kappa pode

chegar a 1, sendo uma concordância menor que “0”, pobre, até 0,2 leve, até 0,4 razoável, até 0,6 moderada, até 0,8 substancial e maior que 0,8, quase perfeita. O índice Kappa foi calculado no ambiente SIG de geração de mapas de suscetibilidade do QGIS.

Além desse, a própria qualidade do modelo será avaliada pelo uso do indicador AUC-ROC. A avaliação AUC (*Area Under the Curve*), da curva ROC (*Receiver Operating Characteristic*), avalia, de maneira robusta o desempenho de modelos de predição presença/ausência. Esta não depende de um limiar de decisão específico, ou seja, não depende de uma definição de linha de corte. A AUC-ROC, indica a probabilidade de o modelo classificar corretamente uma presença aleatória como mais provável do que uma ausência aleatória. Assim, um valor de AUC de 0,5 indica desempenho aleatório do modelo, enquanto uma AUC de 1,0 indica um modelo sem falhas. No entanto, valores acima de 0,7 já são considerados aceitáveis (Fielding; Bell, 1997). A estatística AUC-ROC é calculada automaticamente para o MaxEnt e é gerada no ambiente SIG de geração de mapas de probabilidade dos algoritmos Floresta Aleatória, Máquina de Vetores de Suporte, Redes Neurais Artificiais e Regressão Logística.

4 RESULTADOS E DISCUSSÕES

Neste capítulo, são apresentados os resultados deste trabalho, inicialmente para as variáveis ambientais que alimentam os algoritmos, descrevendo as características da região. Na sequência, seguem os resultados para os algoritmos Floresta Aleatória, Máquina de Vetores de Suporte, Maxent, Redes Neurais Artificiais e Regressão Logística, sendo apresentados em itens separados. Os itens são compostos por mapas, gráficos e tabelas, juntamente com os textos descritivos.

4.1 VARIÁVEIS AMBIENTAIS

Nos itens a seguir são descritas separadamente as variáveis ambientais utilizadas para caracterizar a área de estudo e alimentar os algoritmos, com o objetivo de avaliar a suscetibilidade de ocorrência de *huaicos* no departamento de Lima. As variáveis são variáveis associadas ao relevo: altitude (forma do relevo), declividade, aspecto, TRI (*Terrain Ruggedness Index*) e TWI (*Topographic Wetness Index*); variáveis associadas a formação do terreno: geologia, geomorfologia, pedologia; variáveis associadas ao uso e cobertura do solo:

ecossistemas, distância a partir de cursos hídricos e distância a partir de áreas urbanizadas; e variáveis climáticas: precipitação média mensal histórica.

4.1.1 Altitude

Em relação às características físicas da área de estudo (Figura 16), a altitude, segundo o modelo digital de elevação, varia de 0 a 6.248 metros. Identifica-se uma altitude baixa na porção que se encontra o Oceano Pacífico, sendo que aumenta à medida que adentra o continente. Observa-se que as ocorrências de *huaicos* não ocorrem nas porções de altitudes máximas e próximas a essas. Na Tabela 6 constam os quantitativos relacionados a altitude.

Cruzando as informações sobre ocorrências de huaicos e altitude, 57% dos registros ocorreram até 1.000 metros de altitude. Considerando os registros até 2.000 metros de altitude, 80% dos huaicos estão nessa faixa.

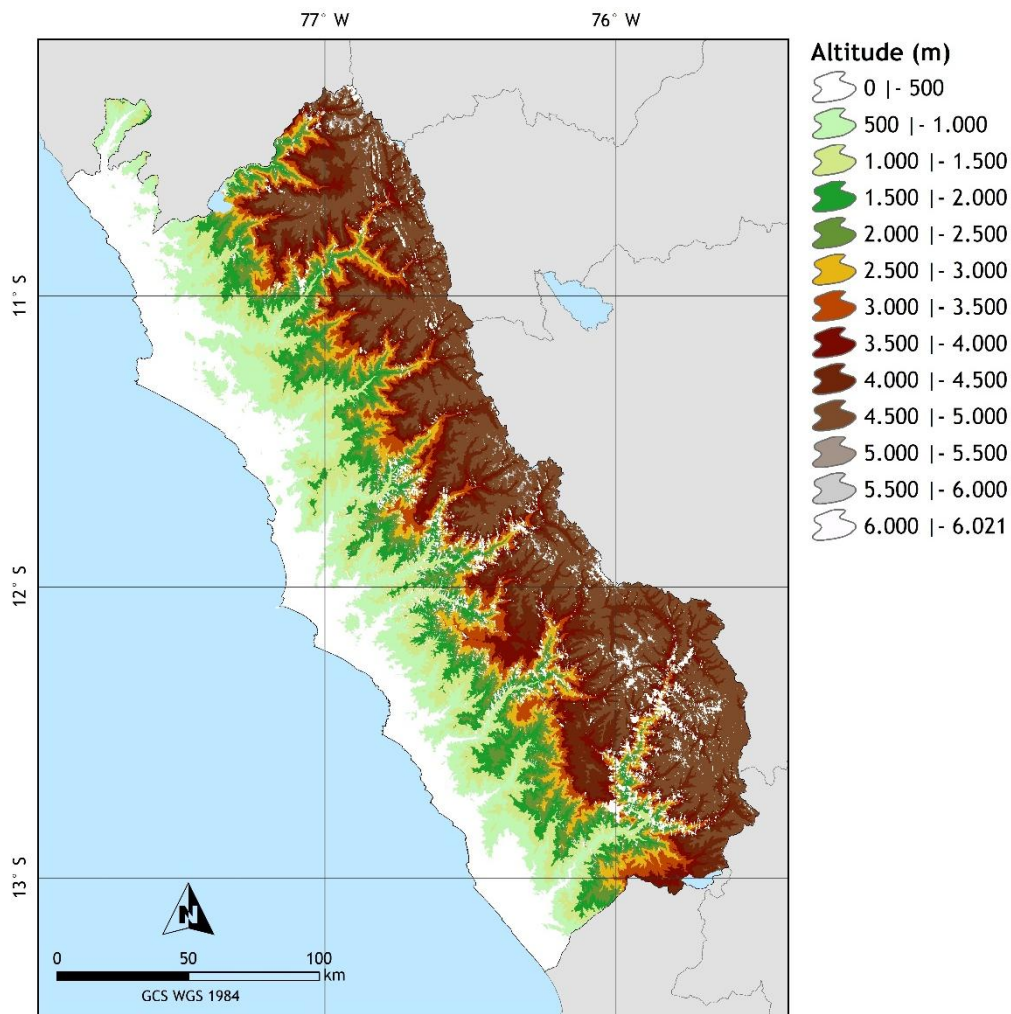


Figura 16. Altitude. Fonte: adaptado de SRTM (2000).

Tabela 6. Quantificação da altitude

Altitude (m)	Área (km²)	%	Altitude (m)	Área (km²)	%
0 - 500	7.734,73	22,09	3.500 - 4.000	2.245,31	6,41
500 - 1.000	3.495,72	9,98	4.000 - 4.500	4.143,31	11,83
1.000 - 1.500	2.797,18	7,99	4.500 - 5.000	6.153,04	17,57
1.500 - 2.000	2.378,92	6,79	5.000 - 5.500	191,99	0,55
2.000 - 2.500	2.074,90	5,93	5.500 - 6.000	1,39	0,00
2.500 - 3.000	1.823,34	5,21	6.000 - 6.248	0,00	0,00
3.000 - 3.500	1.971,21	5,63			
Total				35.011,04	100

Fonte: a autora (2026).

4.1.2 Declividade

A declividade mapeada na Figura 17 é um fator influente na ocorrência de huaicos. O dado gerado no ambiente do QGIS, com base no Modelo Digital de Elevação (MDE) proveniente do SRTM (*Shuttle Radar Topography Mission*), permitiu a representação das características topográficas da área de estudo. Esta varia de 0° à 86° de inclinação, que em porcentagem correspondente à 192% de inclinação. Ainda assim, pela Tabela 7 observa-se que 98,41% da área de estudo possui declividades de até 45°.

Os *huaicos* ocorrem especialmente em áreas com inclinação de até 20°, concentrando 75% das ocorrências nesse intervalo. Se consideradas as declividades de até 35°, somam-se 96% das ocorrências de *huaicos* nesse intervalo. Isso indica que as ocorrências foram registradas nas porções de acúmulo do fluxo. Entende-se que a geração do fluxo ocorre em áreas de maior inclinação, no entanto, seria necessária a avaliação de cada microbacia isoladamente para compreender o gradiente de declividade relacionado a cada ocorrência de huaicos.

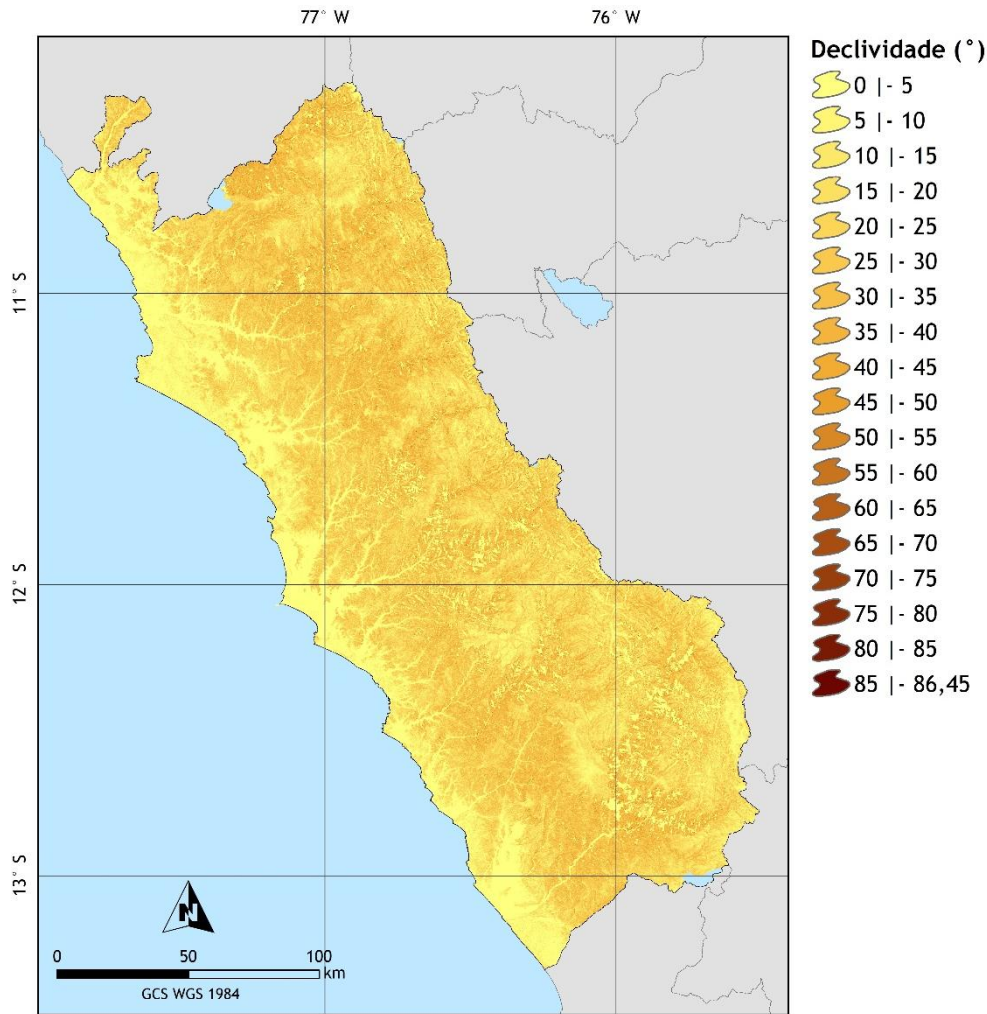


Figura 17. Declividade. Fonte: a autora (2026).

Tabela 7. Quantificação da declividade

Declividade (°)	Área (km²)	%	Declividade (°)	Área (km²)	%
0 - 5	5.712,98	16,32	45 - 50	366,90	1,05
5 - 10	3.233,74	9,24	50 - 55	123,01	0,35
10 - 15	3.604,91	10,30	55 - 60	43,87	0,13
15 - 20	4.249,28	12,14	60 - 65	15,49	0,04
20 - 25	4.799,06	13,71	65 - 70	4,88	0,01
25 - 30	4.853,62	13,86	70 - 75	1,06	0,00
30 - 35	4.193,08	11,98	75 - 80	0,14	0,00
35 - 40	2.689,89	7,68	80 - 85	0,00	0,00
40 - 45	1.119,12	3,20	85 - 86,45	0,00	0,00
Total				35.011,04	100,00

Fonte: a autora (2026).

4.1.3 Aspecto

O aspecto representa a orientação do relevo, em ângulo a partir do norte, variando de 0° a 360° (Figura 18). Isso tem relação especialmente a insolação, que influencia na temperatura e umidade de uma área. Na área de estudo, o aspecto varia de 0° a 360°, sendo difícil identificar

um padrão específico, como pode ser observado pela distribuição dos quantitativos de ângulos do aspecto na Tabela 8, onde cada faixa de 30° englobam entre 5,18% da área de estudo até 11,76% dessa. Mas ainda assim, observa-se uma certa orientação geral do terreno costeiro em 300°.

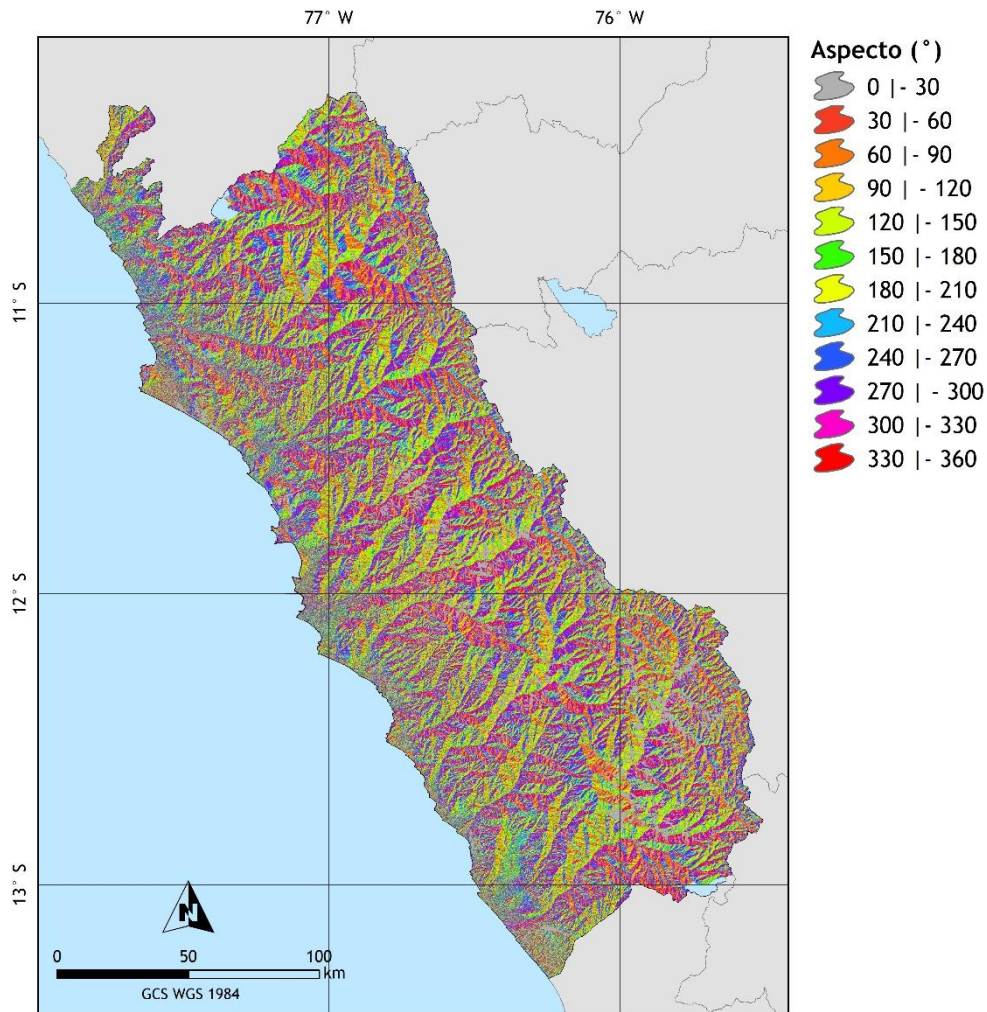


Figura 18. Aspecto. Fonte: a autora (2026).

Tabela 8. Quantificação do aspecto

Aspecto (°)	Área (km²)	%	Aspecto (°)	Área (km²)	%
0 - 30	4.116,67	11,76	180 - 210	3.386,61	9,67
30 - 60	2.000,60	5,71	210 - 240	3.362,47	9,60
60 - 90	1.812,24	5,18	240 - 270	3.096,96	8,85
90 - 120	2.612,22	7,46	270 - 300	3.776,36	10,79
120 - 150	2.776,60	7,93	300 - 330	3.048,67	8,71
150 - 180	2.587,86	7,39	330 - 360	2.433,80	6,95
Total				35.011,04	100,00

Fonte: a autora (2026).

4.1.4 Terrain Ruggedness Index - TRI

O índice TRI (*Terrain Ruggedness Index*) avalia o grau de acidentamento do relevo e o classifica em nivelado, aproximadamente nivelado, levemente acidentado, intermediariamente acidentado, moderadamente acidentado, altamente acidentado e extremamente acidentado (Riley; DeGLoria; Elliot, 1999). Na área de estudo, o terreno classifica-se majoritariamente como nivelado, correspondendo a 95,12% da área, conforme a Tabela 9 e a Figura 19. Ainda assim, podem ser observadas áreas com terreno classificado como altamente acidentado, ao longo de toda a área de estudo.

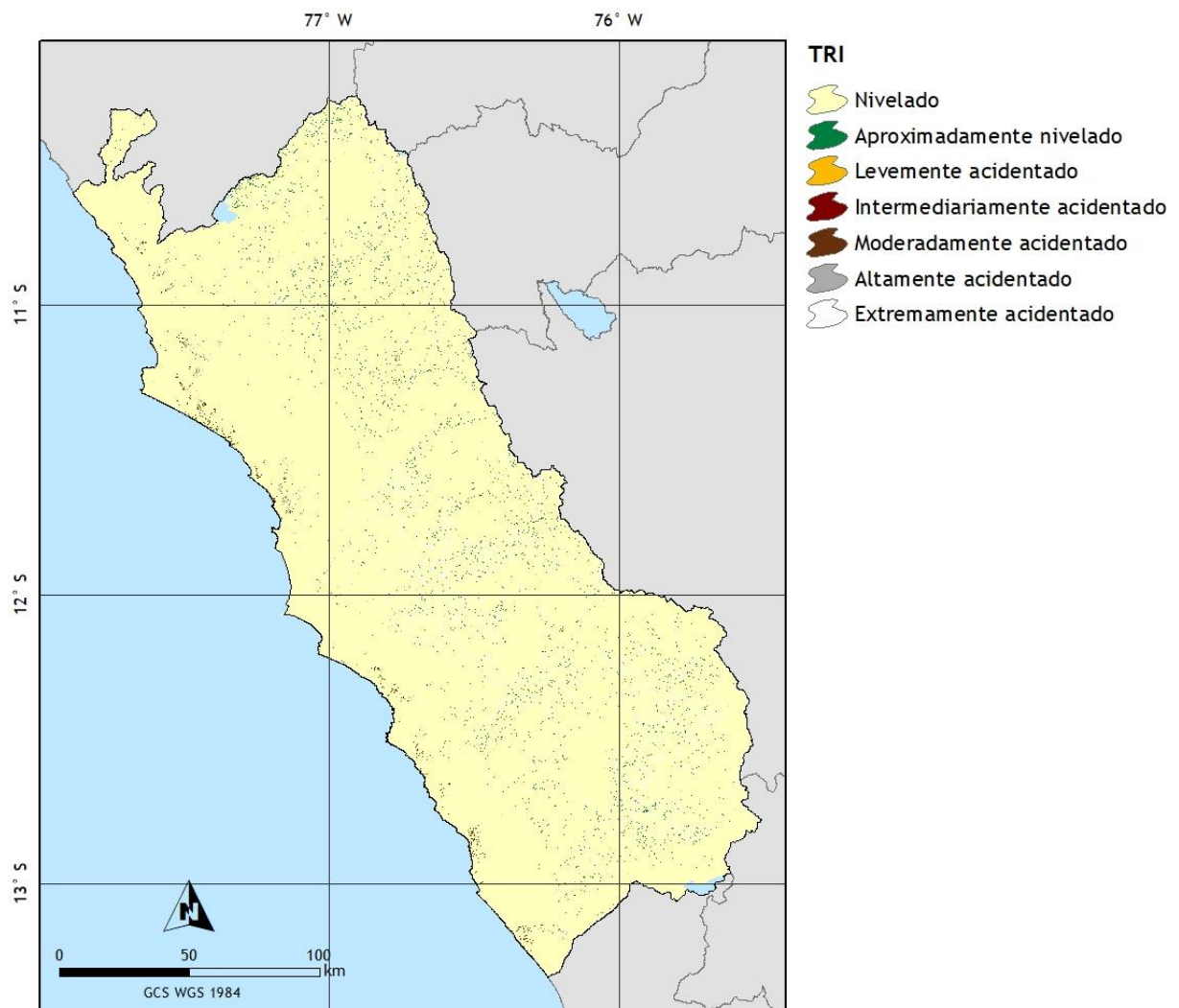


Figura 19. TRI – Terrain ruggednes index. Fonte: a autora (2026).

Tabela 9. Quantificação do TRI

Classificação do TRI	Área (km²)	%
Nivelado	33.300,93	95,12
Aproximadamente nivelado	401,40	1,15
Levemente acidentado	56,88	0,16
Intermediariamente acidentado	25,42	0,07
Moderadamente acidentado	43,97	0,13
Altamente acidentado	37,83	0,11
Extremamente acidentado	1.144,60	3,27
Total	35.011,04	100,00

Fonte: a autora (2026).

4.1.5 Topographic Wetness Index - TWI

O TWI (*Topographic Wetness Index*) indica onde a água tende a se acumular. Na área de estudo, o TWI variou de 0 a 27 (Figura 20), com os maiores valores nas áreas mais baixas e porções menos íngremes do terreno, indicando assim essas áreas como as mais suscetíveis à deposição de água. Ainda assim, conforme os quantitativos da Tabela 10, a área de estudo é composta principalmente por área com baixa deposição.

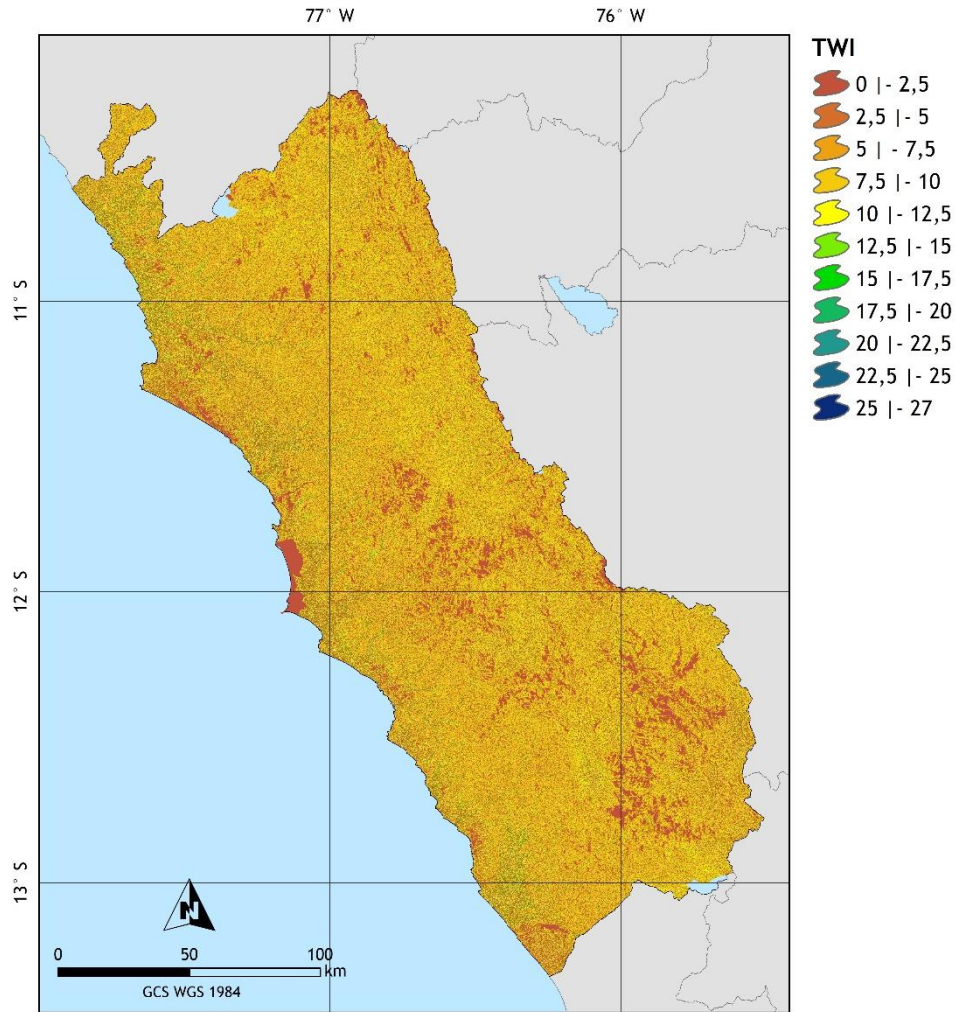


Figura 20. TWI – Topographic wetness index. Fonte: a autora (2026).

Tabela 10. Quantificação do TWI

TWI	Área (km ²)	%
0 - 2,5	11.598,20	33,13
2,5 - 5,0	0,93	0,00
5,0 - 7,5	4.637,20	13,24
7,5 - 10	11.151,33	31,85
10,0 - 12,5	4.999,64	14,28
12,5 - 15,0	1.735,61	4,96
15,0 - 17,5	686,95	1,96
17,5 - 20,0	182,06	0,52
20,0 - 22,5	18,71	0,05
22,5 - 25,0	0,41	0,00
25,0 - 27,0	0,00	0,00
Total	35.011,04	100

Fonte: a autora (2026).

4.1.6 Geologia

Dos dados geológicos mapeados na Figura 21, disponibilizados pelo INGEMMET (2019), observam-se 38 classes geológicas na área de estudo. Dessas, 31 estão relacionadas às

ocorrências de huaicos; 16% das ocorrências estão relacionadas à classificação Superunidade Santa Rosa; 14% relacionadas à classificação Cretáceo Superior e Inferior vulcânico-sedimentar; e mais 12% das ocorrências correspondentes à classificação Paleogêneo-Neogêneo, vulcânico-sedimentar. No Quadro 6, são discriminadas todas as classificações geológicas identificadas na área de estudo, bem como a área de abrangência de cada uma. Destaca-se que além de apresentarem as maiores correspondências entre as ocorrências de huaicos, as classificações geológicas supracitadas também possuem as maiores áreas de quantificação dentro da área de estudo.

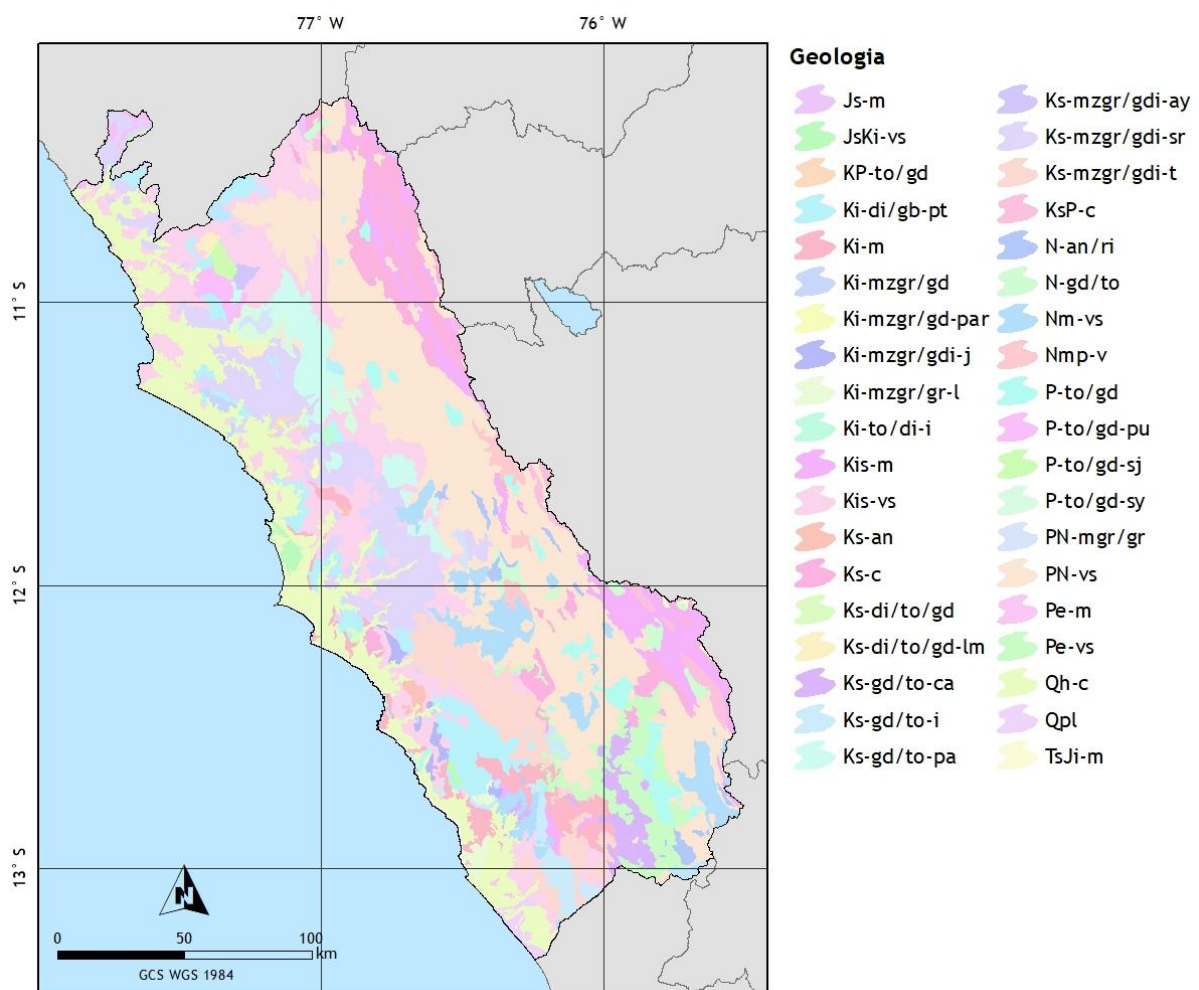


Figura 21. Classificação geológica. Fonte: INGEMMET (2022).

Quadro 6. Caracterização geológica

Sigla	Classificação	Área (km²)	%
Js-m	Jurássico superior marinho	22,80	0,07
JsKi-vs	Jurássico superior - Cretáceo inferior, vulcânico-sedimentar	128,14	0,37
Ki-di/gb-pt	Super unidade, Patap	1.382,33	3,95
Ki-m	Cretáceo inferior, marinho	892,40	2,55
Ki-mzgr/gd	Cretáceo inferior, monzogranito/granodiorita	65,25	0,19
Ki-mzgr/gd-par	Cretáceo inferior, microgranito/granodiorita	18,37	0,05
Ki-mzgr/gdi-j	Cretáceo inferior, microgranito/granodiorita	155,69	0,44
Ki-mzgr/gr-l	Super unidade, Linga	128,89	0,37
Ki-to/di-i	Super unidade, Ilo	32,11	0,09
Kis-m	Cretáceo inferior superior, marinho	1.689,17	4,82
Kis-vs	Cretáceo superior inferior, vulcânico-sedimentar	4.355,93	12,44
KP-to/gd	Cretáceo Paleogênico, tonalita/granodiorita	32,38	0,09
Ks-na	Cretáceo superior, subvolcânico Lurin	81,03	0,23
Ks-c	Cretáceo superior continental	1.789,69	5,11
Ks-di/to/gd	Super unidade, La Mina, Humaya	18,53	0,05
Ks-di/to/gd-lm	Super unidade, La Mina	104,19	0,30
Ks-gd/to-ca	Cretáceo, granodiorita, tonalita	511,96	1,46
Ks-gd/to-i	Cretáceo superior, granodiorita/tonalita	414,87	1,18
Ks-gd/to-pa	Super unidade, Pampahuasi	1.055,95	3,02
Ks-mzgr/gdi-ay	Super unidade, Aynaca	59,41	0,17
Ks-mzgr/gdi-sr	Super unidades, Santa Rosa	2.938,36	8,39
Ks-mzgr/gdi-t	Super unidades, Tiabaya	1.572,42	4,49
KsP-c	Cretáceo superior, Paleogênico, continental	362,42	1,04
N-an/ri	Neogeno, andesito-riolito	221,45	0,63
N-gd/to	Neogeno, granodiorita-tonalita	280,84	0,80
Nm-vs	Neogeno Mioceno, vulcânico-sedimentar	1.296,68	3,70
Nmp-v	Neogeno Mioceno Plioceno, vulcânico	246,98	0,71
P-to/gd	Tonalitas e granodioritas paleogênicas	676,98	1,93
P-to/gd-pu	Plutão Puscao	268,33	0,77
P-to/gd-sj	Plutão, San Jeronimo	99,81	0,29
P-to/gd-sy	Paleozoico, tonalita/granodiorita	268,25	0,77
Pe-m	Paleogeno Eoceno, marinho	10,97	0,03
Pe-vs	Paleogeno Eoceno, vulcânico-sedimentar	1.128,07	3,22
PN-mgr/gr	Stocks da região Andina	127,34	0,36
PN-vs	Paleogeno-Neogeno, vulcânico-sedimentar	8.505,90	24,29
Qh-c	Quaternário Holoceno-continental	3.935,29	11,24
Qpl	Quaternário Pleistoceno continental	110,00	0,31
TsJi-m	Triássico superior, Jurássico inferior, marinho	21,88	0,06

Fonte: INGEMMET (2022).

4.1.7 Geomorfologia

São identificadas na área de estudo, conforme a Figura 22, 45 classes geomorfológicas, segundo o INGEMMET (2016). As ocorrências de *huaicos* cruzadas com a geomorfologia evidenciam a relação dessas com 24 classificações. Dessas, destaca-se a classe geomorfológica de vertente ou sopé aluvio-torrencial, com 29% das ocorrências de *huaicos*, 20% em montanhas de rochas intrusivas e 14% relacionadas à classificação geomorfológica de montanha em rocha vulcânica. Somadas, essas classes representam 63% das ocorrências. A

classe de vertente ou sopé aluvio-torrencial é uma planície inclinada que se estende ao pé dos contrafortes ou dos sistemas montanhosos andinos. São o resultado do acúmulo de materiais mobilizados pelos huaicos (INGEMMET, 2021; 2023).

A classe de montanha em rocha intrusiva, pode apresentar fraturamento dos afloramentos, gerando assim material granular que alimenta os fluxos de detritos (INGEMMET, 2024). As montanhas em rochas vulcânicas, também são suscetíveis a geração de deslizamentos e quedas de rochas (INGEMMET, 2020).

No Quadro 7 são explicitadas as classificações geomorfológicas da área de estudo e suas quantificações. As classes de montanhas de rochas intrusivas e montanha em rocha vulcânica ocupam respectivamente 23,92% e 26,04%.

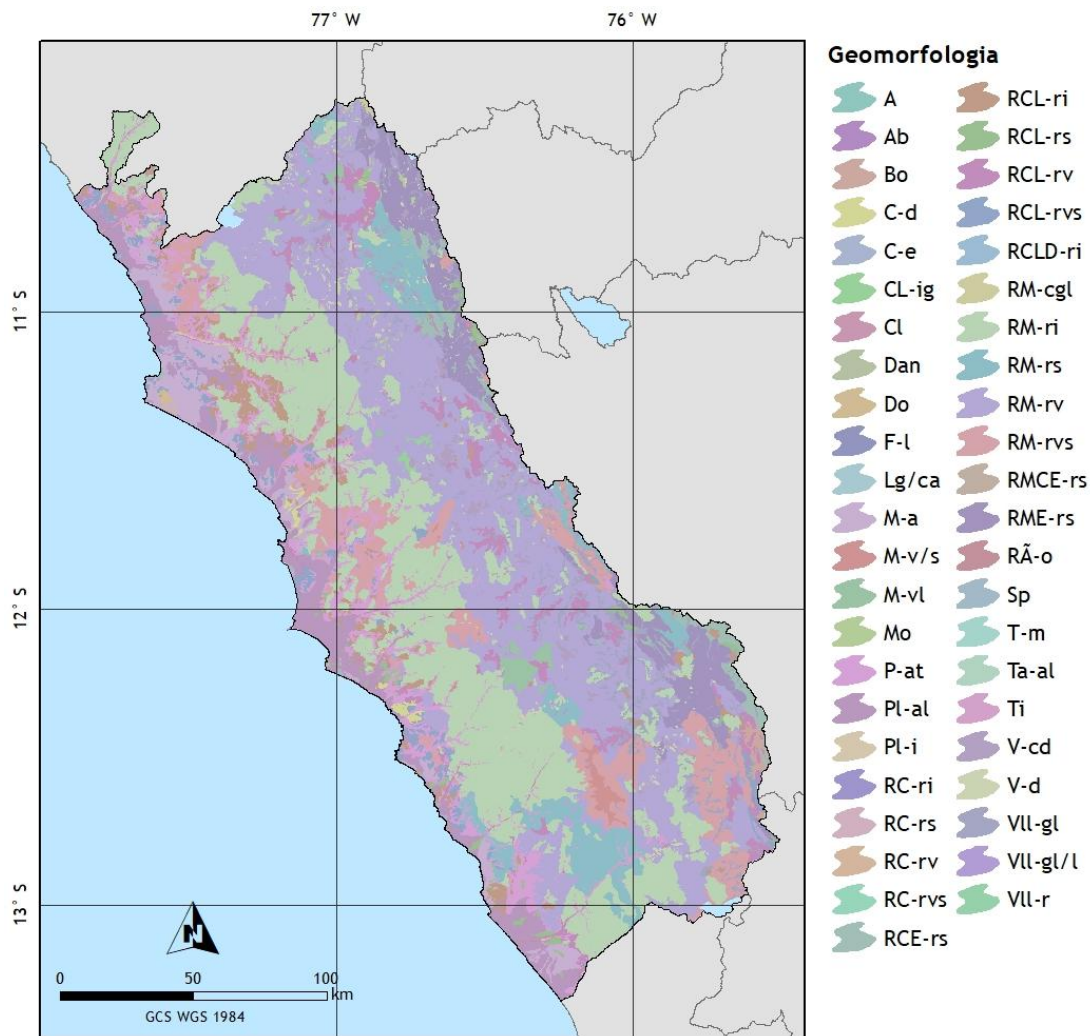


Figura 22. Classificação geomorfológica. Fonte: INGEMMET (2016).

Quadro 7. Caracterização geomorfológica

Sigla	Classificação geomorfológica	Área (km²)	%
A	Albufeira	7,39	0,02
Ab	Leque aluvial de sopé	0,38	0,00
Bo	Bofedales (pântano alto-andino)	86,82	0,25
C-d	Campo de dunas	90,94	0,26
C-e	Costa emergente ou de regressão recente	9,31	0,03
Cl	Cordão litorâneo	4,88	0,01
CL-ig	Colina e lomba ignimbríticas	39,77	0,11
Dan	Depósito antrópico	8,16	0,02
Do	Domo de sal	23,96	0,07
F-l	Faixa litorânea	17,89	0,05
Lg/ca	Lagoa e corpos d'água	7,47	0,02
M-a	Mantos de areia	1.135,73	3,24
M-v/s	Planalto vulcânico/sedimentar	137,45	0,39
M-vl	Planalto vulcânico lávico	234,60	0,67
Mo	Morenas	109,71	0,31
P-at	Vertente ou sopé aluvio-torrencial	2.081,54	5,95
Pl-al	Planície aluvial	1.834,14	5,24
Pl-i	Planície inundável	30,66	0,09
Río	Leito do rio	69,50	0,20
RC-ri	Colina em rocha intrusiva	0,11	0,00
RC-rs	Colina em rocha sedimentar	0,29	0,00
RC-rv	Colina em rocha vulcânica	10,06	0,03
RC-rvs	Colina em rocha vulcano-sedimentar	11,13	0,03
RCE-rs	Colina estrutural em rocha sedimentar	447,21	1,28
RCL-ri	Colina e lomba em rocha intrusiva	944,76	2,70
RCL-rs	Colina e lomba em rocha sedimentar	224,74	0,64
RCL-rv	Colina e lomba em rocha vulcânica	1.175,80	3,36
RCL-rvs	Colina e lomba em rocha vulcano-sedimentar	638,21	1,82
RCLD-ri	Colina e lomba dissecadas em rocha intrusiva	4,51	0,01
RM-cgl	Montanha com cobertura glacial	21,45	0,06
RM-ri	Montanha em rocha intrusiva	8.374,21	23,92
RM-rs	Montanha em rocha sedimentar	1.761,46	5,03
RM-rv	Montanha em rocha vulcânica	9.117,63	26,04
RM-rvs	Montanha em rocha vulcano-sedimentar	2.429,48	6,94
RMCE-rs	Montanhas e colinas estruturais em rocha sedimentar	5,06	0,01
RME-rs	Montanha estrutural em rocha sedimentar	1.823,55	5,21
Sp	Sistema de pântanos e aguajás	15,55	0,04
T-m	Terraço marinho	15,38	0,04
Ta-al	Terraço alto aluvial	12,17	0,03
Ti	Terraço indiferenciado	152,60	0,44
V-cd	Vertente ou sopé colúvio-deluvial	543,39	1,55
V-d	Vertente coluvial de detritos	258,98	0,74
Vll-gl	Vale glacial	1.091,29	3,12
Vll-gl/l	Vale glacial com lagoa	0,14	0,00
Vll-r	Vale represado	1,59	0,00

Fonte: INGEMMET (2016).

4.1.8 Pedologia

Na área de estudo há 7 classificações pedológicas que podem ser observadas na Figura 23. A identificação da classificação, tipo de paisagem e declividade da área e o quantitativo de cada classe podem ser observados no Quadro 8. As tipologias de solo de Leptosolo distrófico

- Afloramento lítico, Leptossolo lítico - Afloramento lítico e Regossolo dístico - Afloramento lítico, possuem as maiores áreas dentro da área de estudo. Ou seja, são áreas de leptossolos e regossolos, que apresentam solos pouco desenvolvidos, com forte presença do material de origem (Embrapa, 2018).

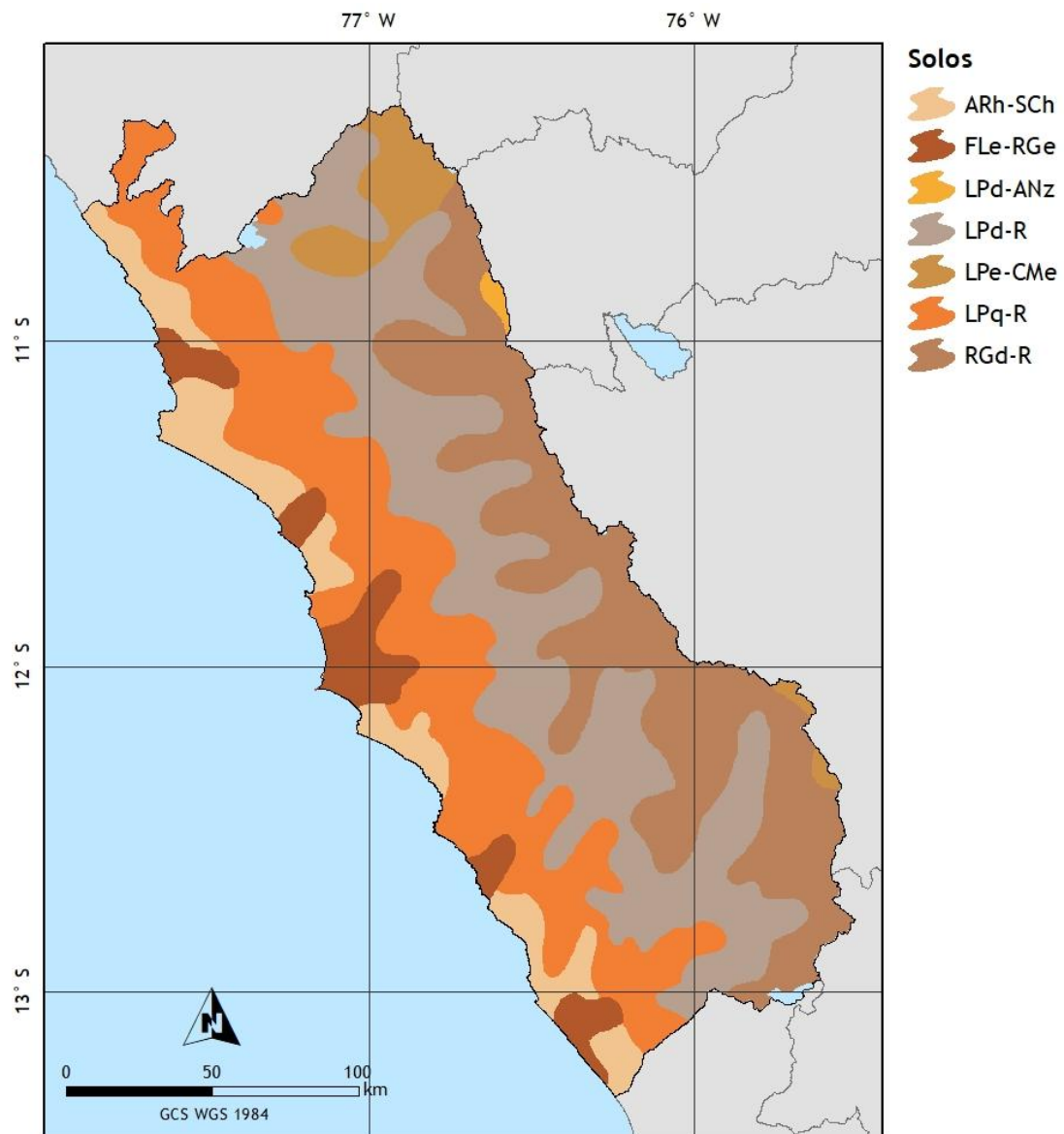


Figura 23. Classificação pedológica. Fonte: MINAM (2009).

Quadro 8. Caracterização pedológica

Sigla	Classificação	Paisagem	Declividade (%)	Área (km²)	%
ARh-SCh	Arenossolo háplico - Solonchaco háplico	Planícies marítimas e eólicas	0-8 e 8-25	2.958,06	8,45
FLe-Gle	Fluviossolo eutrófico - Gleissolo eutrófico	Terraços aluviais baixos, geralmente inundáveis	0-8	1.904,29	5,44
LPd-Anz	Leptossolo distrófico - Andossolo vítrico	Colinas e montanhas, em menor proporção morros	25 a +75, 8-25	96,35	0,28
LPd-R	Leptossolo distrófico - Afloramento lítico	Encostas da vertente ocidental da cordilheira andina	25 a +75	10.914,71	31,18
LPe-Cme	Leptossolo eutrófico - Cambissolo eutrófico	Morros, colinas e montanhas	8-25, 25 a +75	1.402,33	4,01
LPq-R	Leptossolo lítico - Afloramento lítico	Encostas da vertente ocidental da cordilheira andina	25 a +75	9.112,00	26,03
RGd-R		Montanhas da cadeia ocidental dos Andes	25 a +75	8.623,30	24,63
Total				35.011,04	100,00

Fonte: MINAM (2009).

4.1.9 Ecossistemas

O mapa de ecossistemas da área de estudo pode ser observado na Figura 24, sendo contabilizados 3 diferentes ecossistemas (Tabela 11). Segundo o memorial do MINAM (2019), o ecossistema de Puna, que possui a maior área de cobertura na área de estudo, é um ecossistema andino com vegetação herbácea, constituída principalmente por gramíneas de baixo crescimento e pastagens de gramíneas, que crescem dispersas e apresentam caules e folhas duras, e por algumas associações de arbustos dispersos. O ecossistema de Serrania Esteparia apresenta características semelhantes às do ecossistema de Puna. Já a Selva Alta (Yunga), possui uma paisagem fisiográfica dominada pelo sistema montanhoso com encostas íngremes, onde o clima é caracterizado por alta umidade, com áreas de neblina permanente, as florestas são densas, com a presença marcante de epífitas (bromélias, orquídeas), samambaias e algumas palmeiras (MINAM, 2019).

O deserto costeiro é um ecossistema que varia de árido a hiperárido com áreas em sua maioria desprovidas de vegetação. São constituídos por solos arenosos ou por afloramentos rochosos que ocupam áreas planas, onduladas e dissecadas, sujeitas à erosão eólica. Estende-se desde as praias e falésias até as primeiras elevações das encostas, podendo ocupar áreas significativas. Algumas formações vegetais destacáveis são os cactos (colunares, prostrados e globulares), as moitas e os arbustos (MINAM, 2019).

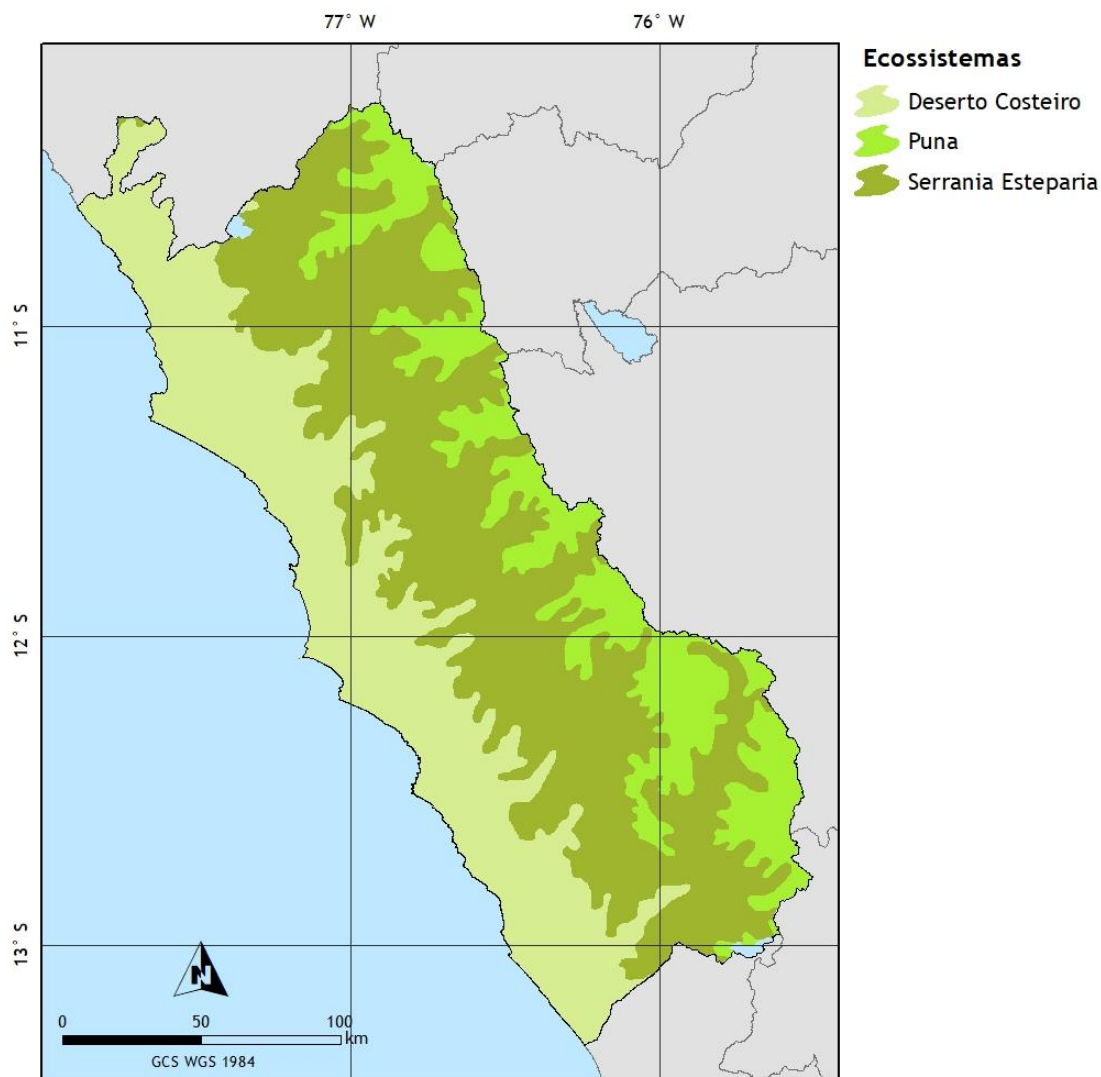


Figura 24. Ecosistemas do Peru. Fonte: MINAM (2019).

Tabela 11. Quantificação dos ecossistemas

Ecosistemas	Área (km ²)	%
<i>Puna</i>	6.467,54	18,47
<i>Desierto Costanero del Pacifico Peruano</i>	12.001,40	34,28
<i>Serrania Esteparia</i>	16.542,10	47,25
Total	35.011,04	100

Fonte: a autora: (2026).

4.1.10 Distância a partir de cursos hídricos

A distância a partir dos cursos hídricos (perenes) variou, ao longo da área de estudo, de 0 metros a 32.591 metros, conforme a Figura 25. Na maior parte das áreas, observam-se distâncias de até 10.000 metros a partir dos cursos hídricos.

Na Tabela 12, que contém os quantitativos de distâncias a partir dos cursos hídricos perenes, 35,01% das áreas de estudo estão a menos de 2.500 metros deles.

A proximidade de cursos hídricos pode estar mais ou menos associada a *huaicos*, considerando a formação do relevo da região. Embora sejam poucos os cursos hídricos perenes identificados, nessas bacias constam também bacias de cursos hídricos efêmeros, delineadas naturalmente ao longo do tempo. No entanto, destaca-se que a presença de cursos hídricos perenes não é um fator que desencadeia diretamente um *huaico*.

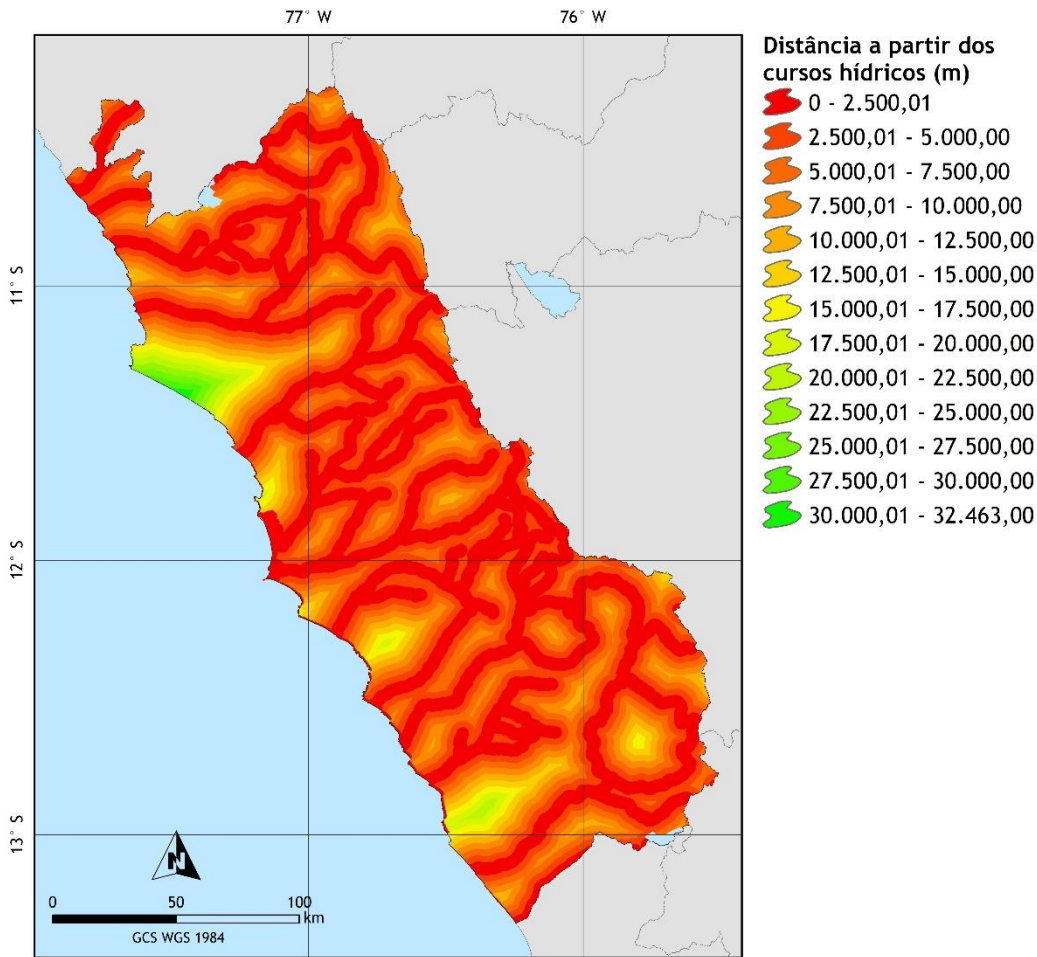


Figura 25. Distância a partir dos cursos hídricos. Fonte: adaptado de NGI (2024).

Tabela 12. Quantificação das distâncias a partir dos cursos hídricos

Distância a partir dos cursos hídricos (m)	Área (km²)	%	Distância a partir dos cursos hídricos (m)	Área (km²)	%
0 - 2.500	12.258,17	35,01	17.500 - 20.000	334,41	0,96
2.500 - 5.000	9.511,28	27,17	20.000 - 25.500	184,38	0,53
5.000 - 7.500	6.280,44	17,94	22.500 - 25.000	119,98	0,34
7.500 - 10.000	3.421,07	9,77	25.000 - 27.500	89,33	0,26
10.000 - 12.500	1.454,28	4,15	27.500 - 30.000	58,32	0,17
12.500 - 15.000	767,03	2,19	30.000 - 32.591	18,53	0,05
15.000 - 17.500	513,84	1,47			0
Total				35.011,04	100,00

Fonte: a autora (2026).

4.1.11 Distância a partir de áreas urbanizadas

A distância a partir das áreas urbanizadas do mapa da Figura 26 variou de 0 a 24.578 metros. Em sua maioria, as distâncias a partir dos cursos hídricos variaram cerca de 12.500 metros. Pela Tabela 13 constata-se que há uma distribuição espacial entre as áreas urbanizadas. Nas porções norte e leste, e em duas áreas, estão contabilizadas as maiores distâncias entre as áreas. Destaca-se que essa informação reflete a distância de cada pixel com relação ao ponto de interesse, neste caso, áreas urbanizadas.

Com relação à proximidade das áreas urbanizadas, essas são associadas ao impacto dos *huaicos*, mas também podem estar associadas a localização dessas áreas em áreas sujeitas a ocorrência de *huaicos*.

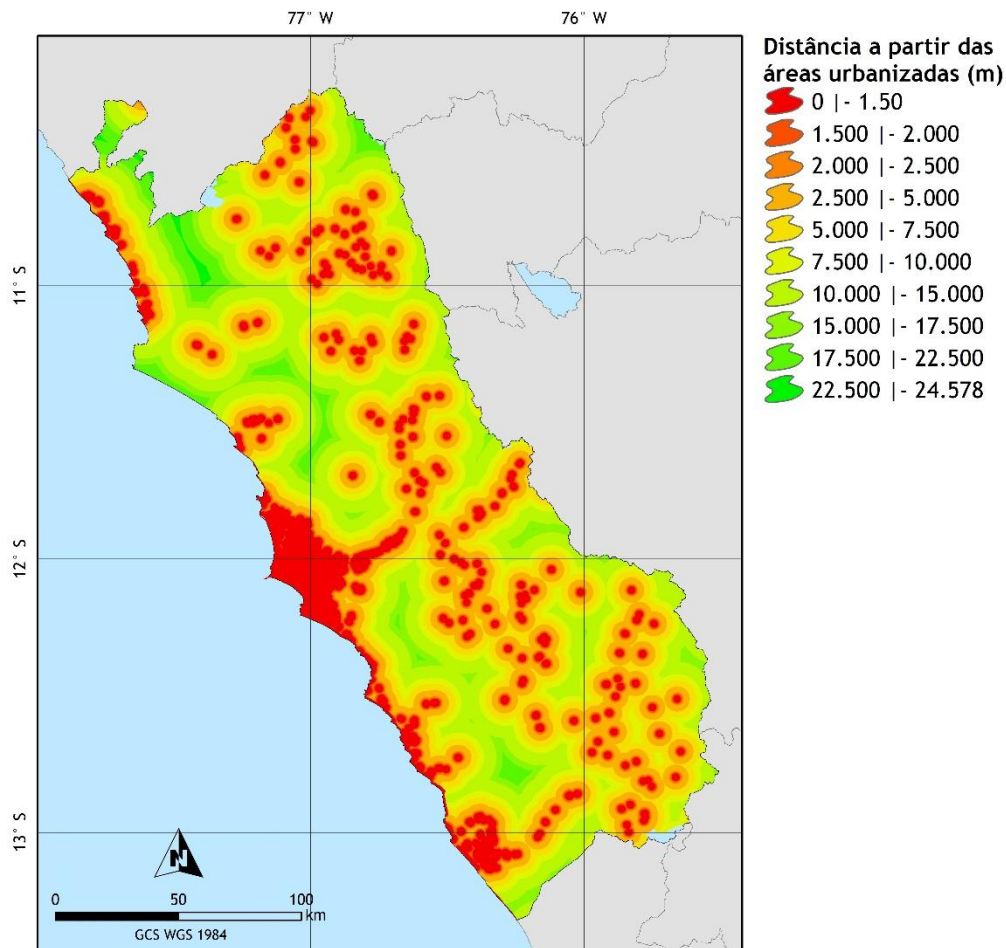


Figura 26. Distância a partir das áreas urbanizadas. Fonte: adaptado de NGI (2024).

Tabela 13. Quantitativos com relação as distâncias a partir de áreas urbanizadas

Distância a partir das áreas urbanizadas (m)	Área (km²)	%	Distância a partir das áreas urbanizadas (m)	Área (km²)	%
0 - 2.500	5.947,27	16,99	12.500 - 15.000	3.037,43	8,68
2.500 - 5.000	7.027,08	20,07	15.000 - 17.500	1.648,13	4,71
5.000 - 7.500	6.314,76	18,04	17.500 - 20.000	704,64	2,01
7.500 - 10.000	5.447,62	15,56	20.000 - 25.500	319,36	0,91
10.000 - 12.500	4.477,10	12,79	22.500 - 24.578	87,66	0,25
Total				35.011,04	100,00

Fonte: a autora (2026).

4.1.12 Precipitação média histórica

Conforme a Figura 27, observam-se que as médias históricas (normais climatológicas) entre o período de 1991-2020 apontam os meses de dezembro, janeiro, fevereiro e março como os mais chuvosos. Dessa forma, infere-se que nesses períodos a ocorrência de *huaicos* é mais acentuada. Ainda assim, mesmo nos meses mais chuvosos, o nível de precipitação é baixo, já sendo, o suficiente para desencadear *huaicos*.

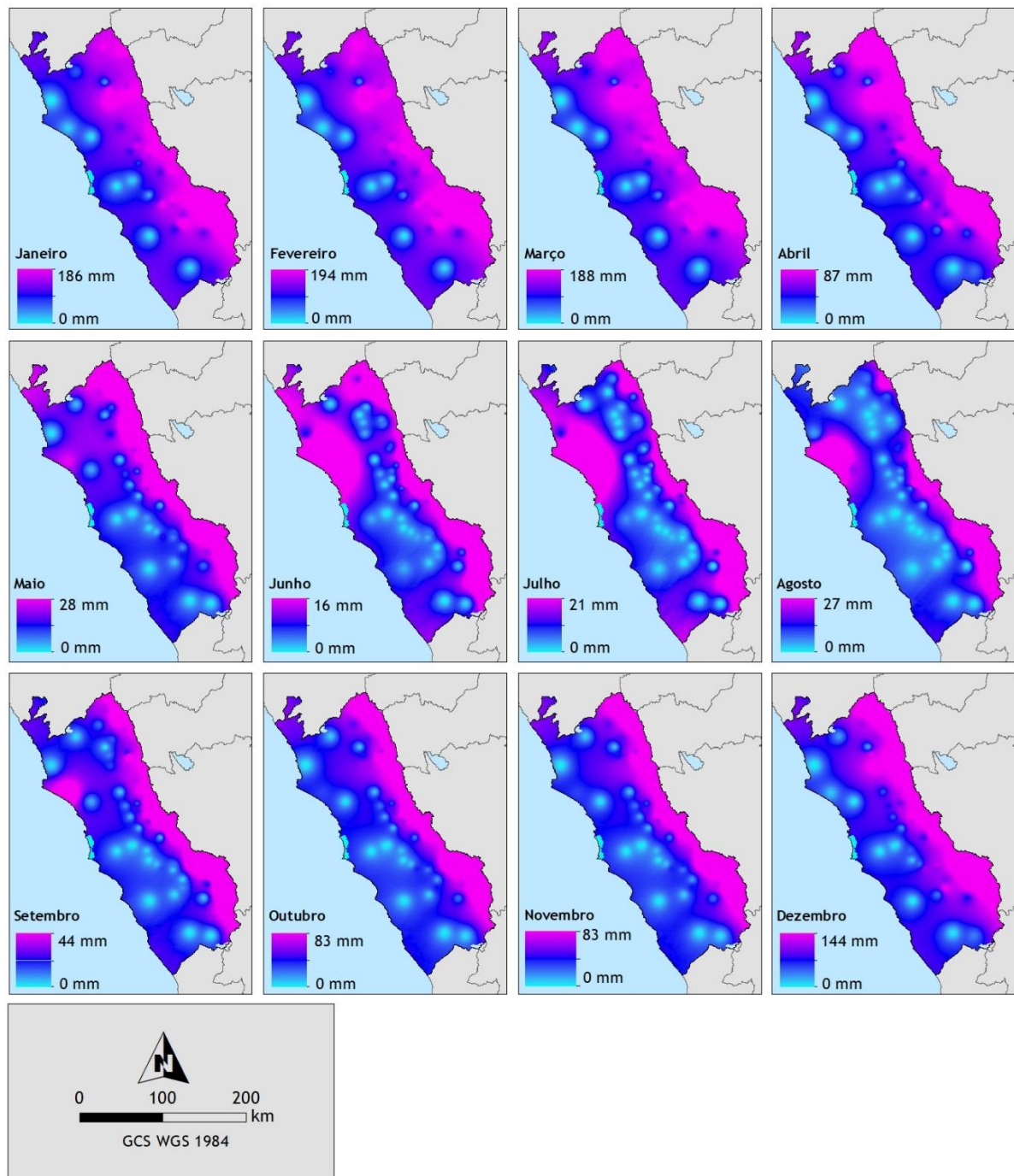


Figura 27. Precipitação média mensal histórica. Fonte: SENAMHI (2024).

Em relação ao aumento dos níveis de precipitação no Peru, destaca-se o fenômeno *El Niño* responsável pelo aquecimento das águas do Pacífico (INGEMMET, 2020). No campo da exemplificação, nos ciclos de 1982-1983, 1997-1998 e 2017, os níveis de precipitação foram extraordinários, atingindo níveis acima de 1.000 mm considerando um ciclo de três meses, desencadeando *huaicos* e inundações que levaram a perdas econômicas consideráveis e até mortes (INGEMMET, 2020).

Gonzales e Ingol (2021) apontaram a particularidade da formação do *Niño Costero* de 2017 que atingiu direta e fortemente a região costeira do Peru. Os autores concluíram que houve uma sucessão de eventos, onde a Oscilação da Década do Pacífico, em sua fase positiva, deu origem a processos acoplados oceano-atmosfera gerando ENSO Modoki e La Niña MODOKI, sendo que o ENSO costeiro é um subprocesso de La Niña Modoki da parte inferior do Modoki ENSO de 2017. O evento Modoki representa um padrão único de pressão tripolar ao nível do mar (2021).

Eventos como o El Niño Costeiro de 2017 são difíceis de prever e de estar preparado. E, ainda assim, englobar seus dados na modelagem permite uma avaliação da influência de seus níveis na ocorrência dos huaicos.

4.1.13 Correlação

Os resultados obtidos para as variáveis ambientais no QGIS apresentaram correlação máxima para cada variável com si mesma, resultando em 1. A correlação foi calculada com os coeficientes de correlação de Pearson e de Spearman. As variáveis apontadas em vermelho na Tabela 14 e Tabela 15 apresentaram alta correspondência entre si. Já os valores inferiores a 0,7 indicam níveis de correlação aceitáveis (Kumar et al., 2023). Ainda se destaca que, embora sejam apresentadas correlações entre as mesmas variáveis e entre as precipitações mensais, tais correlações não existem nos modelos. Portanto, são observadas as relações entre as variáveis geomorfologia, geologia, altitude, declividade e cada precipitação média mensal, isso para os modelos que consideram as variáveis do INGEMMET. Já para a avaliação com as variáveis segundo a bibliografia, acrescentam-se as relações entre as distâncias a partir de cursos hídricos e áreas urbanizadas, solos, ecossistemas, aspecto, *Terrain Ruggedness Index* (TRI) e *Topographic Wetness Index* (TWI).

A correlação de Pearson indica correlações especialmente altas entre as variáveis comparadas e as precipitações médias mensais. Entretanto, a correlação por Spearman descreve dados altamente correlacionados. Assim, um modelo com variáveis altamente relacionadas entre si aumenta a complexidade do modelo, segundo Kornejady, Ownegh e Bahremand (2017). Ainda assim, o algoritmo MaxEnt apresenta a configuração de não ser tão afetado pela correlação entre variáveis (Adab et al., 2018). No entanto, os resultados dos modelos podem ser afetados por essas condições.

Tabela 14. Matriz de correlação de Pearson

	Altitude	Aspecto	Declividade	Distância a partir de áreas urbanizadas	Distância a partir de cursos hídricos	Ecossistemas	Geologia	Geomorfologia	PM Jan	PM Fev	PM Mar	Pm Abr	PM Mai	PM Jun	PM Jul	PM Ago	PM Set	PM Out	PM Nov	PM Dez	Solos	TRI	TWI
Altitude	1	0,64	0,73	0,56	0,4	0,67	0,6	0,33	0,87	0,86	0,86	0,86	0,8	0,49	0,39	0,51	0,75	0,82	0,82	0,86	0,87	0,09	0,59
Aspecto	0,64	1	0,69	0,64	0,55	0,74	0,61	0,57	0,72	0,74	0,74	0,7	0,63	0,51	0,45	0,49	0,59	0,62	0,62	0,69	0,74	0,11	0,65
Declividade	0,73	0,69	1	0,63	0,45	0,78	0,69	0,32	0,74	0,76	0,76	0,72	0,64	0,44	0,36	0,43	0,57	0,64	0,64	0,71	0,81	0,15	0,63
Distância a partir de áreas urbanizadas	0,56	0,64	0,63	1	0,61	0,65	0,62	0,49	0,69	0,7	0,71	0,69	0,68	0,62	0,55	0,58	0,64	0,65	0,65	0,68	0,72	0,11	0,58
Distância a partir de cursos hídricos	0,4	0,55	0,45	0,61	1	0,55	0,47	0,55	0,53	0,54	0,53	0,51	0,52	0,69	0,67	0,64	0,54	0,48	0,48	0,51	0,55	0,07	0,48
Ecossistemas	0,67	0,74	0,78	0,65	0,55	1	0,74	0,54	0,79	0,82	0,82	0,76	0,63	0,47	0,4	0,45	0,57	0,63	0,63	0,74	0,87	0,15	0,66
Geologia	0,6	0,61	0,69	0,62	0,47	0,74	1	0,35	0,68	0,7	0,7	0,64	0,56	0,43	0,37	0,42	0,52	0,56	0,56	0,64	0,77	0,14	0,56
Geomorfologia	0,33	0,57	0,32	0,49	0,55	0,54	0,35	1	0,55	0,56	0,56	0,53	0,52	0,55	0,5	0,5	0,5	0,49	0,49	0,53	0,49	0,05	0,5
PM Jan	0,87	0,72	0,74	0,69	0,53	0,79	0,68	0,55	1	1	0,99	0,99	0,92	0,62	0,53	0,66	0,88	0,94	0,94	0,99	0,92	0,18	0,66
PM Fev	0,86	0,74	0,76	0,7	0,54	0,82	0,7	0,56	1	1	1	0,98	0,9	0,6	0,51	0,62	0,85	0,92	0,92	0,98	0,94	0,18	0,68
PM Mar	0,86	0,74	0,76	0,71	0,53	0,82	0,7	0,56	0,99	1	1	0,98	0,9	0,59	0,49	0,6	0,83	0,91	0,91	0,98	0,94	0,18	0,68
Pm Abr	0,86	0,7	0,72	0,69	0,51	0,76	0,64	0,53	0,99	0,98	0,98	1	0,95	0,64	0,53	0,66	0,9	0,96	0,96	0,99	0,9	0,18	0,65
PM Mai	0,8	0,63	0,64	0,68	0,52	0,63	0,56	0,52	0,92	0,9	0,9	0,95	1	0,76	0,66	0,77	0,96	0,98	0,98	0,96	0,81	0,17	0,59
PM Jun	0,49	0,51	0,44	0,62	0,69	0,47	0,43	0,55	0,62	0,6	0,59	0,64	0,76	1	0,98	0,97	0,82	0,71	0,71	0,64	0,56	0,1	0,46
PM Jul	0,39	0,45	0,36	0,55	0,67	0,4	0,37	0,5	0,53	0,51	0,49	0,53	0,66	0,98	1	0,97	0,76	0,62	0,62	0,54	0,47	0,09	0,4
PM Ago	0,51	0,49	0,43	0,58	0,64	0,45	0,42	0,5	0,66	0,62	0,6	0,66	0,77	0,97	0,97	1	0,88	0,76	0,76	0,67	0,56	0,11	0,44
PM Set	0,75	0,59	0,57	0,64	0,54	0,57	0,52	0,5	0,88	0,85	0,83	0,9	0,96	0,82	0,76	0,88	1	0,97	0,97	0,91	0,75	0,16	0,54
PM Out	0,82	0,62	0,64	0,65	0,48	0,63	0,56	0,49	0,94	0,92	0,91	0,96	0,98	0,71	0,62	0,76	0,97	1	1	0,97	0,82	0,17	0,58
PM Nov	0,82	0,62	0,64	0,65	0,48	0,63	0,56	0,49	0,94	0,92	0,91	0,96	0,98	0,71	0,62	0,76	0,97	1	1	0,97	0,82	0,17	0,58
PM Dez	0,86	0,69	0,71	0,68	0,51	0,74	0,64	0,53	0,99	0,98	0,98	0,99	0,96	0,64	0,54	0,67	0,91	0,97	0,97	1	0,9	0,18	0,64
Solos	0,87	0,74	0,81	0,72	0,55	0,87	0,77	0,49	0,92	0,94	0,94	0,9	0,81	0,56	0,47	0,56	0,75	0,82	0,82	0,9	1	0,19	0,68
TRI	0,09	0,11	0,15	0,11	0,07	0,15	0,14	0,05	0,18	0,18	0,18	0,18	0,17	0,1	0,09	0,11	0,16	0,17	0,17	0,18	0,19	1	0,03
TWI	0,59	0,65	0,63	0,58	0,48	0,66	0,56	0,5	0,66	0,68	0,68	0,65	0,59	0,46	0,4	0,44	0,54	0,58	0,58	0,64	0,68	0,03	1

Legenda: P.M. – precipitação média. Fonte: a autora (2026).

Tabela 15. Matriz de correlação de Spearman

	Altitude	Aspecto	Declividade	Distância a partir de áreas urbanizadas	Distância a partir de cursos hídricos	Ecossistemas	Geologia	Geomorfologia	PM Jan	PM Fev	PM Mar	Pm Abr	PM Mai	PM Jun	PM Jul	PM Ago	PM Set	PM Out	PM Nov	PM Dez	Solos	TRI	TWI
Altitude	1	0,91	0,95	0,9	0,9	0,91	0,91	0,9	0,94	0,95	0,94	0,94	0,93	0,91	0,9	0,91	0,93	0,94	0,94	0,94	0,94	0,94	0,75
Aspecto	0,91	1	0,92	0,84	0,85	0,85	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,84	0,91	0,73
Declividade	0,95	0,92	1	0,9	0,9	0,92	0,92	0,88	0,91	0,91	0,91	0,91	0,91	0,9	0,89	0,89	0,9	0,91	0,91	0,91	0,92	0,98	0,76
Distância a partir de áreas urbanizadas	0,9	0,84	0,9	1	0,94	0,93	0,94	0,92	0,93	0,93	0,93	0,94	0,94	0,95	0,95	0,95	0,94	0,94	0,94	0,93	0,93	0,91	0,72
Distância a partir de cursos hídricos	0,9	0,85	0,9	0,94	1	0,93	0,94	0,94	0,93	0,93	0,93	0,93	0,94	0,95	0,95	0,95	0,94	0,93	0,93	0,93	0,93	0,91	0,72
Ecossistemas	0,91	0,85	0,92	0,93	0,93	1	0,95	0,93	0,94	0,94	0,94	0,94	0,93	0,92	0,92	0,92	0,93	0,93	0,93	0,94	0,97	0,93	0,73
Geologia	0,91	0,84	0,92	0,94	0,94	0,95	1	0,91	0,94	0,94	0,94	0,94	0,93	0,93	0,93	0,93	0,93	0,94	0,94	0,94	0,96	0,93	0,72
Geomorfologia	0,9	0,84	0,88	0,92	0,94	0,93	0,91	1	0,94	0,94	0,94	0,94	0,95	0,96	0,96	0,96	0,95	0,95	0,95	0,94	0,93	0,89	0,72
PM Jan	0,94	0,84	0,91	0,93	0,93	0,94	0,94	0,94	1	1	1	1	0,98	0,96	0,95	0,96	0,98	0,99	0,99	1	0,97	0,92	0,73
PM Fev	0,95	0,84	0,91	0,93	0,93	0,94	0,94	0,94	1	1	1	1	0,98	0,96	0,95	0,96	0,98	0,99	0,99	1	0,97	0,92	0,73
PM Mar	0,94	0,84	0,91	0,93	0,93	0,94	0,94	0,94	1	1	1	1	0,98	0,96	0,95	0,95	0,98	0,99	0,99	1	0,97	0,92	0,73
Pm Abr	0,94	0,84	0,91	0,94	0,93	0,94	0,94	0,94	1	1	1	1	0,99	0,96	0,96	0,96	0,98	1	1	1	0,97	0,92	0,73
PM Mai	0,93	0,84	0,91	0,94	0,94	0,93	0,93	0,95	0,98	0,98	0,98	0,99	1	0,99	0,98	0,98	1	1	1	0,99	0,95	0,91	0,72
PM Jun	0,91	0,84	0,9	0,95	0,95	0,92	0,93	0,96	0,96	0,96	0,96	0,96	0,99	1	1	0,99	0,99	0,98	0,98	0,97	0,93	0,9	0,72
PM Jul	0,9	0,84	0,89	0,95	0,95	0,92	0,93	0,96	0,95	0,95	0,95	0,96	0,98	1	1	1	0,99	0,97	0,97	0,96	0,93	0,9	0,72
PM Ago	0,91	0,84	0,89	0,95	0,95	0,92	0,93	0,96	0,96	0,96	0,95	0,96	0,98	0,99	1	1	0,99	0,98	0,98	0,96	0,93	0,9	0,72
PM Set	0,93	0,84	0,9	0,94	0,94	0,93	0,93	0,95	0,98	0,98	0,98	0,98	1	0,99	0,99	0,99	1	0,99	0,99	0,98	0,95	0,91	0,72
PM Out	0,94	0,84	0,91	0,94	0,93	0,93	0,94	0,95	0,99	0,99	0,99	1	1	0,98	0,97	0,98	0,99	1	1	1	0,96	0,92	0,72
PM Nov	0,94	0,84	0,91	0,94	0,93	0,93	0,94	0,95	0,99	0,99	0,99	1	1	0,98	0,97	0,98	0,99	1	1	1	0,96	0,92	0,72
PM Dez	0,94	0,84	0,91	0,93	0,93	0,94	0,94	0,94	1	1	1	1	0,99	0,97	0,96	0,96	0,98	1	1	1	0,97	0,92	0,72
Solos	0,94	0,84	0,92	0,93	0,93	0,97	0,96	0,93	0,97	0,97	0,97	0,97	0,95	0,93	0,93	0,93	0,95	0,96	0,96	0,97	1	0,93	0,73
TRI	0,94	0,91	0,98	0,91	0,91	0,93	0,93	0,89	0,92	0,92	0,92	0,92	0,91	0,9	0,9	0,9	0,91	0,92	0,92	0,92	0,93	1	0,74
TWI	0,75	0,73	0,76	0,72	0,72	0,73	0,72	0,72	0,73	0,73	0,73	0,73	0,72	0,72	0,72	0,72	0,72	0,72	0,72	0,72	0,73	0,74	1

Legenda: P.M. – precipitação média. Fonte: a autora (2026).

4.2 MODELAGENS

Reforça-se que a abordagem adotada envolveu a elaboração de modelos mensais, devido à variação da precipitação no país, que chega a nula em parte dos meses do ano. Dessa forma, espera-se uma associação entre a ocorrência de huaicos e os níveis de precipitação.

Além disso, foram elaborados dois conjuntos de modelos mensais para cada algoritmo, conforme já apresentado na metodologia deste trabalho: um primeiro conjunto de modelos considerando as variáveis ambientais identificadas como determinantes da ocorrência de huaicos pelo INGEMMET e outro conjunto de modelos utilizando as variáveis ambientais desencadeantes de inundações, deslizamentos e fluxo de detritos, conforme outros trabalhos desenvolvidos com machine learning para a modelagem de fenômenos decorrentes de desastres naturais.

Além disso, para os demais algoritmos, regressão logística, floresta aleatória, máquina de vetores de suposições e redes neurais artificiais realizou-se um teste de aplicação, já para a área completa de estudo, considerando a precipitação média histórica de janeiro e as variáveis ambientais do INGEMMET, com o intuito de verificar o funcionamento dos algoritmos.

A interpretação dos resultados para cada algoritmo seguirá a organização conforme Santos et al. (2025) (Apêndice A).

4.2.1 Floresta Aleatória

Na Tabela 16 apresenta-se a qualidade dos modelos de floresta aleatória. Todos os modelos apresentam qualidade adequada no “AUC de Teste”, que deve ser superior a 0,7. A área sob a curva varia de 0 a 1, sendo 1 indicativo da maior precisão. As faixas de AUC de 0,5–0,6, 0,6–0,7, 0,7–0,8, 0,8–0,9 e 0,9–1 classificam os modelos como fracos, medianos, bons, muito bons e excelentes, respectivamente, de acordo com Yesilnacar (2005), sendo que os modelos foram classificados como muito bons (INGEMMET) e excelentes (bibliografia).

O índice Kappa representa cinco classes — $k < 0,4$; $0,4 < k < 0,55$; $0,55 < k < 0,85$; $0,85 < k < 0,99$, respectivamente — com validação fraca, mediana, boa, excelente e muito excelente (Cohen, 1960). No índice Kappa, os modelos gerados com todas as variáveis apresentaram desempenho excelente, assim como os gerados com base apenas nas variáveis ambientais descritas pelo INGEMMET.

Mesmo apresentando classificações excelentes, observa-se um desempenho menos satisfatório dos modelos com base apenas nas variáveis do INGEMMET, considerando, à

primeira vista, mais adequados os modelos que utilizam as variáveis conforme apontadas na bibliografia, quando do uso do algoritmo floresta aleatória.

Tabela 16. Qualidade dos modelos de Floresta Aleatória

Mês	INGEMMET		Variáveis segundo a bibliografia	
	AUC	Kappa	AUC	Kappa
Janeiro	0,86	0,60	0,92	0,78
Fevereiro	0,85	0,61	0,92	0,77
Março	0,85	0,60	0,92	0,78
Abril	0,83	0,61	0,92	0,78
Maio	0,86	0,56	0,93	0,81
Junho	0,89	0,58	0,93	0,82
Julho	0,90	0,59	0,93	0,86
Agosto	0,89	0,58	0,92	0,83
Setembro	0,86	0,57	0,92	0,79
Outubro	0,85	0,59	0,92	0,79
Novembro	0,85	0,59	0,91	0,79
Dezembro	0,86	0,65	0,92	0,78

Fonte: a autora (2026).

As Figuras 28 e 29 apresentam os resultados dos modelos, reclassificados para representar as áreas com suscetibilidade de ocorrência de huaicos. Em média, 11% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos quando considerados os modelos que utilizam todas as variáveis. Em média, 14% do Departamento de Lima apresenta suscetibilidade de ocorrência de huaicos nos modelos que consideram apenas as variáveis de acordo com o INGEMMET

Para os modelos que consideram as variáveis do INGEMMET, no mês de agosto, o modelo resultou em uma área de 4.666,35 km², 13,32% da área do departamento de Lima, que apresenta suscetibilidade de ocorrência de huaicos, chegando a um percentual máximo de 14,38% (5.033,11 km²) de áreas com suscetibilidade no mês de abril. Nos modelos que consideram as variáveis de acordo com a bibliografia, o mês de janeiro apresenta a menor quantidade de áreas com suscetibilidade de ocorrência de huaicos no departamento de Lima, 11,17% (3.911,97 km²). O mês de maio registrou a maior porcentagem de áreas com suscetibilidade existente, 11,78% (4.125,00 km²).

Com os modelos do INGEMMET, há maior variação nas porcentagens de áreas com suscetibilidade de huaicos. Nos modelos a partir das variáveis da bibliografia, embora haja mais variáveis, as respostas foram mais semelhantes entre si, em termos de área.

Espacialmente, observa-se a concentração de suscetibilidade de huaicos nas proximidades dos leitos encaixados de cursos hídricos, similar à configuração dos próprios dados de ocorrência, isso para a modelagem considerando as variáveis conforme a bibliografia.

Já nos dados dos modelos que consideram apenas as variáveis do INGEMMET, observa-se um aumento da suscetibilidade de ocorrência de huaicos em áreas costeiras.

Em relação à coincidência entre os modelos mensais, a Tabela 17 e a Tabela 18 contém as matrizes de coincidência das áreas com suscetibilidade de ocorrência para cada modelo, tomando os meses na horizontal como referência. Em média, há uma coincidência de 75% entre os modelos mensais considerando as variáveis da bibliografia. A coincidência entre os modelos mensais, considerando as variáveis do INGEMMET, é de 71%.

Isso revela que há sim um comportamento variável conforme a precipitação mensal vai sendo alterada. Especialmente nos meses mais secos, de maio a agosto, a correspondência entre os modelos é reduzida até atingir apenas 56%. No entanto, retoma-se o fato de que para os modelos com as variáveis da bibliografia, as maiores porcentagens de áreas com suscetibilidade de ocorrência de huaicos, foram registradas nos meses de menores precipitações e as menores porcentagens, registradas nos meses de maiores precipitações. Destaca-se que, mesmo no conjunto de modelos com base em variáveis bibliográficas, que apresentam menor variação entre as áreas dos resultados mensais, ainda assim há uma porcentagem relevante de áreas que não convergem.

Para além da falta de correspondência linear entre o nível de precipitação e as áreas com suscetibilidade de ocorrência de huaicos, pontua-se, espacialmente, que, embora a precipitação aumente de forma geral à medida que se adentra no continente, não se observa essa correspondência nas áreas apontadas como com suscetibilidade de ocorrência de huaicos.

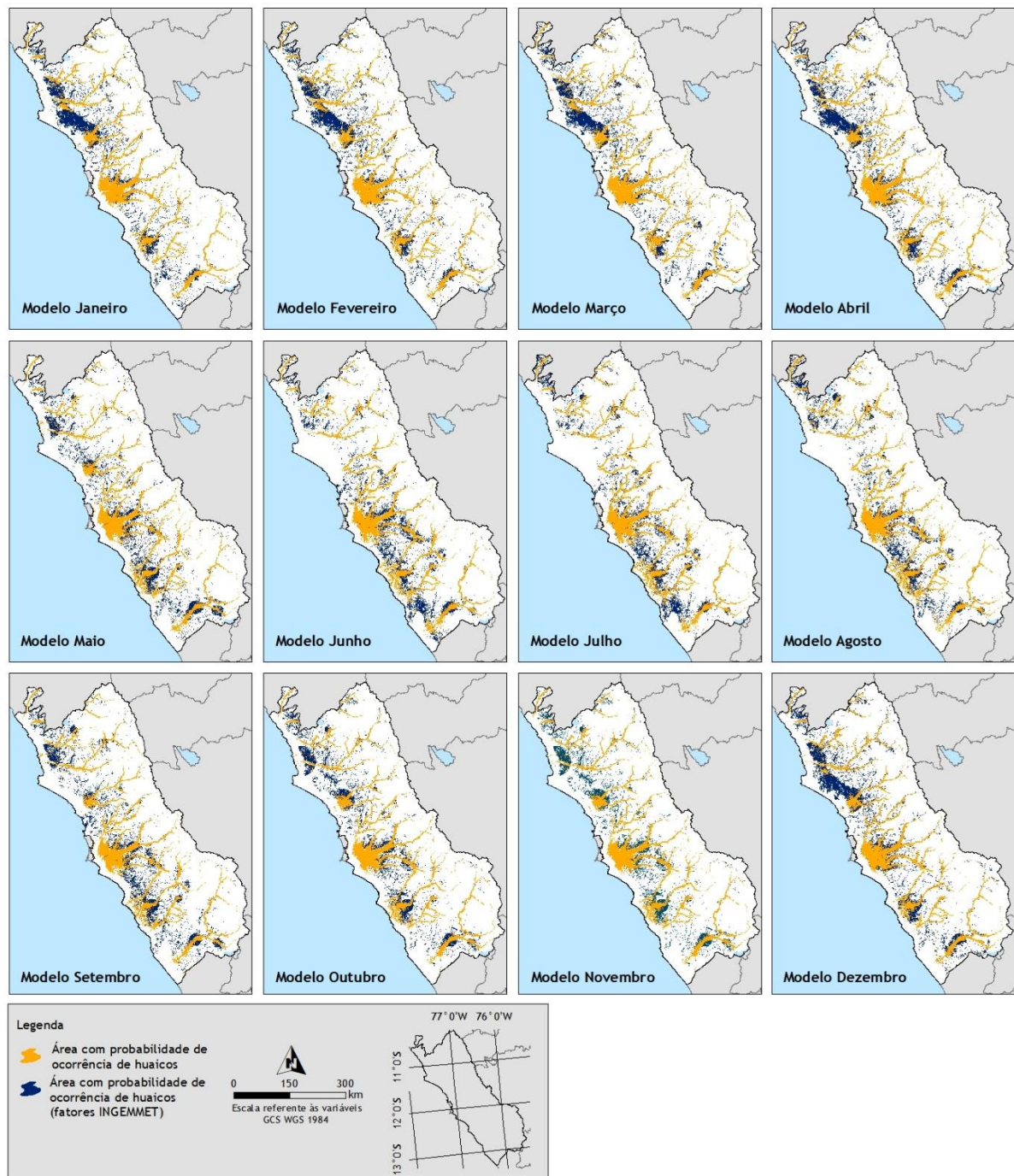


Figura 28. Resultados de suscetibilidade considerando os modelos mensais de Floresta Aleatória. Fonte: a autora (2026).

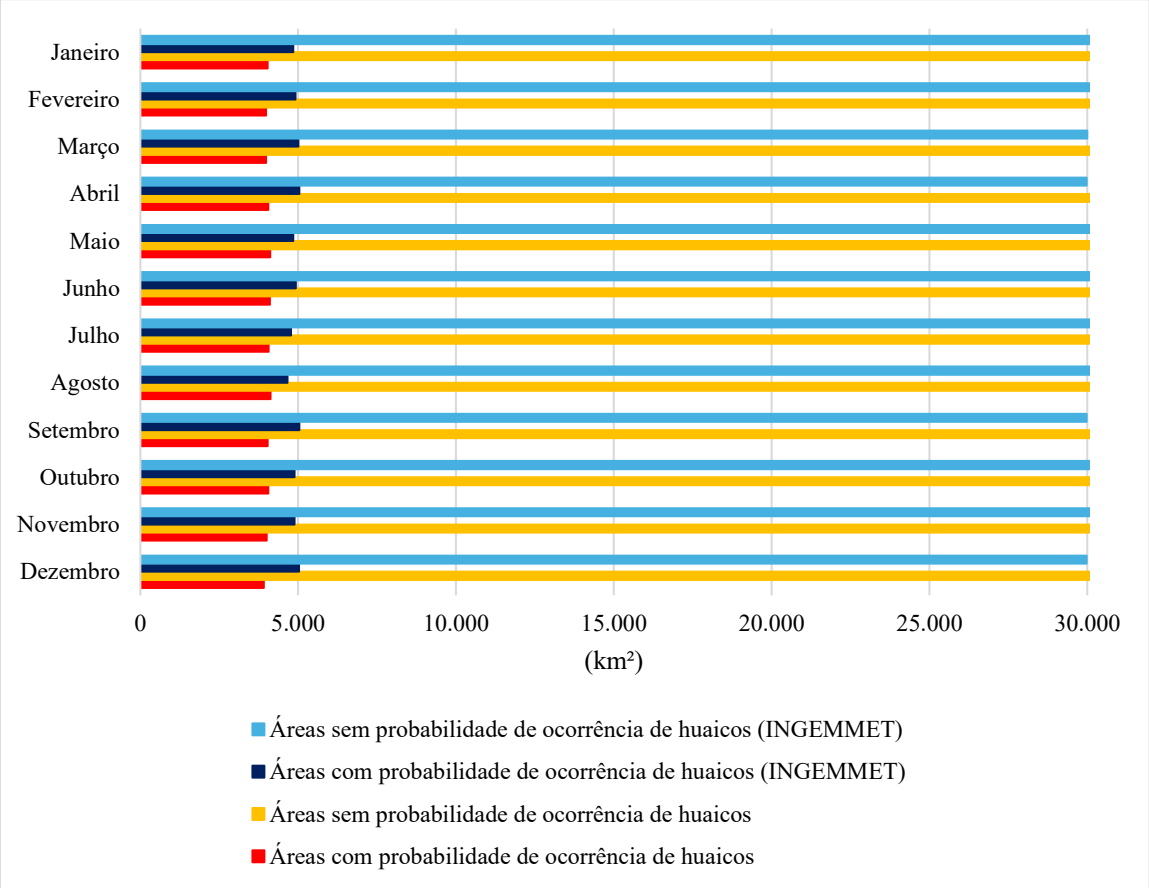


Figura 29. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos Floresta Aleatória.
Fonte: a autora (2026).

Tabela 17. Coincidência entre modelos mensais Floresta Aleatória – Variáveis segundo o INGEMMET (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	89	85	82	69	61	59	62	68	71	71	100
Fev	90	100	87	82	69	59	58	61	67	71	71	100
Mar	88	89	100	84	68	59	59	62	67	72	72	100
Abr	86	84	84	100	73	62	61	65	70	76	76	100
Maio	69	68	66	70	100	72	71	75	83	83	83	100
Jun	61	59	58	60	74	100	86	84	75	68	68	100
Jul	58	57	56	58	70	83	100	85	72	64	64	100
Ago	59	58	58	60	72	79	83	100	76	68	68	100
Set	70	69	67	70	86	77	76	82	100	82	82	100
Out	72	71	70	74	84	68	66	71	79	100	100	100
Nov	72	71	70	74	84	68	66	71	79	100	100	100
Dez	72	71	70	74	84	68	66	71	79	100	100	100

Fonte: a autora (2026).

Tabela 18. Coincidência entre modelos mensais Floresta Aleatória – Variáveis segundo a bibliografia (%)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	89	85	82	69	61	59	62	68	71	71	83
Fev	90	100	87	82	69	59	58	61	67	71	71	82
Mar	88	89	100	84	68	59	59	62	67	72	72	83
Abr	86	84	84	100	73	62	61	65	70	76	76	87
Mai	69	68	66	70	100	72	71	75	83	83	83	71
Jun	61	59	58	60	74	100	86	84	75	68	68	60
Jul	58	57	56	58	70	83	100	85	72	64	64	57
Ago	59	58	58	60	72	79	83	100	76	68	68	60
Set	70	69	67	70	86	77	76	82	100	82	82	71
Out	72	71	70	74	84	68	66	71	79	100	100	76
Nov	72	71	70	74	84	68	66	71	79	100	100	76
Dez	86	84	84	87	74	61	60	64	71	78	78	100

Fonte: a autora (2026).

Com relação às variáveis que mais contribuem para os modelos, na Figura 30 observa-se que a distância a partir dos cursos d'água é o fator que mais contribui para os modelos mensais, na ordem de 24%, seguida pela altitude, com cerca de 16% de contribuição. A precipitação média mensal e a distância a partir de áreas urbanizadas possuem contribuições similares, com a terceira e quarta maior contribuição, com aproximadamente 11% cada. Os ecossistemas têm menor influência nos modelos, de cerca de 1%.

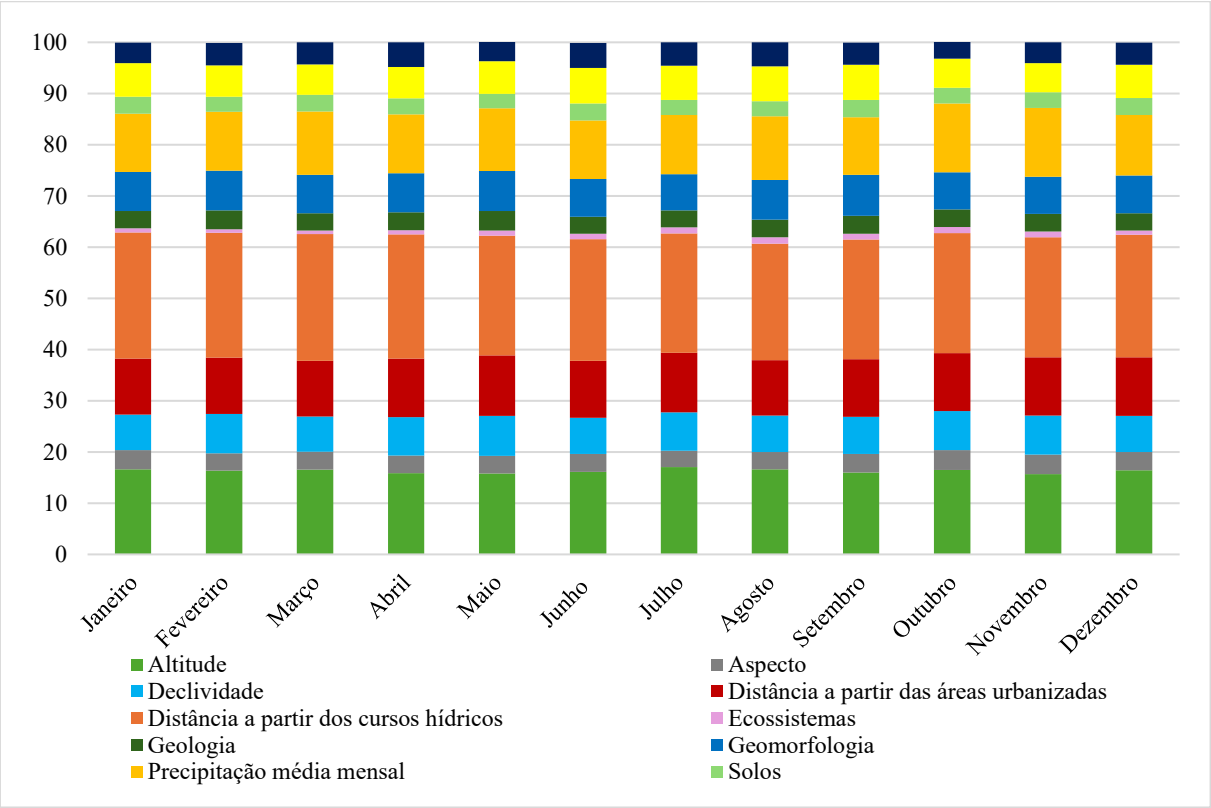
Ainda, para a distância a partir dos cursos hídricos, as áreas com suscetibilidade de ocorrência se concentram a até 1.500 metros. Embora a altitude vá aumentando gradualmente, visualmente identificam-se nas formas do relevo, os leitos de rios, sejam eles perenes ou efêmeros como áreas com suscetibilidade de ocorrência. Observa-se, ainda, a presença de assentamentos humanos nas proximidades das áreas de maior suscetibilidade. O comportamento da precipitação é bastante variável. E, justamente em pontos próximos de áreas com suscetibilidade de ocorrência, observam-se precipitações baixas.

Das variáveis consideradas como deflagrações dos huaicos pelo INGEMMET, nos modelos que consideram as variáveis influentes no fenômeno pela bibliografia, a declividade, a geologia e geomorfologia possuem baixa contribuição, entre 3 a 7%. Já a altitude e a precipitação média mensal apresentaram contribuições significativas.

Das variáveis apontadas somente pela bibliografia, assim como os ecossistemas, possuem baixa importância, o aspecto, os solos e os índices do terreno, TRI e TWI, retornaram uma baixa influência nos modelos.

Com relação ao comportamento comparado à distribuição normal, nos dados que mais contribuíram para os modelos, observa-se maior sensibilidade. Destaca-se que, embora somente o dado de precipitação média mensal tenha sido alterado a cada novo modelo mensal, isso foi suficiente para modificar as respostas das demais variáveis quanto à contribuição e à variação

na distribuição normal.



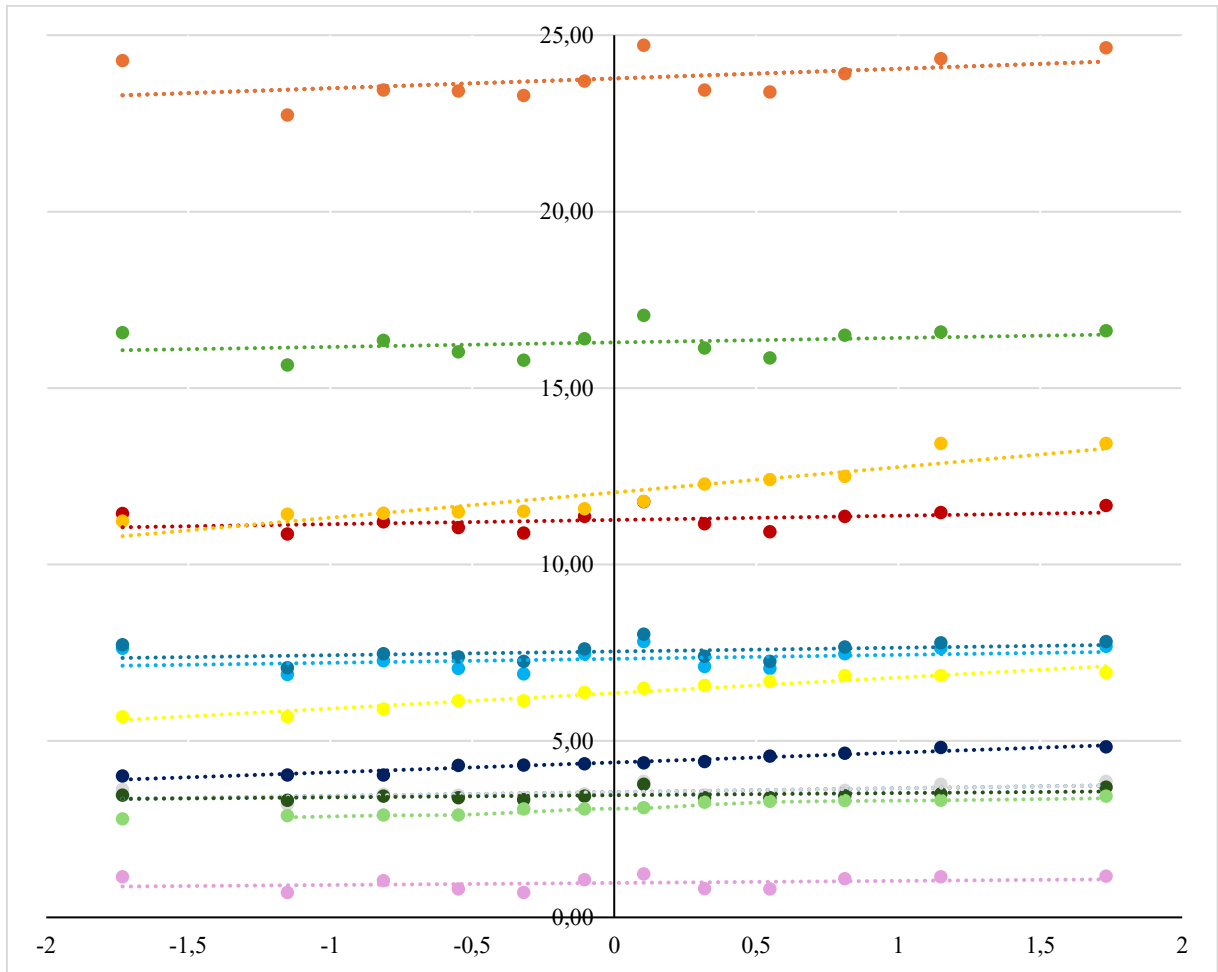


Figura 30. Contribuição das variáveis para cada modelo mensal de Floresta Aleatória, considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada variável em relação a uma distribuição normal.
Fonte: a autora (2026).

Nos modelos mensais que consideram apenas as variáveis indicadas pelo INGEMMET como relevantes para a ocorrência de huaicos, todos os meses apresentaram uma distribuição semelhante das contribuições das variáveis ambientais, com a mais significativa sendo a altitude. O segundo fator mais importante é a precipitação média mensal. Os demais fatores têm menor influência e contribuições semelhantes (Figura 31).

Embora a altitude tenha representado o maior valor de contribuição, com uma participação média de 45%, essa contribuição variou de 35% em agosto, a 52% em abril. Esses meses foram os que, respectivamente, apresentaram as menores e as maiores áreas com suscetibilidade de ocorrência de huaicos. Para a precipitação média mensal, a contribuição média foi de 28%, sendo a menor no mês de fevereiro (17%), seguida de abril (19%) e a maior contribuição da variável ocorrendo no mês de agosto (35%), com valores bastante próximos entre maio, julho, outubro e novembro.

No entanto, esses meses possuem níveis de precipitação diferentes, sendo os meses de outono e inverno menos chuvosos historicamente, que os meses de outubro e novembro. E quanto aos meses mais chuvosos, outubro e novembro não correspondem aos meses mais chuvosos do ano, que são dezembro, janeiro, fevereiro e março. Dessa forma, a relação entre a participação nos modelos e os níveis de precipitação não é linear, sendo então a mudança do nível de precipitação, espacialmente mais importante.

A respeito do comportamento da distribuição normal, a precipitação média mensal e a altitude predominam nos modelos mensais e apresentam variações inversas ao longo do ano. Novamente, destaca-se o impacto da precipitação média mensal nos modelos e nas demais variáveis. Neste caso, de avaliação com base somente nas variáveis do INGEMMET, o impacto é basicamente na altitude. Dessa forma, observa-se que, embora o aumento de variáveis nos modelos tende a tornar os modelos mais complexos, na situação deste trabalho, os modelos que consideram mais variáveis, tem a participação da precipitação média mensal suavizada ao longo dos meses, tornando os resultados mais uniformes.

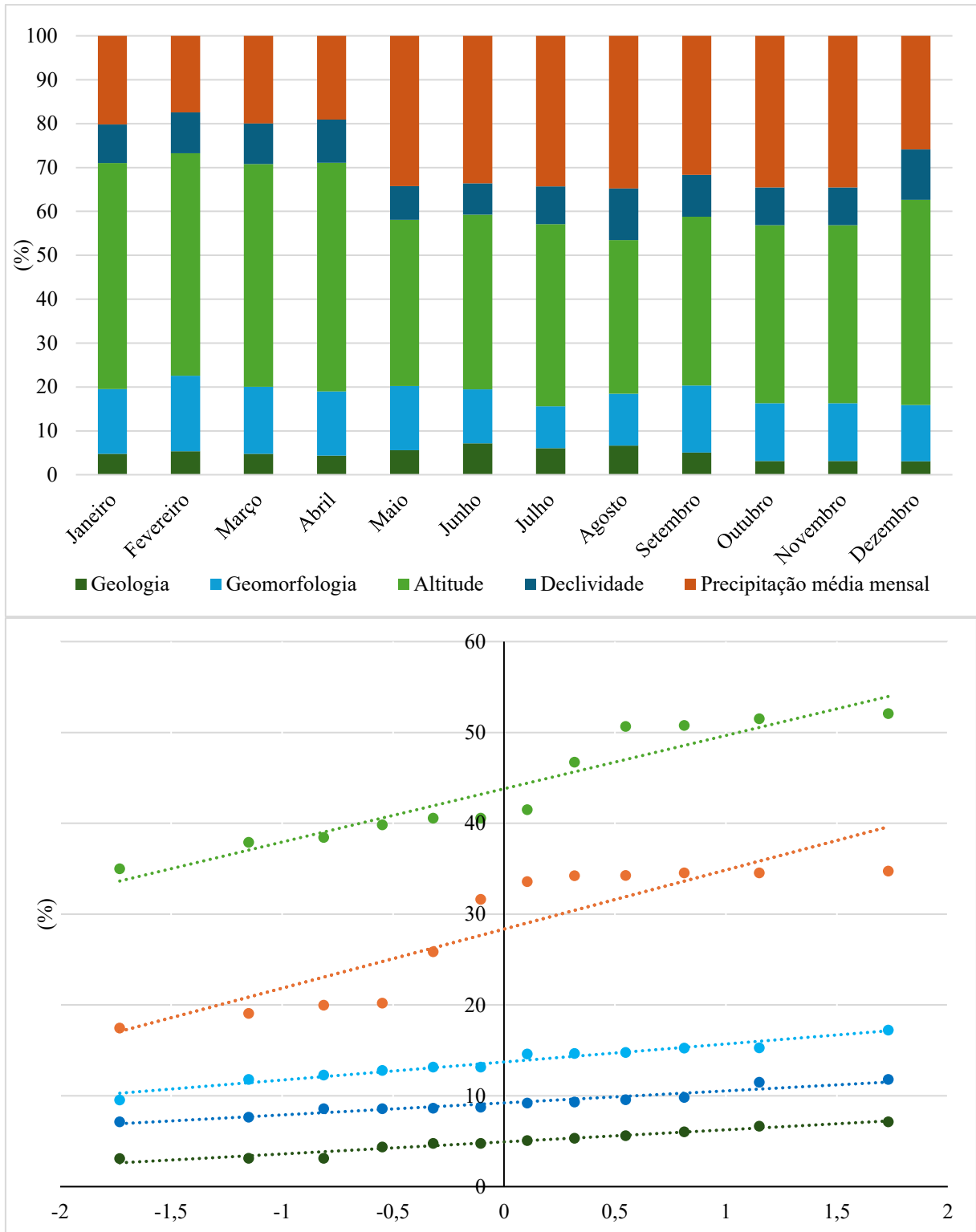


Figura 31. Contribuição das variáveis para cada modelo mensal de Floresta Aleatória, considerando as variáveis ambientais indicadas pelo INGEMMET (a) e o comportamento de cada variável em relação a uma distribuição normal (b). Fonte: a autora (2026).

Huang et al. (2022) consideraram como variáveis ambientais a altitude, declividade, orientação, precipitação, índice de vegetação por diferença normalizada (NDVI), uso do solo, relevo, geologia, distância até as falhas, densidade de aldeias, distância até os rios e distância

até as estradas. O modelo elaborado pelos autores atingiu excelentes níveis de qualidade ($AUC = 0,97$). No modelo de Huang et al. (2022) a distância a partir das estradas foi o fator predominante. Os autores destacam que antes, os impactos das atividades humanas raramente eram considerados em modelos de suscetibilidade, mas que gradualmente a antropização se tornou um gatilho que não pode ser ignorado na formação de fluxos de detritos.

Si et al. (2020) destacam que o uso do algoritmo floresta aleatória tem como uma das principais vantagens determinar com eficiência as regiões do terreno com condições adequadas para a ocorrência de fluxos de detritos, tendo como base um inventário de desastres e dados geoespaciais suficientes. Ao mesmo tempo, os autores descrevem que o modelo de floresta aleatória resultante se limitou a localizar as regiões que atendessem aos critérios adequados à ocorrência de fluxo de detritos. Porém, o modelo não considerou os mecanismos que desencadeiam o fenômeno ou as propriedades físico-mecânicas. Em seu trabalho, o modelo levou algumas regiões com terrenos adequados a serem consideradas propensas a fluxos de detritos, mas, devido às restrições nas propriedades físico-mecânicas, como a condutividade hidráulica ou a espessura do sedimento solto, essas regiões dificilmente se tornarão propensas a fluxos de detritos na realidade. Neste trabalho, verificam-se condições similares, pois os modelos que consideram as variáveis segundo a bibliografia, consideraram como mais importante a distância a partir dos cursos hídricos.

Xiong et al. (2020) abordam o uso do algoritmo floresta aleatória de uma maneira diferente. A própria seleção das variáveis já fez uso do algoritmo. Os autores adotaram um método seleção de variáveis regressiva, mantendo para a geração dos resultados os fatores ótimos. Nessa modalidade, 18 variáveis foram utilizadas para alimentar o modelo de floresta aleatória. Em seguida, o fator com a menor importância foi eliminado e um novo modelo foi implementado, repetindo esse procedimento até que restar apenas um fator. Selecionou-se o conjunto de fatores com o melhor desempenho. Esta pode ser uma alternativa para melhorar os resultados para o conjunto de dados com as variáveis segundo a bibliografia, visto que, fatores como ecossistemas, geologia, solos e aspecto apresentaram baixa contribuição nos modelos. Ainda, é uma forma de avaliar a validade de consideração tanto das variáveis apontadas pelo INGEMMET, quanto da relevância da precipitação média mensal.

Ainda, Xiong et al. (2020), assim como Huang et al. (2022), optaram por reclassificar as áreas com suscetibilidade não somente em áreas com ou sem suscetibilidade de ocorrência de fluxo de detritos, como em áreas de muito baixa, baixa, moderada, alta e muito alta suscetibilidade. Essa pode ser uma possibilidade futura de exibição dos resultados.

4.2.2 Máquina de Vetores de Suporte

Na Tabela 19, apresenta-se a qualidade dos modelos de máquinas de vetores de suporte. Todos os modelos apresentam qualidade adequada, no “AUC de Teste” apresentando resultados superiores a 0,7. Para esse parâmetro, os modelos de máquinas de vetores de suporte, tanto considerando as variáveis segundo a bibliografia quanto segundo o INGEMMET, foram classificados como muito bons, de acordo com a classificação de Yesilnacar (2005).

No índice Kappa, os modelos gerados com as variáveis da bibliografia apresentaram desempenho excelente. Os modelos gerados com base apenas nas variáveis ambientais descritas pelo INGEMMET como influentes na formação de huaicos apresentaram, em sua maioria, índices Kappa quase máximos, classificados como muito excelentes.

Diferentemente dos modelos que utilizaram floresta aleatória, para os modelos que utilizaram o algoritmo de máquinas de vetores de suporte, o índice Kappa apresentou índices melhores, de forma geral, para ambos os conjuntos de modelos, e os modelos que consideram somente as variáveis do INGEMMET apresentaram uma representatividade melhor da realidade.

Tabela 19. Qualidade dos modelos de Máquina de Vetores de Suporte

Mês	INGEMMET		Variáveis segundo a bibliografia	
	AUC	Kappa	AUC	Kappa
Janeiro	0,84	0,88	0,88	0,76
Fevereiro	0,84	0,88	0,88	0,76
Março	0,84	0,88	0,88	0,76
Abril	0,83	0,89	0,88	0,76
Maio	0,85	0,89	0,88	0,75
Junho	0,88	0,83	0,88	0,75
Julho	0,88	0,83	0,88	0,75
Agosto	0,88	0,85	0,88	0,76
Setembro	0,85	0,88	0,88	0,76
Outubro	0,84	0,90	0,88	0,76
Novembro	0,84	0,90	0,88	0,76
Dezembro	0,84	0,88	0,88	0,76

Fonte: a autora (2026).

As Figuras 32 e 33 apresentam os resultados dos modelos, reclassificados para representar as áreas com suscetibilidade de ocorrência de huaicos. Em média, 29% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos quando considerados os modelos que utilizam as variáveis conforme a bibliografia.

Ao analisar os modelos gerados com base nas variáveis do INGEMMET identificadas como fatores de ocorrência de huaicos, a qualidade dos modelos melhorou e as áreas com suscetibilidade de ocorrência diminuíram. Em média, 26% do Departamento de Lima apresenta

suscetibilidade de ocorrência de huaicos. O uso de variáveis que não contribuem efetivamente para a ocorrência de huaicos tornou os modelos mais complexos, levando o algoritmo a buscar relações inexistentes entre os dados de amostra e as variáveis ambientais. Isso resultou em uma qualidade de modelo inferior àqueles que utilizaram variáveis de contribuição relevantes.

Ainda assim, destaca-se que o aumento do índice Kappa também se deveu ao fato de as áreas com suscetibilidade de ocorrência de huaicos terem sido praticamente dobradas, quando comparadas aos resultados dos modelos com o algoritmo floresta aleatória.

Explicita-se que para os modelos com as variáveis do INGEMMET, o mês de outubro retornou a menor quantidade de áreas com suscetibilidade de ocorrência de huaicos, com 23,93% (8.378,21 km²) das áreas do departamento de Lima com suscetibilidade. O mês de junho apresentou a maior quantidade de áreas com suscetibilidade, 27,54% (9.643,71 km²). Para os modelos com as variáveis conforme a bibliografia, o mês de junho apresentou as menores áreas com suscetibilidade de ocorrência de huaicos, 27,45% do departamento de Lima (9.610,09 km²) e o mês de abril retornou as maiores áreas com suscetibilidade de ocorrências, 30,30% (10.609,42 km²). Este comportamento é similar aos modelos com o algoritmo floresta aleatória.

Para os modelos de máquinas de vetores de suporte, o comportamento espacial é similar aos dados espaciais dos modelos com o algoritmo floresta aleatória. Os modelos com as variáveis conforme a bibliografia, apresentam uma configuração de proximidade com cursos hídricos, enquanto os modelos com as variáveis do INGEMMET também ocorrem nas porções mais costeiras.

Em relação à coincidência entre os modelos mensais, a Tabela 20 e a Tabela 21 **Tabela 23** contém as matrizes de coincidência das áreas com suscetibilidade de ocorrência para cada modelo, tomando os meses na horizontal como referência. Em média, há uma coincidência de 98% entre os modelos mensais que consideram as variáveis conforme a bibliografia. A coincidência para os modelos segundo o INGEMMET s é 97% em média.

Há uma elevada convergência entre os modelos mensais, tanto pelo INGEMMET, quanto pelas variáveis da bibliografia. Isso indica uma baixa variação a partir dos dados precipitações médias mensais.

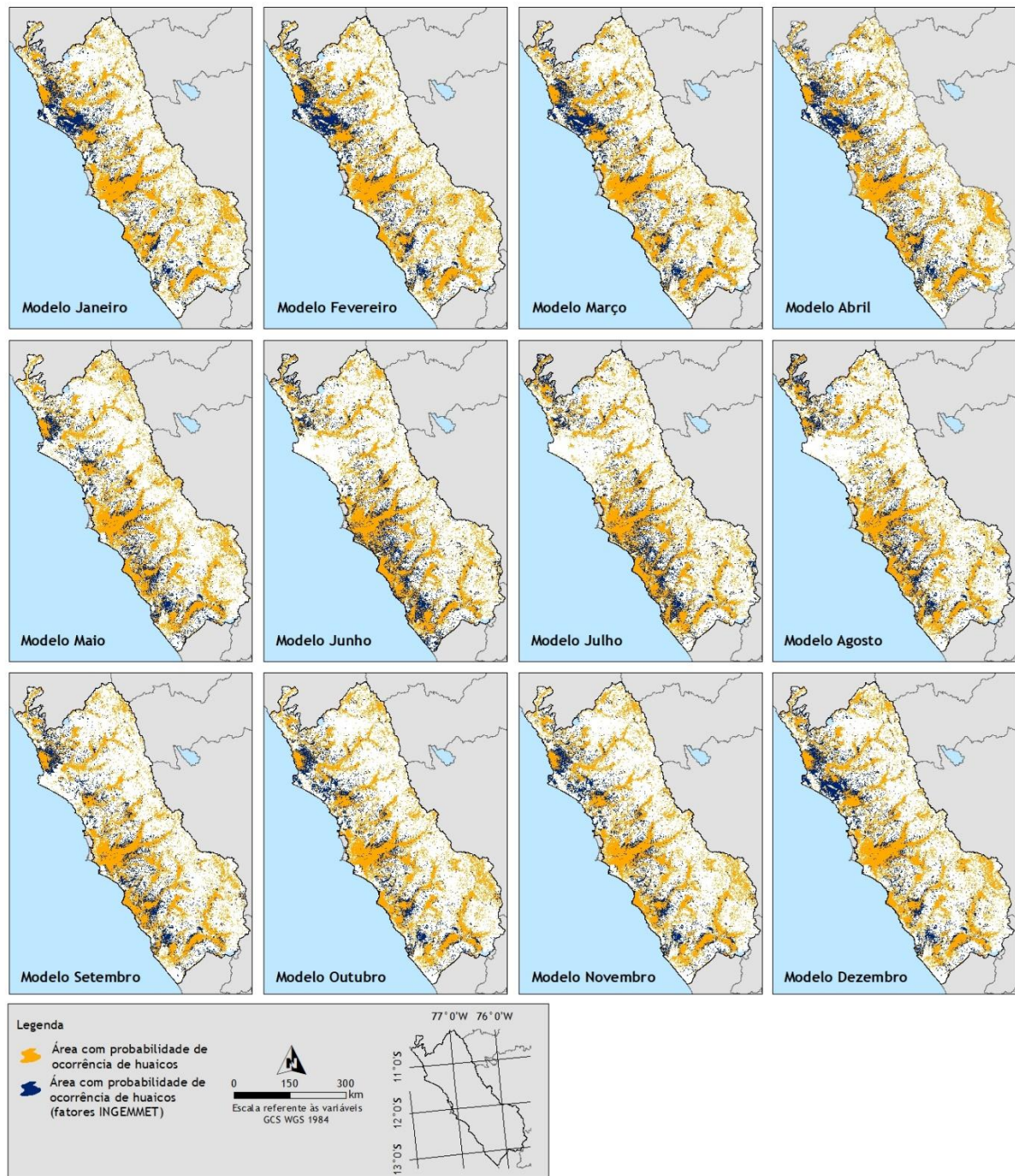


Figura 32. Resultados de suscetibilidade considerando os modelos mensais de Máquina de Vetores de Suporte.
 Fonte: a autora (2026).

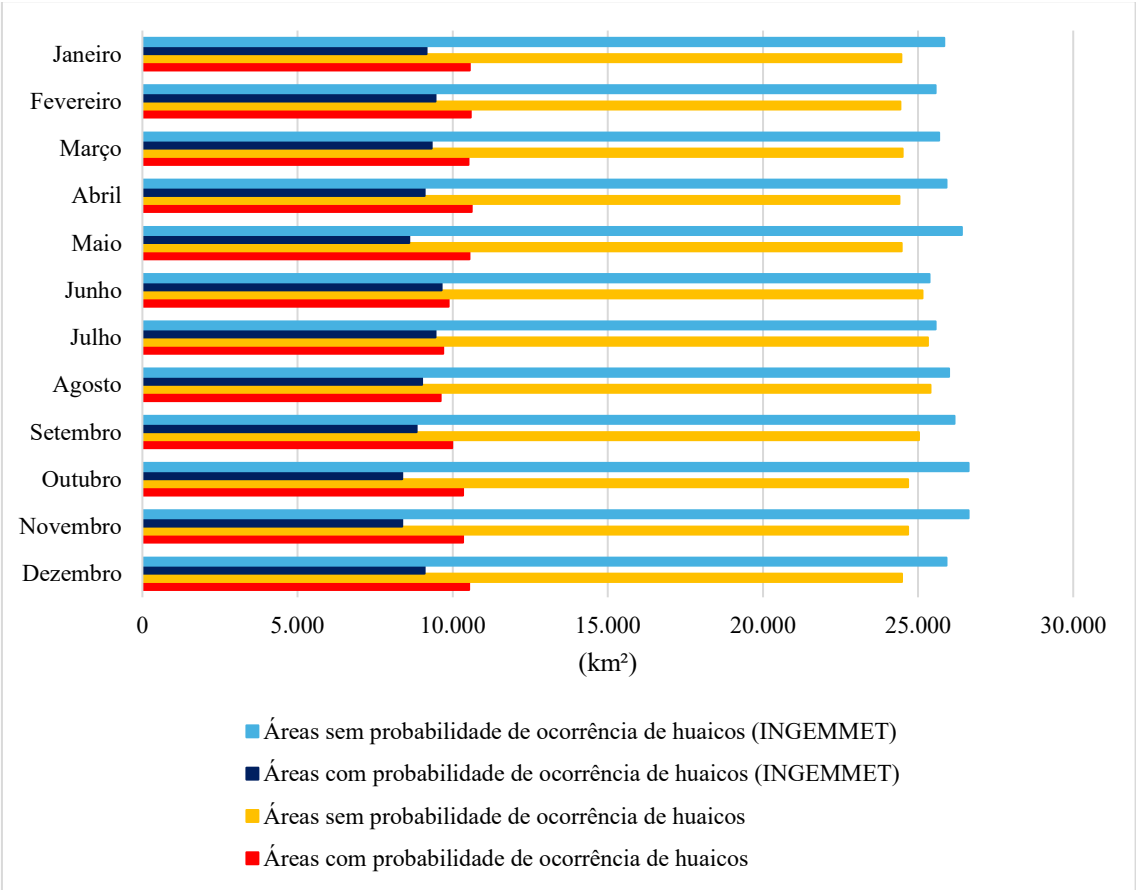


Figura 33. Áreas com e sem suscetibilidade de ocorrência de huaicos, segundo os modelos de Máquina de Vetores de Suporte. Fonte: a autora (2026).

Tabela 20. Coincidência entre modelos mensais Máquina de Vetores de Suporte – Variáveis segundo o INGEMMET (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	99	99	99	98	95	96	96	98	99	99	99
Fev	100	100	99	99	98	96	96	97	98	98	98	99
Mar	99	99	100	98	98	95	96	96	97	98	98	99
Abr	99	99	99	100	98	95	96	96	97	98	98	98
Maio	97	96	97	97	100	96	96	97	98	99	99	97
Jun	96	96	96	12	98	100	99	99	98	97	97	96
Jul	96	96	96	96	98	99	100	99	98	97	97	96
Ago	96	96	96	96	98	98	99	100	98	97	97	96
Set	97	97	97	97	99	97	97	98	100	99	99	98
Out	97	97	97	97	99	96	96	96	98	100	100	98
Nov	97	97	97	97	99	96	96	96	98	100	100	98
Dez	99	98	98	98	98	96	96	97	98	99	99	100

Fonte: a autora (2026).

Tabela 21. Coincidência entre modelos mensais Máquina de Vetores de Suporte – Variáveis segundo a bibliografia (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	99	99	99	98	98	98	98	98	98	98	99
Fev	99	100	99	99	98	98	98	98	98	98	98	99
Mar	99	99	100	99	98	98	98	98	98	98	98	99
Abr	99	99	99	100	98	98	98	98	98	99	99	99
Maio	98	98	98	98	100	98	98	98	99	99	99	98
Jun	97	97	97	97	97	100	99	99	98	97	97	97
Jul	97	97	97	97	97	99	100	99	98	97	97	97
Ago	97	97	97	97	97	99	99	100	98	97	97	97
Set	98	97	97	98	98	98	98	99	100	99	99	98
Out	98	98	98	98	99	98	98	98	99	100	100	99
Nov	98	98	98	98	99	98	98	98	99	100	100	99
Dez	99	99	99	99	98	98	98	98	99	99	99	100

Fonte: a autora (2026).

Com relação às variáveis que mais contribuem para os modelos, na Figura 34 observa-se que a altitude é o fator que mais contribui para os modelos mensais que consideram as variáveis de acordo com a bibliografia, com média de 27% de contribuição. A segunda e a terceira variáveis com as participações mais significativas são a distância a partir de áreas urbanizadas e a distância a partir de cursos hídricos, com 18% e 16% de participação média. As demais variáveis apontadas como deflagradoras do fenômeno pelo INGEMMET apresentaram contribuições entre 7% (declividade, precipitação média mensal) e 4% (geologia e geomorfologia), ainda superiores às de fatores como TRI, TWI, ecossistemas e aspecto.

Embora se observem áreas com suscetibilidade de ocorrência de huaicos até em altitudes elevadas, visualmente identificam-se, nas formas do relevo, os leitos de rios, sejam eles perenes ou efêmeros. Ainda, à distância das áreas urbanizadas, as áreas com suscetibilidade de ocorrência se concentram até 5.000 metros delas. Para a distância a partir dos cursos hídricos, as áreas com suscetibilidade de ocorrência se concentram até 2.500 metros. Novamente, observa-se que a precipitação média mensal apresenta comportamento bastante variável. E, justamente em pontos próximos de áreas com suscetibilidade de ocorrência, observam-se precipitações mais baixas.

Com relação ao comportamento comparado a distribuição normal, nos dados que mais contribuíram para os modelos, observa-se a maior sensibilidade. E, mais uma vez, ainda que somente o dado de precipitação média mensal tenha sido alterado a cada novo modelo mensal, essa alteração foi suficiente para modificar as respostas das demais variáveis em termos de contribuição e de variação na distribuição normal, sendo o dado de precipitação média mensal o que apresenta a maior variação em relação à distribuição normal.

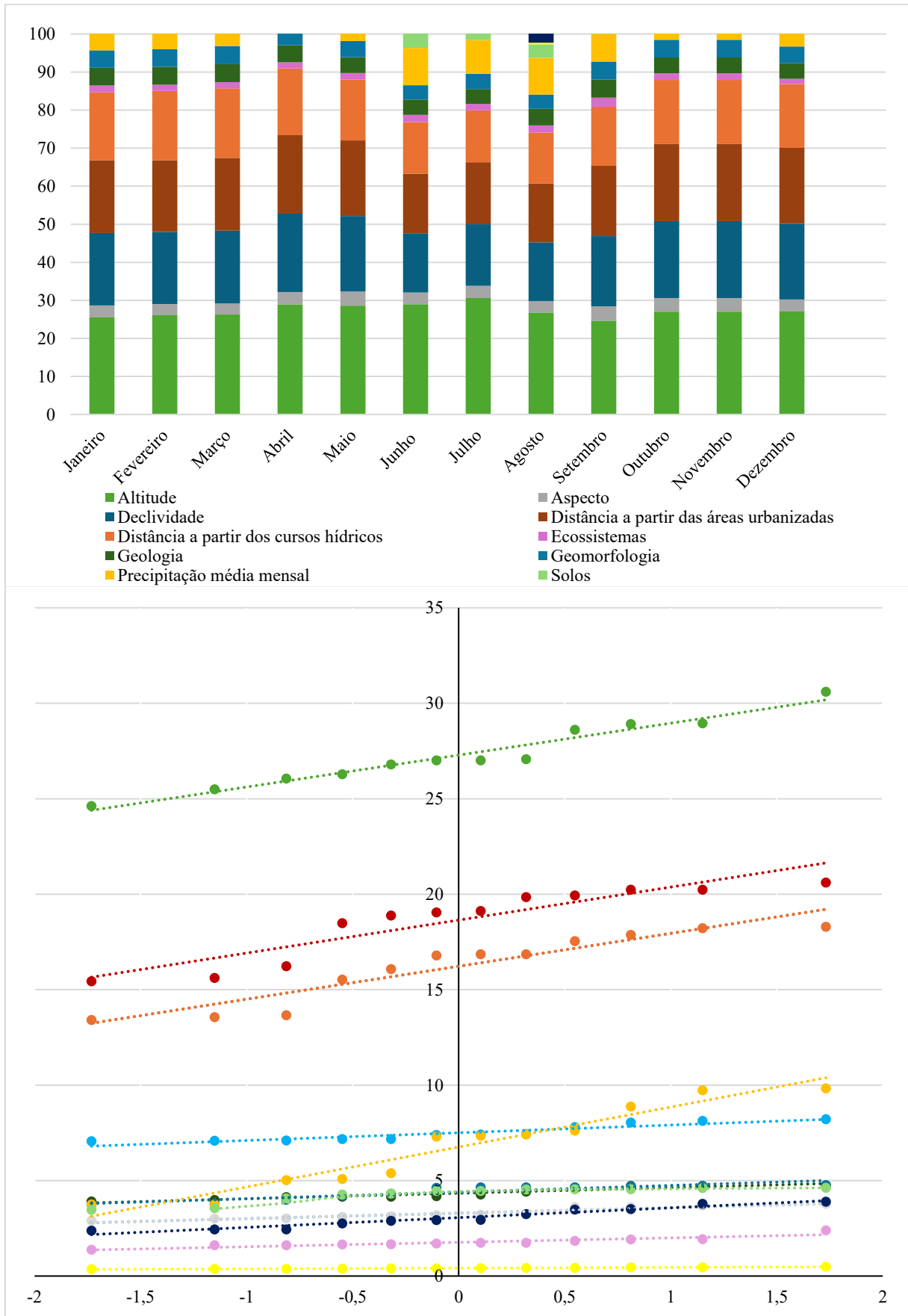


Figura 34. Contribuição das variáveis para cada modelo mensal de Máquina de Vetores de Suporte considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).

Nos modelos mensais que consideram apenas as variáveis indicadas pelo INGEMMET como relevantes para a ocorrência de huaicos, todos os meses apresentaram uma distribuição semelhante das contribuições das variáveis ambientais, com a mais significativa sendo a altitude. O segundo fator mais importante é a precipitação média mensal. Os demais fatores têm menor influência e contribuições semelhantes (Figura 35).

Embora a altitude tenha representado o maior valor de contribuição, com uma participação média de 36%, essa contribuição variou de 30% em setembro, a 43% em julho. Diferentemente dos modelos de floresta aleatória com variáveis do INGEMMET, que apresentaram a menor e a maior área com suscetibilidade de ocorrência de huaicos nos meses de menor e maior participação da variável que mais contribuiu para o modelo, essa relação não se repetiu nos modelos de máquina de vetores de suporte. Para a precipitação média mensal, a contribuição média foi de 27%, sendo a menor no mês de março (20%) e a maior na variável ocorrendo no mês de setembro (32%).

No entanto, esses meses possuem níveis de precipitação diferentes, sendo os meses de outono e inverno menos chuvosos historicamente, que os meses de março e setembro. E quanto aos meses mais chuvosos, o mês de setembro não corresponde aos meses mais chuvosos do ano, que são dezembro, janeiro, fevereiro e março; este último, sim, apresenta relação positiva com um aumento de áreas com suscetibilidade de ocorrência de huaicos. Ainda assim, a relação entre a participação nos modelos e os níveis de precipitação não é linear, sendo, então, a mudança espacial do nível de precipitação mais relevante.

Do comportamento pela distribuição normal, a precipitação média mensal e a altitude predominam nos modelos mensais. Novamente, destaca-se o impacto da precipitação média mensal nos modelos e nas demais variáveis. Neste caso, de avaliação somente com as variáveis do INGEMMET, o impacto ocorre sobre a altitude e a geologia, embora com uma participação inferior no modelo.

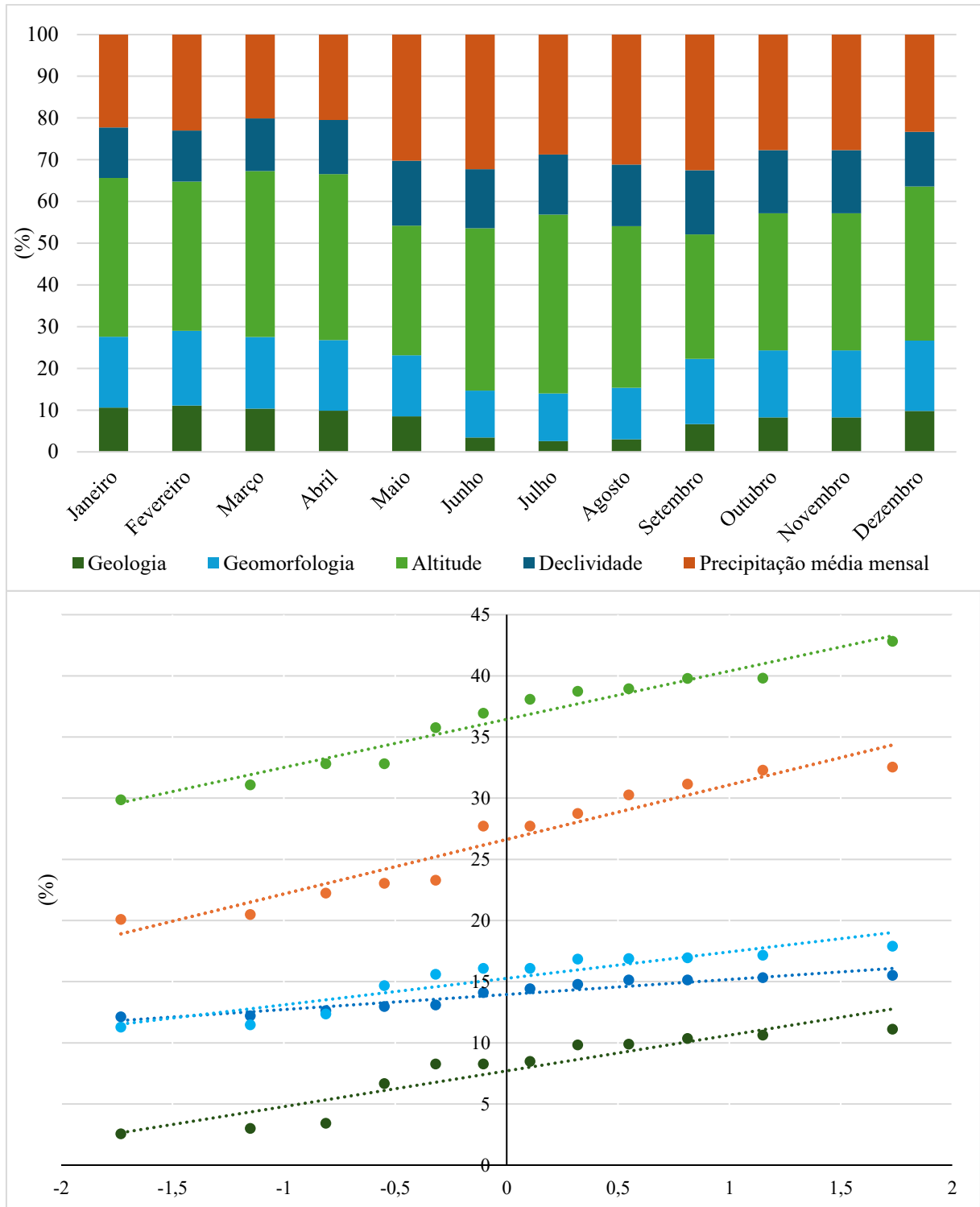


Figura 35. Contribuição das variáveis para cada modelo mensal de Máquina de Vetores de Suporte, considerando as variáveis ambientais indicadas pelo INGEMMET (a) e o comportamento de cada variável em relação a uma distribuição normal (b). Fonte: a autora (2026).

Quin et al. (2020) obtiveram os melhores resultados para a previsão de fluxo de detritos. Os autores ainda adotaram as bacias hidrográficas como unidades de estudo. Para cada bacia, também se adotou a classificação da suscetibilidade de ocorrência de muito baixa a muito alta. Tendo em vista a característica das áreas de estudo, formadas por bacias hidrográficas de

rios efêmeros, materiais fragmentáveis, desprovidas de vegetação de cobertura, seria interessante a adoção de bacias hidrográficas para a avaliação. Porém, no formato atual do trabalho, podem existir diferentes faixas de suscetibilidade de ocorrência de huaicos em uma mesma bacia.

Li et al. (2023) obtiveram como fatores que mais influenciam na ocorrência de huaicos, o relevo relativo da bacia e o comprimento do canal principal. Assim como neste trabalho, a altitude também se mostrou o fator mais influente. Mas não a altitude por si só, a altitude como forma do relevo, que associa o fluxo as áreas de leitos de rios, sejam eles perenes ou efêmeros. Li et al. (2023) utilizam uma abordagem um tanto distinta, considerando apenas a densidade de drenagem, o relevo relativo da bacia, a proporção de leito móvel e o comprimento do canal principal. No entanto, além das informações sobre a localização das áreas com ocorrência de fluxos de detritos, os autores possuem dados referentes à origem dos materiais soltos na área de estudo. Consequentemente, os resultados foram gerados em termos da escala dos eventos — uma possibilidade improvável para a maioria dos estudos devido à falta de dados.

Gao et al. (2021) utilizaram altitude, declividade, orientação, curvatura em planta, curvatura em perfil, TWI, NDVI, litologia, distância a rios, distância a estradas e distância a falhas. Com essas variáveis, o resultado em termos de qualidade AUC foi de 0,99, um resultado excelente e muito superior aos encontrados aqui. Isso se deve à abordagem diferenciada adotada pelos autores em relação à inserção de dados no modelo, incluindo apenas informações que contribuiriam positivamente para o fenômeno. Gao et al. (2021) constaram que a topografia da encosta do vale associada à baixa altitude contribuiu mais para a formação e o desenvolvimento de fluxos de detritos, corroborando para os resultados aqui obtidos, assim como Chen et al. (2020).

4.2.3 MaxEnt

Na Tabela 22 apresenta-se a qualidade dos modelos MaxEnt. Todos os modelos apresentam qualidade apropriada no “AUC de Teste”, resultando em valores superior a 0,7. Conforme a classificação de Yesilnacar (2005), ambos os conjuntos de modelos gerados (INGEMMET e bibliografia) foram classificados como excelentes.

Para o índice Kappa, os dois conjuntos de modelos gerados (INGEMMET e bibliografia) apresentaram desempenho muito excelente. Ambos os conjuntos de dados apresentam respostas similares quanto ao índice Kappa, sendo estes os melhores resultados

obtidos, em comparação aos dados dos algoritmos floresta aleatória e máquinas de vetores de suporte.

Tabela 22. Qualidade dos modelos Maxent

Mês	INGEMMET		Variáveis segundo a bibliografia	
	AUC	Kappa	AUC	Kappa
Janeiro	0,92	0,96	0,93	0,96
Fevereiro	0,92	0,97	0,93	0,95
Março	0,91	0,97	0,93	0,94
Abril	0,90	0,94	0,93	0,95
Maio	0,92	0,97	0,93	0,96
Junho	0,91	0,96	0,93	0,93
Julho	0,91	0,96	0,93	0,94
Agosto	0,91	0,96	0,93	0,95
Setembro	0,91	0,97	0,93	0,95
Outubro	0,91	0,96	0,93	0,95
Novembro	0,91	0,96	0,93	0,97
Dezembro	0,92	0,97	0,93	0,95

Fonte: a autora (2026).

As Figuras 36 e 37 apresentam os resultados dos modelos, reclassificados para representar as áreas com suscetibilidade de ocorrência de huaicos. Em média, 39% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos quando considerados os modelos que utilizam todas as variáveis. Há pouca diferença entre os valores mensais, o que sugere que a precipitação não influencia significativamente os resultados.

Ao analisar os modelos gerados com base nas variáveis do INGEMMET identificadas como fatores de ocorrência de huaicos, em média, 26% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos.

O mês de fevereiro apresentou a menor área com suscetibilidade de ocorrência de huaicos no conjunto de modelos do INGEMMET, com 6.565,81 km² (18,75%). O mês de abril retornou a maior área com suscetibilidade, ainda para as variáveis do INGEMMET, 33,07% (11.579,20 km²). Para os modelos elaborados com as variáveis, conforme a bibliografia, as áreas atingem, em novembro, um mínimo de 11.640,40 km² (33,25%). O mês de outubro apresentou a maior área, correspondente a 43,07% do departamento de Lima (15.079,00 km²).

Espacialmente, os modelos apresentam abrangência por toda a porção costeira, nos modelos do INGEMMET. Com os modelos com todas as variáveis da bibliografia, amplia-se ainda mais a abrangência das áreas com suscetibilidade de ocorrência na porção costeira. Ainda assim, destaca-se que, nas porções de maiores níveis de precipitação média mensal, não se observa aumento no nível de suscetibilidade de ocorrência.

Em relação à coincidência entre os modelos mensais, a Tabela 23 e Tabela 24 contém as matrizes de coincidência das áreas com suscetibilidade de ocorrência para cada modelo,

tomando os meses na horizontal como referência. Em média, há uma coincidência de 95% entre os modelos mensais que consideram as variáveis do INGEMMET e de 94% entre os modelos que consideram as variáveis conforme a bibliografia. Quanto maior a correspondência entre os modelos ao longo do ano, maior a chance de que, mesmo com variações na precipitação, as áreas com suscetibilidade de ocorrência permaneçam basicamente as mesmas.

Embora os resultados sejam de uma alta convergência entre os dois conjuntos de dados mensais, destaca-se que a convergência entre os modelos com o algoritmo de vetores de suporte de máquina obteve um resultado ainda superior.

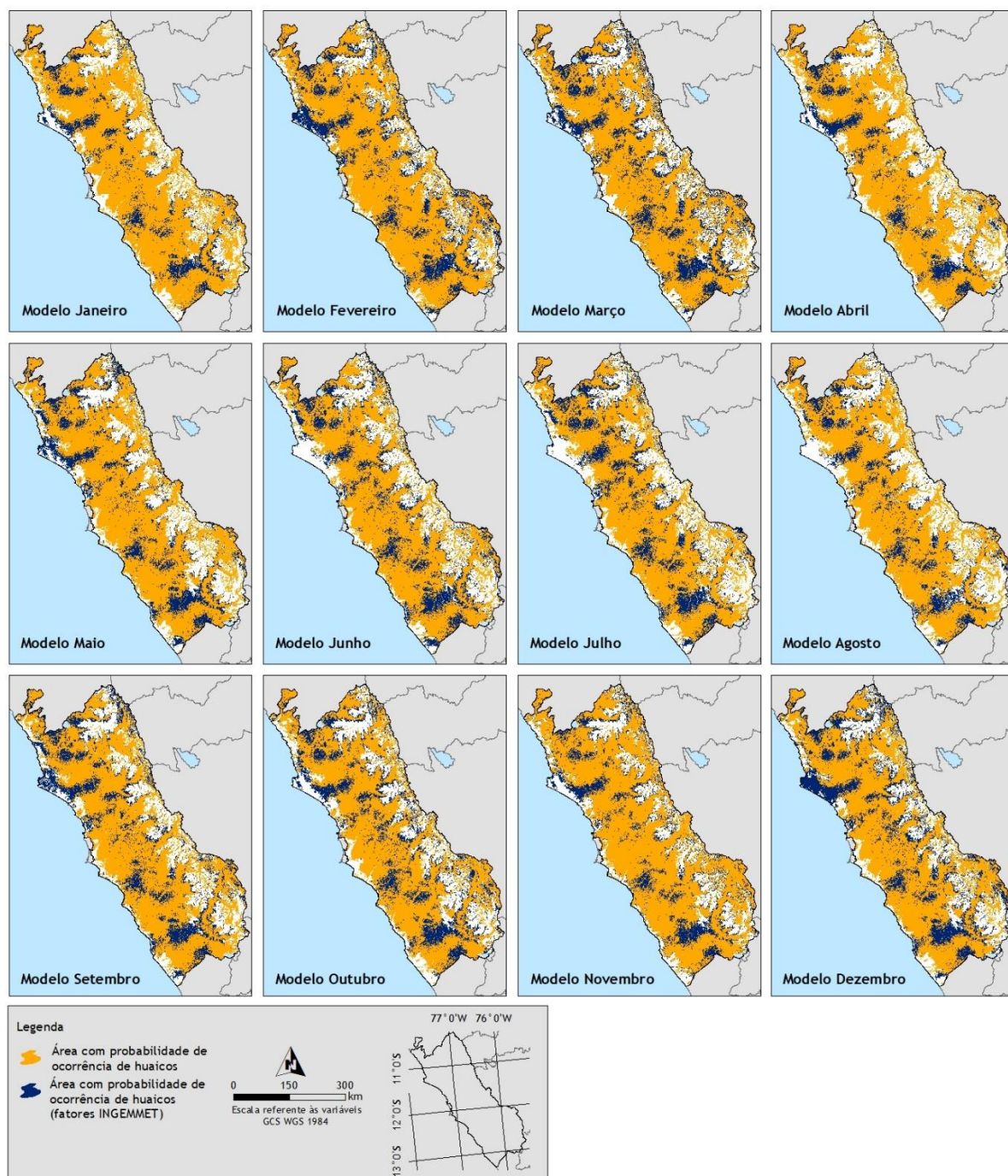


Figura 36. Resultados de suscetibilidade considerando os modelos mensais Maxent. Fonte: a autora (2026).

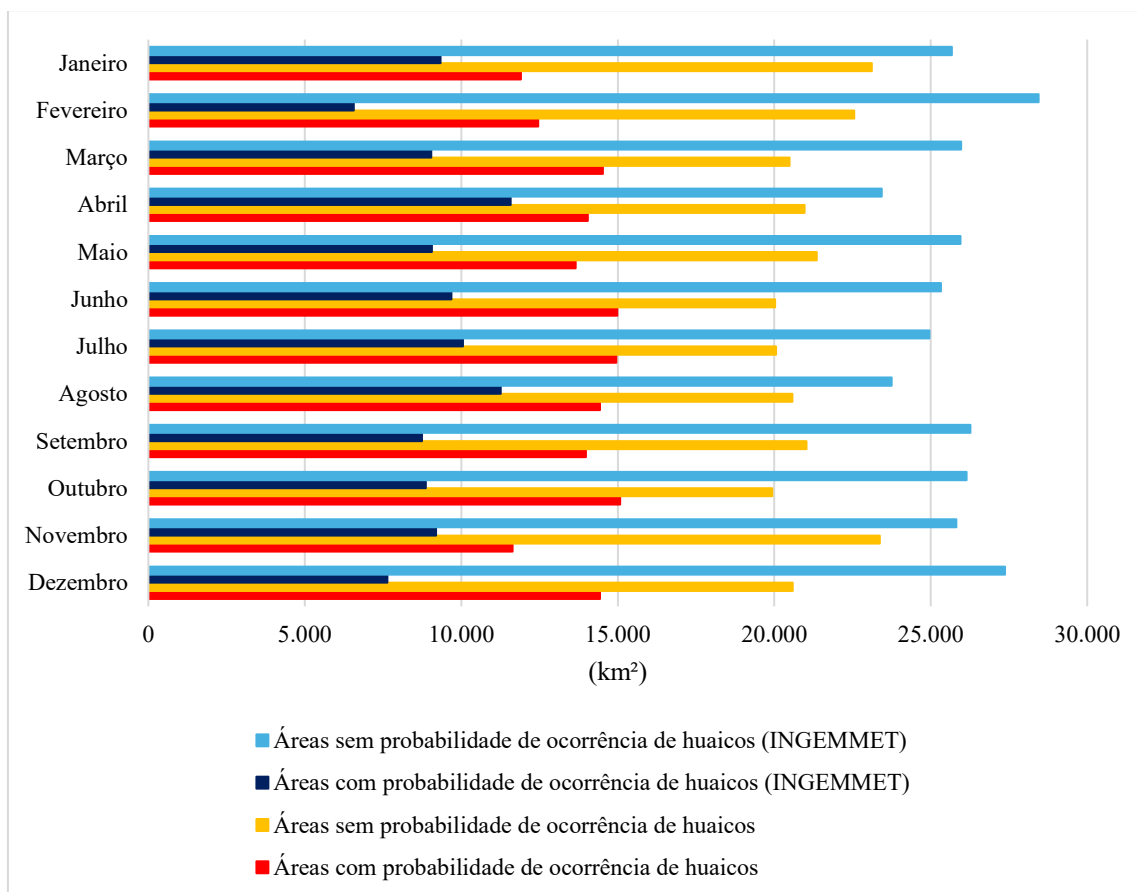


Figura 37. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos Maxent. Fonte: a autora (2026).

Tabela 23. Coincidência entre modelos mensais Maxent – Variáveis segundo o INGEMMET (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	90	96	100	95	96	97	99	96	96	97	93
Fev	100	100	99	100	99	100	100	100	100	98	99	98
Mar	97	91	100	99	95	96	96	98	95	94	96	93
Abr	91	82	89	100	89	90	90	93	88	89	90	86
Maio	96	90	95	99	100	95	95	98	96	96	97	94
Jun	95	89	93	97	93	100	97	99	94	93	94	90
Jul	94	88	93	96	92	96	100	99	92	91	92	89
Ago	91	83	90	95	89	93	94	100	89	88	90	86
Set	98	92	96	99	97	97	97	99	100	96	97	94
Out	97	90	95	99	97	96	95	97	96	100	98	94
Nov	97	90	95	99	96	96	95	97	95	97	100	93
Dez	99	95	98	100	99	98	98	99	98	98	99	100

Fonte: a autora (2026).

Tabela 24. Coincidência entre modelos mensais Maxent – Variáveis segundo a bibliografia (%)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	96	99	98	95	98	98	98	97	98	92	98
Fev	94	100	99	98	95	98	98	98	97	97	91	97
Mar	88	90	100	94	91	96	95	95	93	95	85	94
Abr	89	91	96	100	93	96	96	95	95	97	87	95
Mai	88	90	94	95	100	96	95	94	94	96	88	96
Jun	85	87	94	92	90	100	94	93	92	94	83	92
Jul	85	87	93	92	89	95	100	93	92	93	83	91
Ago	88	90	96	94	91	96	96	100	94	95	85	93
Set	89	91	96	95	93	97	96	96	100	97	87	95
Out	85	86	93	92	90	93	92	92	92	100	84	93
Nov	93	94	97	97	96	97	97	97	97	98	100	97
Dez	87	88	95	94	92	94	93	93	93	96	85	100

Fonte: a autora (2026).

Com relação às variáveis que mais contribuem para os modelos, na Figura 38, observa-se que a geologia é o fator que mais contribui para os modelos mensais, seguida pela altitude e pela precipitação média mensal. Este é um comportamento diferente dos observados nos demais algoritmos. Embora se observe um aumento significativo na participação da geologia nos modelos com o algoritmo Maxent, a geomorfologia apresenta uma participação média de 3%. Aspecto, ecossistemas, TRI e TWI apresentam participações inferiores a 1%.

Das variáveis que desencadeiam o fenômeno, segundo o INGEMMET, a declividade, ainda não citada, possui uma participação média de 2%.

Assim como os demais algoritmos, foram classificadas áreas com suscetibilidade de ocorrência de huaicos em altitudes elevadas, mas onde a configuração do relevo se apresenta como leito de rios, sejam eles perenes ou efêmeros, e ainda em planícies de inundação. As formações geológicas que representam a suscetibilidade de ocorrência incluem especialmente a Super unidade Santa Rosa, a classificação Cretáceo superior inferior vulcânico-sedimentar e a classificação Paleogeno-Neogeno, vulcânico-sedimentar.

A Super unidade Santa Rosa possui afloramentos de rochas ígneas intrusivas suscetíveis a ocorrência de erosão de encostas, deslizamentos de terra e quedas de rochas (INGEMMET, 2020). A classificação Cretáceo superior inferior vulcânico-sedimentar possui solos siltosos-arenosos, que são facilmente erodidos, quando ocorrem chuvas excepcionais, sendo assim propícias para gerar fluxos de detritos canalizados e não canalizados (INGEMMET, 2014). A classificação Paleogeno-Neogeno, vulcânico-sedimentar, corresponde a rochas piroclásticas de composição andesítica, abundantes lavas andesíticas e ignimbritos dacíticos, altamente fraturados, com espaçamento de fraturas muito próximo (0,05–0,03 m) a amplo (1,0–5,0 mm), e altamente intemperizados (INGEMMET, 2022).

Quanto à distribuição normal, embora a variável de maior participação tenha mudado,

o comportamento geral da relação entre essa variável e a precipitação média mensal permanece.

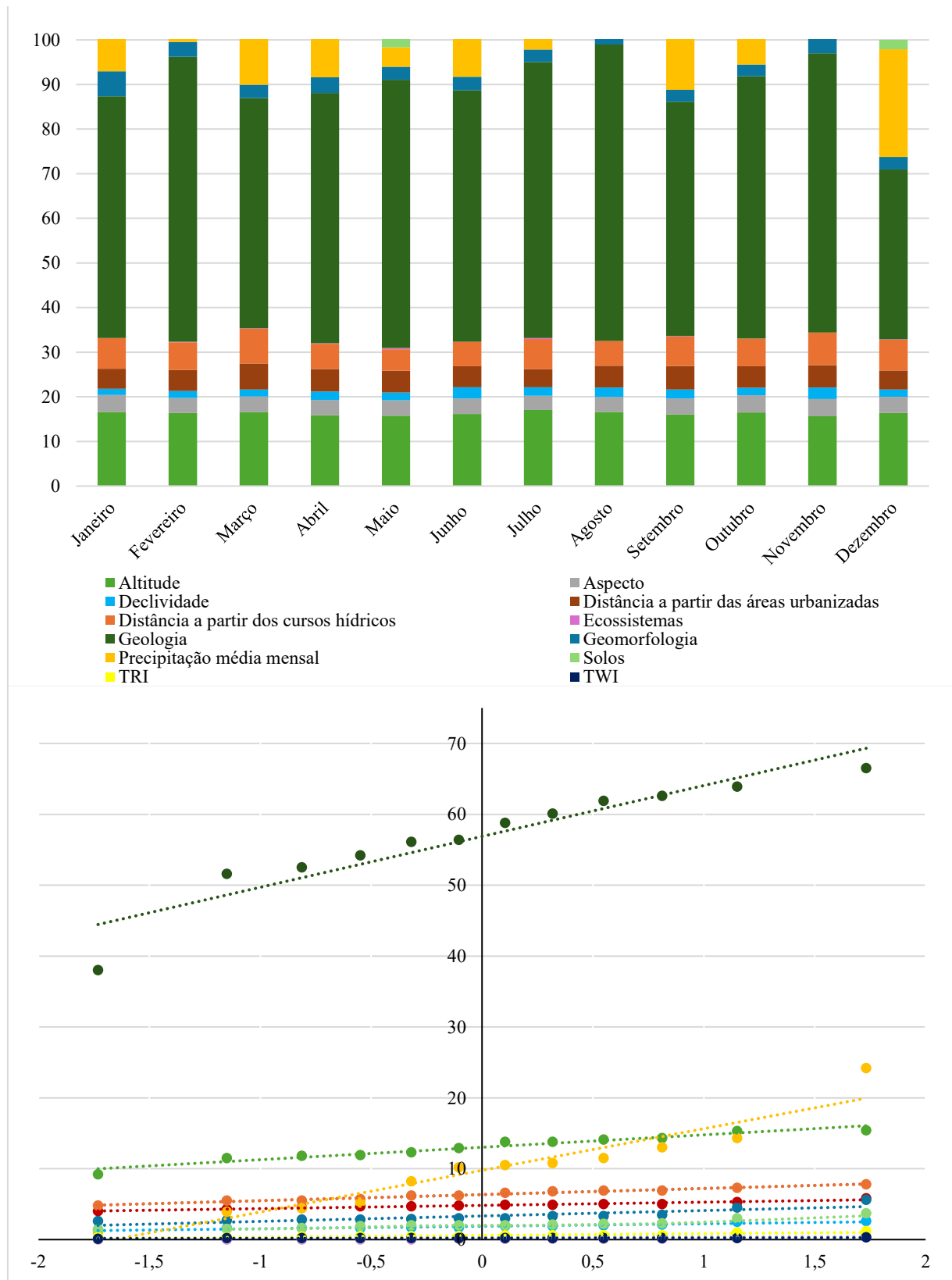


Figura 38. Contribuição das variáveis para cada modelo mensal do Maxent, considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada variável em relação a uma distribuição normal. Fonte: a autora (2026).

Nos modelos mensais que consideram apenas as variáveis indicadas pelo INGEMMET como relevantes para a ocorrência de huaicos, observa-se maior variação nas contribuições das variáveis ambientais. Ainda assim, como os modelos que consideram as variáveis conforme a bibliografia do algoritmo Maxent, neste conjunto, a contribuição mais significativa continua sendo a geologia, que explica as classificações mencionadas anteriormente. O segundo fator mais importante é a altitude, que também segue as faixas descritas na caracterização dos resultados do modelo anterior. A participação da precipitação média mensal está empatada com a altitude, em termos médios (Figura 39).

Embora a geologia tenha representado o maior valor de contribuição, com uma participação média de 58%, essa contribuição variou de 3% em agosto, a 72% em maio. Para a precipitação média mensal, a contribuição média foi de 1%, sendo a menor no mês de fevereiro (1%) e a maior na variável ocorrendo no mês de junho (23%).

Mesmo existindo uma aparente relação inversa da participação no modelo do menor nível de participação da precipitação em um dos meses com maior nível de precipitação e uma participação elevada de precipitação em um mês de baixa precipitação, a relação não se mantém válida para os demais meses.

Do comportamento pela distribuição normal, a precipitação média mensal e a geologia predominam nos modelos mensais. Novamente, destaca-se o impacto da precipitação média mensal nos modelos e nas demais variáveis, neste caso, na geologia.

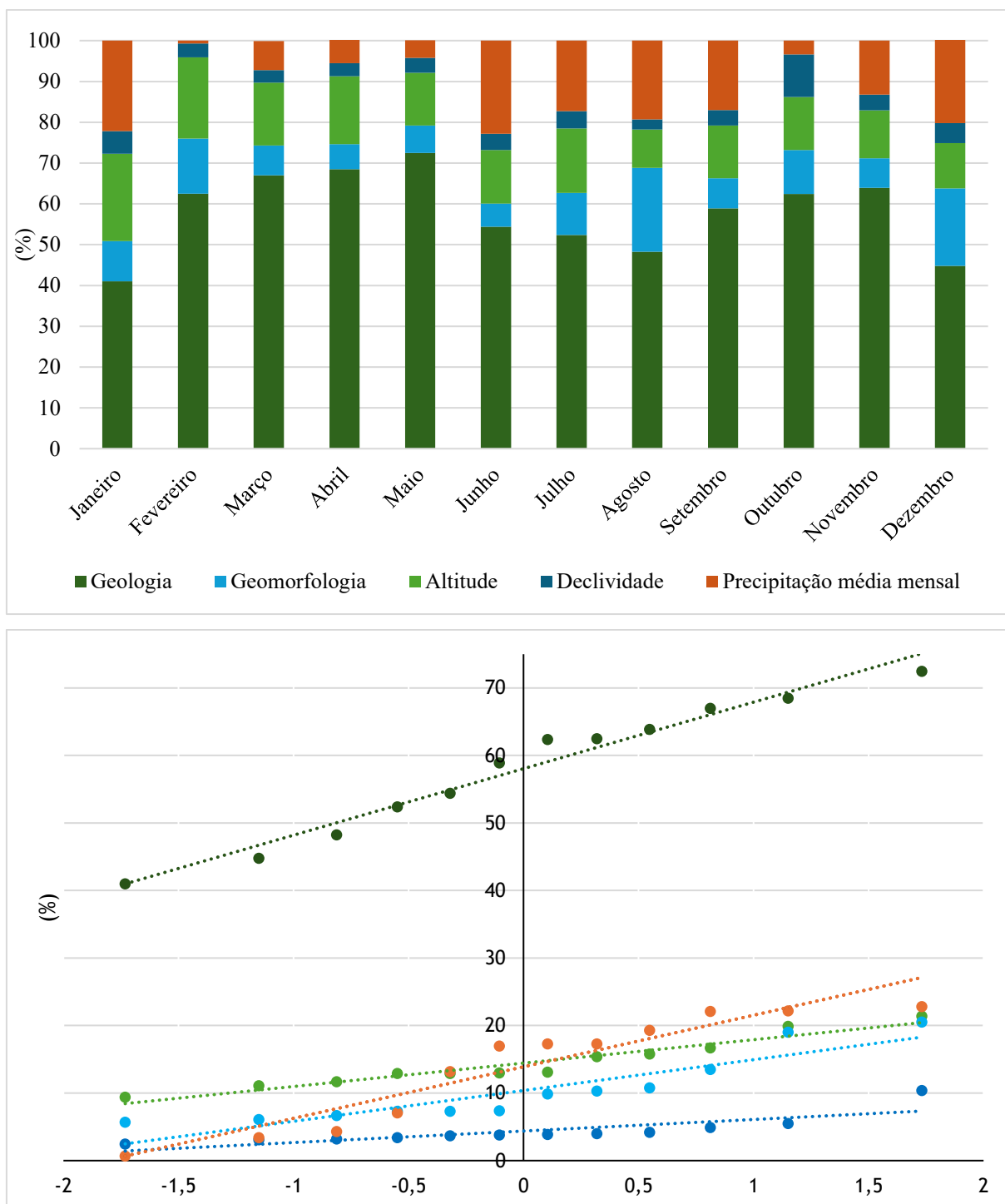


Figura 39. Contribuição das variáveis para cada modelo mensal do Maxent, considerando as variáveis ambientais indicadas pelo INGEMMET (a) e o comportamento de cada variável em relação a uma distribuição normal (b). Fonte: a autora (2026).

O comportamento da geologia e da precipitação média mensal, em relação à distribuição normal, é errático. Kornejady, Ownegh e Bahremmand (2017), assim como Phillips (2017), alertam sobre a interpretação das contribuições dos fatores nos modelos, já que o MaxEnt pode encontrar a melhor solução por caminhos distintos, o que pode resultar em contribuições distintas das variáveis. Em outras palavras, a contribuição não pode ser

considerada um indicativo da importância da variável ambiental.

Embora a identificação das áreas onde ocorrem huaicos não corresponda exatamente à de áreas com deslizamentos ou inundações, Rusk et al. (2022) avaliaram ambos os riscos e adotaram variáveis semelhantes às deste trabalho. Para deslizamentos, fatores relacionados ao relevo e à geomorfologia contribuíram de forma significativa. No entanto, ao contrário deste estudo, para Rusk et al. (2022), a precipitação teve contribuição significativa no modelo para ambos os tipos de risco: aproximadamente 35% para inundações e 25% para deslizamentos. Os resultados de deslizamentos obtidos por Javidan et al. (2021) apontam na mesma direção, onde fatores geomorfológicos relevantes como declividade e altitude, isoladamente, não interferiram no modelo.

Embora a ocorrência de huaicos não dependa diretamente da distância aos cursos d'água, há uma concentração da suscetibilidade de ocorrência próxima aos rios nos modelos que utilizaram todas as variáveis, o que difere dos resultados de Javidan et al. (2021). Eles identificaram a distância aos cursos d'água como um fator preditivo essencial, devido à sua importância na velocidade e na magnitude das inundações, considerando a alta concentração de fluxo na rede de drenagem, o que aumenta a ocorrência de inundações nas proximidades dos riachos. No entanto, essa relação é compreensível, pois algumas "quebradas" pertencem a bacias hidrográficas de cursos d'água perenes.

Chen et al. (2017) também indicaram uma relação entre deslizamentos e a proximidade de cursos d'água e estradas em porções maiores do terreno. Os autores também destacaram a distância das estradas como um fator antrópico que modifica a estabilidade natural das áreas estudadas e ameaça a infraestrutura, reduzindo significativamente a qualidade do modelo. Embora fatores antrópicos não influenciem diretamente a ocorrência de huaicos, os impactos são intensificados pela presença de áreas urbanizadas e de infraestrutura.

De forma semelhante aos resultados de Javidan et al. (2021), há menor incidência de huaicos em áreas de alta altitude. Essa constatação baseou-se nos resultados das curvas de resposta à elevação, que demonstram que huaicos, assim como inundações, ocorrem em áreas baixas e planícies. Vale destacar que o escoamento das bacias que formam os huaicos normalmente converge em um único canal, que pode atingir uma área urbanizada ou seguir por uma planície de inundação até o oceano e, eventualmente, alcançar uma área urbana.

Javidan et al. (2021) constataram que níveis mais altos de precipitação aumentam os deslizamentos (650–690 mm anuais), enquanto níveis mais baixos provocam erosão. Para huaicos, um aumento da precipitação deveria resultar em uma maior área com suscetibilidade de ocorrência, mas esse aumento não se refletiu linearmente nos modelos. De modo geral, a

precipitação não contribuiu significativamente para o fenômeno. Portanto, a constituição dos modelos não permite inferir os níveis de precipitação que desencadeiam os huaicos.

Por outro lado, Rusk et al. (2022) apontam que as mudanças climáticas podem alterar as respostas dos modelos, aumentando o risco de desastres. Um passo importante para a prevenção e mitigação dos impactos dos desastres é a identificação da vulnerabilidade social das populações afetadas.

4.2.4 Redes Neurais Artificiais

A qualidade dos modelos de Redes Neurais Artificiais é apresentada na Tabela 25. Todos os modelos apresentam qualidade apropriada no “AUC de Teste”, com valores acima de 0,7, e foram classificados como muito bons, segundo a escala apresentada por Yesilnacar (2005). Relembra-se que as faixas de AUC de 0,5–0,6, 0,6–0,7, 0,7–0,8, 0,8–0,9 e 0,9–1 classificam os modelos como fraco, mediano, bom, muito bom e excelente, respectivamente.

Já o índice Kappa apresenta cinco classes — $k < 0,4$; $0,4 < k < 0,55$; $0,55 < k < 0,85$; $0,85 < k < 0,99$, que indicam uma validação fraca, mediana, boa, excelente e muito excelente (Cohen, 1960). No índice Kappa, os modelos gerados com todas as variáveis mencionadas na bibliografia apresentaram desempenho excelente. Os modelos gerados com base apenas nas variáveis ambientais descritas pelo INGEMMET como influentes na formação de huaicos apresentaram índice Kappa classificado como muito excelentes.

São resultados interessantes de índice Kappa, similares aos modelos com o algoritmo máquinas de vetores de suporte, mas não tão bons quanto os modelos com o Maxent.

Tabela 25. Qualidade dos modelos Redes Neurais Artificiais

Mês	INGEMMET		Variáveis segundo a bibliografia	
	AUC	Kappa	AUC	Kappa
Janeiro	0,84	0,89	0,89	0,74
Fevereiro	0,84	0,88	0,89	0,76
Março	0,83	0,87	0,88	0,75
Abril	0,83	0,89	0,89	0,76
Maio	0,85	0,93	0,90	0,76
Junho	0,88	0,86	0,90	0,73
Julho	0,88	0,85	0,90	0,72
Agosto	0,88	0,87	0,89	0,72
Setembro	0,88	0,92	0,90	0,76
Outubro	0,88	0,98	0,90	0,77
Novembro	0,88	0,98	0,90	0,77
Dezembro	0,88	0,98	0,90	0,78

Fonte: a autora (2026).

As Figuras 40 e 41 apresentam os resultados dos modelos, reclassificados para representar as áreas com suscetibilidade de ocorrência de huaicos. Em média, 25% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos quando considerados os modelos que utilizam todas as variáveis. Ao analisar os modelos gerados com base nas variáveis do INGEMMET identificadas como fatores de ocorrência de huaicos, a qualidade dos modelos foi melhor, e as áreas com suscetibilidade de ocorrência foram maiores. Em média, 24% do Departamento de Lima apresenta suscetibilidade de ocorrência de huaicos.

Os meses de outubro, novembro e dezembro contêm as menores áreas, 21,45% das áreas do departamento de Lima (7.511,26 km²), enquanto o mês de março possui a maior área para o conjunto de modelos com as variáveis do INGEMMET (9.299,35 km²; 26,56%). Com os modelos das variáveis a partir da bibliografia, o mês de dezembro retornou a menor área dos modelos mensais, 23,69% (8.294,30 km²), enquanto a maior área está associada ao mês de agosto (9.150,88 km²; 26,14%).

Em relação à coincidência entre os modelos mensais, a Tabela 26 e Tabela 27 contêm a matriz de coincidência das áreas com suscetibilidade de ocorrência para cada modelo, tomando os meses na horizontal como referência. Em média, há uma coincidência de 83% entre os modelos mensais quando considerados os do INGEMMET. Para as variáveis da bibliografia, a correspondência média é de 72%. O impacto da mudança da precipitação média mensal é significativo nos modelos de redes neurais artificiais, para alteração dos resultados, que são inferiores em concordância em comparação aos modelos do Maxent e de máquinas de vetores de suporte.

No mapa, observam-se resultados similares aos modelos de floresta aleatória e máquina de vetores de suporte, com relação ao posicionamento mais próximo aos cursos hídricos, para os modelos com as variáveis da bibliografia e áreas de suscetibilidade de ocorrência nas regiões costeiras, pelos modelos com as variáveis do INGEMMET

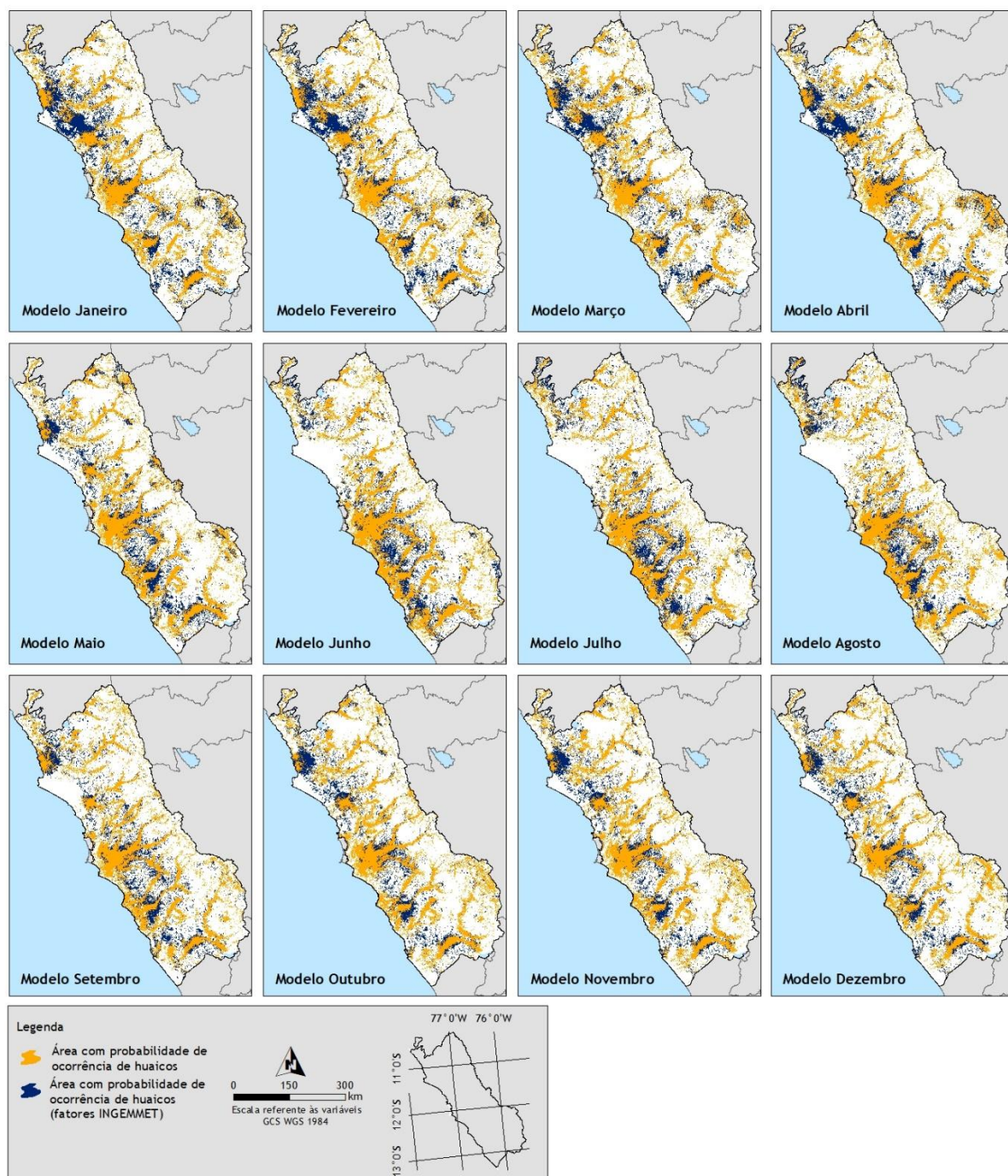


Figura 40. Resultados de suscetibilidade considerando os modelos mensais Redes Neurais Artificiais. Fonte: a autora (2026).

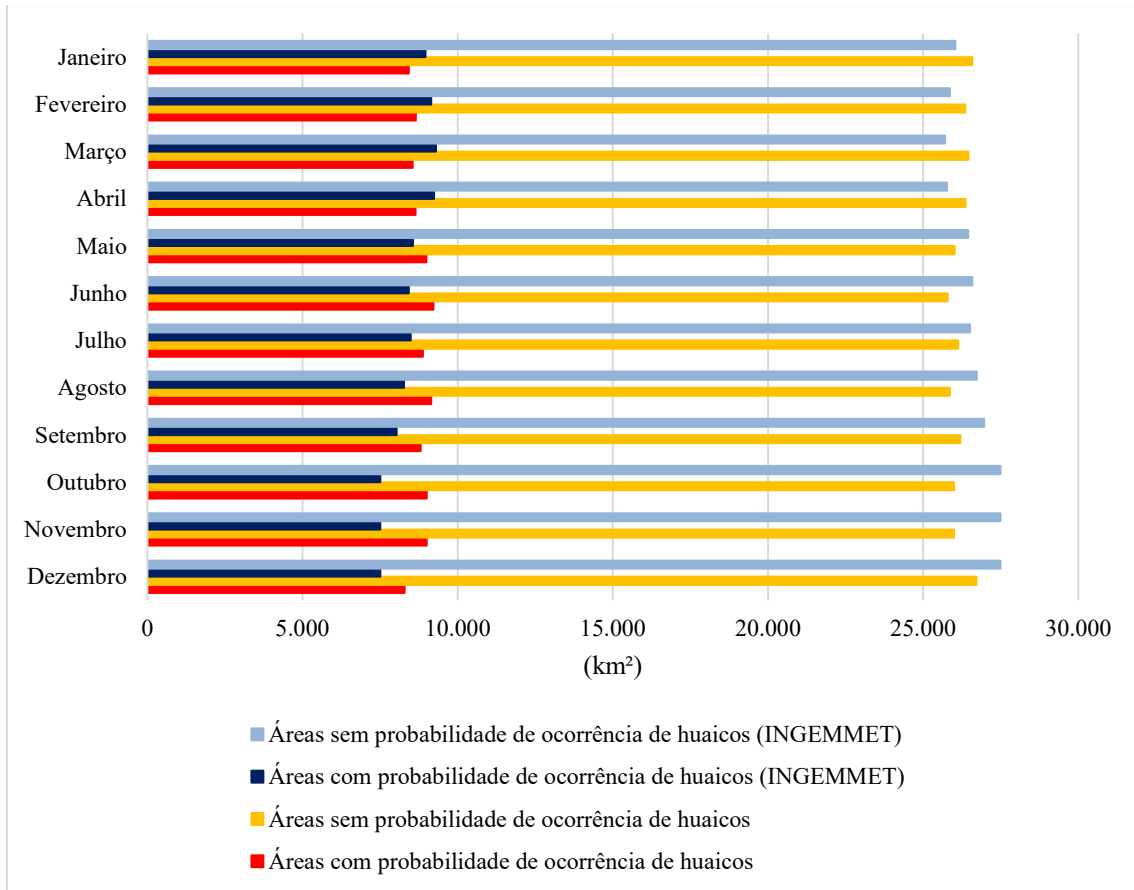


Figura 41. Áreas com e sem suscetibilidade de ocorrência de huaicos pelos modelos Redes Neurais Artificiais.
Fonte: a autora (2026).

Tabela 26. Coincidência entre modelos mensais Redes Neurais Artificiais – Variáveis segundo o INGEMMET (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	99	98	98	98	96	95	96	97	98	98	98
Fev	99	100	98	98	98	96	96	96	97	98	98	98
Mar	98	99	100	98	98	96	96	96	97	98	98	98
Abr	98	98	98	100	96	96	95	96	97	99	99	99
Maio	98	98	98	98	100	96	95	96	97	99	99	99
Jun	95	95	94	95	96	100	99	98	98	97	97	97
Jul	95	95	94	94	95	99	100	99	98	97	97	97
Ago	95	95	94	95	95	98	99	100	98	97	97	97
Set	96	96	95	96	96	97	97	98	100	98	98	98
Out	96	96	96	96	96	96	95	96	98	100	100	100
Nov	96	96	96	96	96	96	95	96	98	100	100	100
Dez	96	96	96	96	96	96	95	96	98	100	100	100

Fonte: a autora (2026).

Tabela 27. Coincidência entre modelos mensais Redes Neurais Artificiais – Variáveis segundo a bibliografia (%)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	85	84	84	75	75	70	71	76	78	78	84
Fev	88	100	86	82	74	75	72	75	76	77	77	84
Mar	85	85	100	83	74	74	69	75	73	76	76	83
Abr	86	82	84	100	79	75	71	71	76	81	81	88
Mai	80	77	78	82	100	10	76	10	84	85	85	84
Jun	76	75	74	76	79	100	87	98	81	78	78	77
Jul	74	73	72	73	75	75	100	11	79	75	75	75
Ago	75	74	73	74	76	98	85	100	81	76	76	75
Set	79	78	76	78	82	83	78	10	100	81	81	82
Out	84	80	80	84	85	85	76	10	83	100	100	86
Nov	84	80	80	84	85	85	76	10	83	100	100	86
Dez	83	80	81	85	78	75	70	75	77	79	79	100

Fonte: a autora (2026).

Com relação às variáveis que mais contribuem para os modelos, na Figura 42 observa-se que a altitude é o fator que mais contribui para os modelos mensais que utilizam as variáveis segundo a bibliografia, e para os modelos de redes neurais artificiais, na ordem dos 20%. A distância dos cursos hídricos é o segundo fator mais relevantes nos modelos, com uma participação média de 17%. Em terceiro e quarto lugar estão a distância a partir de áreas urbanizadas e a precipitação média mensal, com 13% e 12% respectivamente.

O TRI apresentou a menor participação entre todas as categorias, com média de 0,13%. Das variáveis consideradas como deflagrantes dos huaicos pelo INGEMMET, nos modelos que consideram as variáveis da bibliografia, que favorecem a distribuição dos dados de localização do fenômeno, a altitude e a precipitação média mensal são os fatores influentes. A geologia, a geomorfologia e a declividade apresentam contribuições médias de 5%, 6% e 7%, respectivamente.

Novamente, percebe-se a contabilização de áreas com suscetibilidade de ocorrência de huaicos à medida que a altitude aumenta. Mas, visualmente, identificam-se nas formas do relevo os leitos de rios, sejam eles perenes ou efêmeros, essas áreas com suscetibilidade de ocorrência. Ainda, para a distância a partir dos cursos hídricos, as áreas com suscetibilidade de ocorrência se concentram a até 1.500 metros, com exceção da região da capital Lima, que apresenta maior área de abrangência. A respeito da distância a partir de cursos hídricos, a distância máxima é de cerca de 6.000 metros. O comportamento da precipitação é bastante difícil de relacionar às áreas com suscetibilidade de ocorrência de huaicos.

Com relação ao comportamento comparado à distribuição normal, repete-se o padrão de sensibilidade elevada dos dados que mais contribuíram para os modelos ocasionado pela precipitação média mensal.

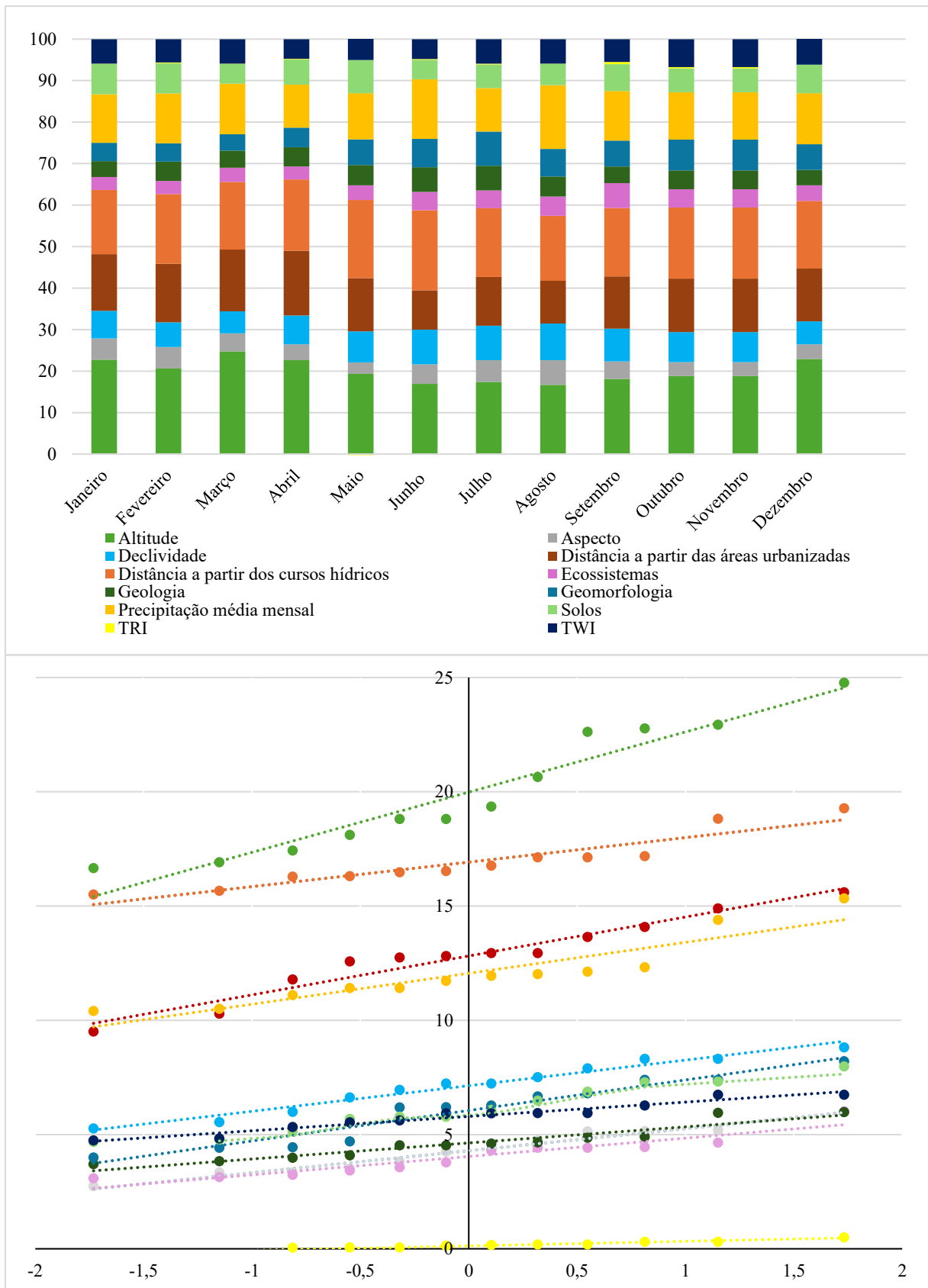


Figura 42. Contribuição das variáveis para cada modelo mensal Redes Neurais Artificiais considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal. Fonte: a autora (2026).

Nos modelos mensais que consideram apenas as variáveis indicadas pelo INGEMMET como relevantes para a ocorrência de huaicos, os meses apresentaram uma distribuição semelhante de contribuições das variáveis ambientais, com a altitude como a mais significativa. O segundo fator mais importante foi a precipitação média mensal. Os demais fatores – declividade, geologia e geomorfologia – têm menor influência e contribuições semelhantes (Figura 43).

Embora a altitude tenha representado o maior valor de contribuição, com uma participação média de 40%, essa contribuição variou de 33% em junho, a 45% em outubro e novembro. Não foi possível identificar uma relação direta entre esses diferentes níveis de precipitação e a área classificada como com suscetibilidade de ocorrência de huaicos. Para a precipitação média mensal, a contribuição média foi de 24%, sendo a menor no mês de fevereiro (16%) e a maior na variável ocorrendo no mês de junho (33%). A relação entre as participações do menor e do maior mês em termos de precipitação é inversamente proporcional ao nível precipitação. Porém, não há uma relação linear inversamente proporcional com os demais meses. Do comportamento pela distribuição normal, a precipitação média mensal e a altitude predominam nos modelos mensais. Novamente, destaca-se o impacto da precipitação média mensal nos modelos e nas demais variáveis, especialmente na que mais contribuiu para o modelo.

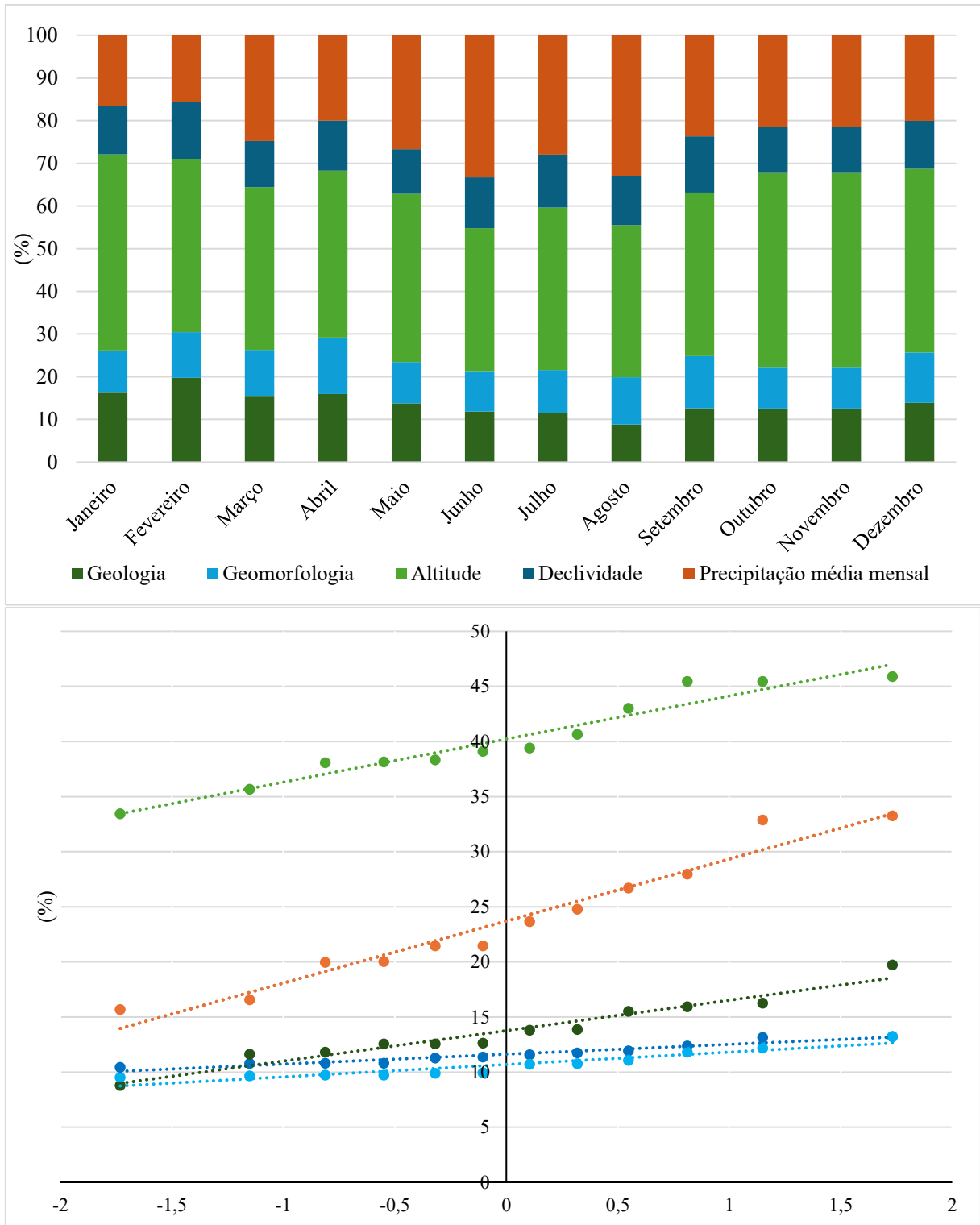


Figura 43. Contribuição das variáveis para cada modelo mensal Redes Neurais Artificiais considerando as variáveis ambientais indicadas pelo INGEMMET (a) e o comportamento de cada dado com relação a uma distribuição normal (b). Fonte: a autora (2026).

Ullah et al. (2022) utilizaram as variáveis ambientais de ângulo de inclinação, elevação, curvatura do plano, curvatura do perfil, geologia, distância das estradas, distância dos cursos d'água, índice de potência fluvial (SPI), índice de transporte de sedimentos (STI), índice

de umidade topográfica (TWI), índice de vegetação por diferença normalizada (NDVI), cobertura do solo, precipitação média anual e orientação em sua modelagem de rede neural artificial. Com base nesses parâmetros, os fatores que mais afetaram o risco de fluxos de detritos foram inclinação, altitude, precipitação e distância das estradas, o que coincide em grande parte com as variáveis mais recorrentes neste estudo. A altitude e precipitação são fatores que se repetem ao longo das aplicações de algoritmos, no geral. É importante notar, no entanto, que, além de avaliar apenas os fluxos de detritos, os autores também levaram em consideração a suscetibilidade de ocorrência de enchentes repentinas e deslizamentos de terra.

Em contraste, Pei e Qiu (2023) observaram que o modelo de rede neural artificial tendeu a superestimar os riscos de fluxos de detritos. Na presente pesquisa, ocorreu o oposto, visto que o algoritmo de rede neural artificial foi responsável por delimitar as menores áreas totais com suscetibilidade de ocorrência de fluxos de detritos em comparação com os outros métodos de aprendizado de máquina avaliados.

4.2.5 Regressão Logística

A Tabela 28 contém a qualidade dos modelos de regressão logística. Todos os modelos apresentam qualidade adequada no “AUC de Teste”, que deve ser superior a 0,7. Os modelos mensais de regressão logística, tanto com as variáveis da bibliografia, quanto do INGEMMET foram classificados como muito bons, sendo que a classificação apresentada por Yesilnacar (2005) varia de fraca, a mediana, a boa, a muito boa e a excelente.

O índice Kappa, que varia de fraco, a mediano, a bom, a excelente e a muito excelente (Cohen, 1960), para os modelos gerados com as variáveis da bibliografia e do INGEMMET apresentou classificação excelente.

Embora classificados como excelentes, os resultados de índices Kappa para os modelos de regressão logística são inferiores, de maneira geral, aos resultados dos algoritmos de máquina de vetores de suporte, Maxent e redes neurais artificiais. Além disso, os resultados entre os modelos com as variáveis do INGEMMET e as variáveis da bibliografia são bastante similares.

Tabela 28. Qualidade dos modelos de Regressão Logística

Mês	INGEMMET		Todas as variáveis	
	AUC	Kappa	AUC	Kappa
Janeiro	0,83	0,77	0,87	0,69
Fevereiro	0,83	0,77	0,88	0,68
Março	0,84	0,77	0,88	0,71
Abril	0,83	0,76	0,88	0,71
Maio	0,84	0,76	0,88	0,72
Junho	0,87	0,77	0,88	0,73
Julho	0,88	0,78	0,88	0,74
Agosto	0,86	0,78	0,88	0,73
Setembro	0,84	0,77	0,88	0,73
Outubro	0,83	0,76	0,88	0,71
Novembro	0,83	0,76	0,88	0,71
Dezembro	0,83	0,77	0,88	0,71

Fonte: a autora. (2026).

As Figuras 44 e 45 apresentam os resultados dos modelos, reclassificados para representar as áreas com suscetibilidade de ocorrência de huaicos. Em média, 28% das áreas do Departamento de Lima apresentam suscetibilidade de ocorrência de huaicos quando considerados os modelos que utilizam todas as variáveis. O mês de fevereiro retornou a menor área com suscetibilidade de ocorrência mensal, são 9.520,54 km², (27,19%). O maior valor de área com suscetibilidade corresponde a 28,46% (9.965,13 km²) no mês de maio.

Em média, 32% do Departamento de Lima apresenta suscetibilidade de ocorrência de huaicos. O mês de julho apresentou o menor valor mínimo de área com suscetibilidade de ocorrência de huaicos, com 10.572,20 km² (30,20%), e o mês com maior suscetibilidade de ocorrência foi maio, com 33,02% do departamento de Lima (11.562,00 km²).

No mapa, o comportamento geral também é semelhante ao dos demais modelos, exceto nos modelos do Maxent. Os resultados dos modelos das variáveis, segundo a bibliografia, seguem, de maneira geral, a proximidade com os cursos hídricos e com os dados de ocorrência. Ainda assim, nos modelos com as variáveis do INGEMMET, o padrão de espalhamento pela porção mais costeira se repete.

Em relação à coincidência entre os modelos mensais, as Tabelas 29 e 30 contêm as matrizes de coincidência das áreas com suscetibilidade de ocorrência para cada modelo, com os meses na horizontal como referência. Em média, há uma coincidência de 99% entre os modelos mensais com base nas variáveis da bibliografia. Já com relação às variáveis do INGEMMET, a coincidência entre os modelos mensais é 98%. Portanto, são observadas poucas diferenças entre os modelos mensais, tanto para as variáveis do INGEMMET, quanto as variáveis segundo a bibliografia.

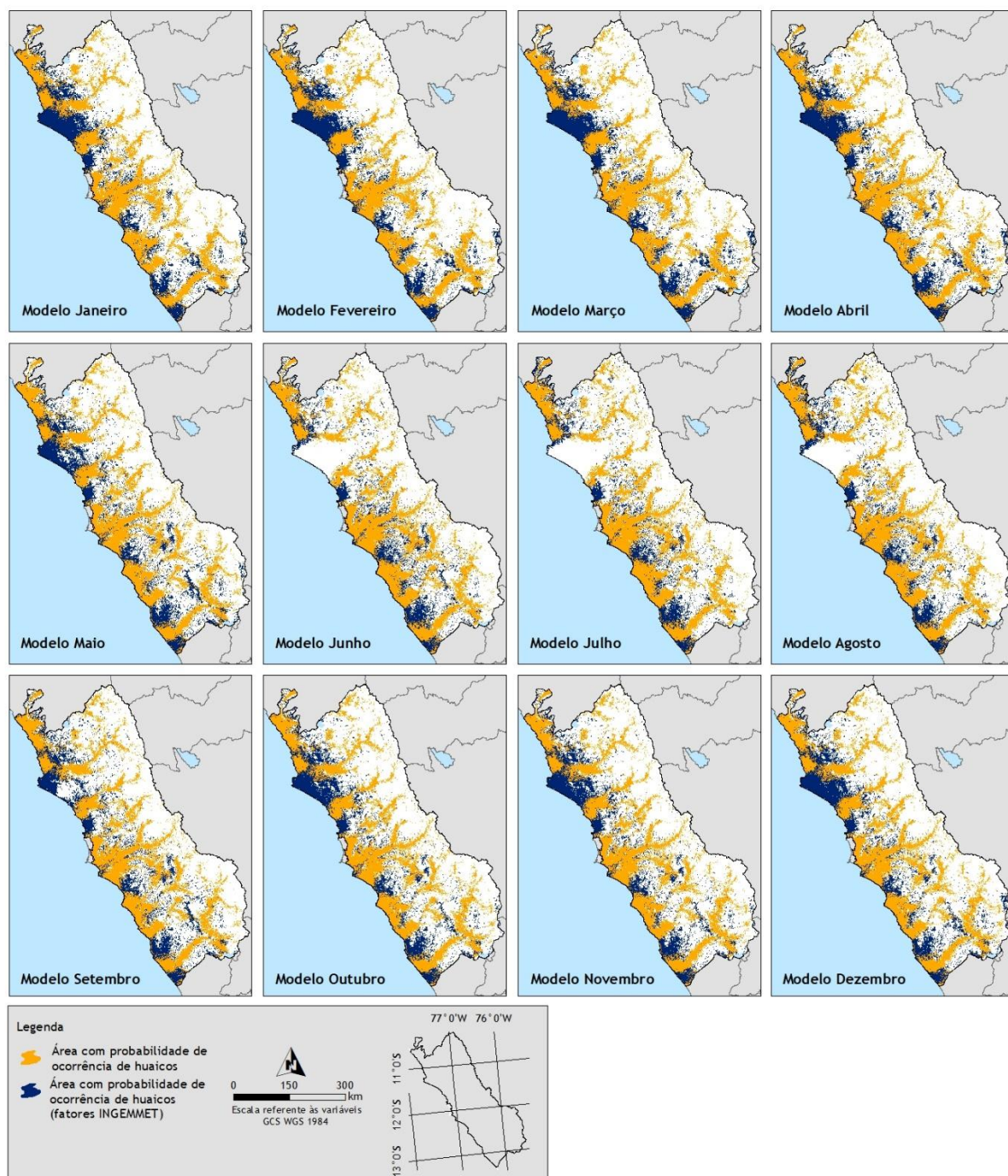


Figura 44. Resultados de suscetibilidade considerando os modelos mensais. Fonte: Santos et al. (2025).

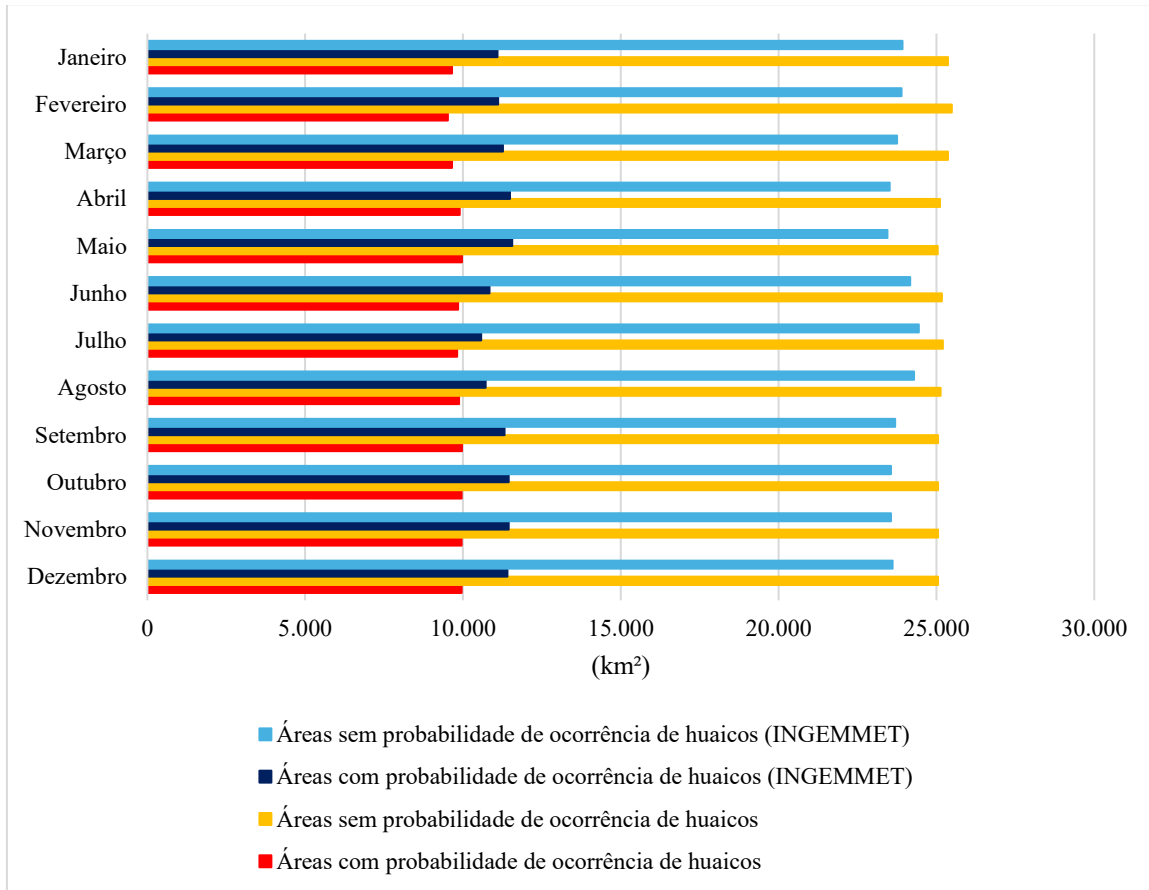


Figura 45. Áreas com e sem suscetibilidade de ocorrência de huaicos segundo os modelos de Regressão Logística. Fonte: a autora (2026).

Tabela 29. Coincidência entre modelos mensais Regressão Logística – Variáveis segundo o INGEMMET (%)

	Jan	Fev	Mar	Abr	Maio	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	100	99	99	98	97	97	97	98	98	98	99
Fev	100	100	100	99	98	96	97	97	98	98	98	99
Mar	100	100	100	99	98	97	97	97	98	98	98	99
Abr	99	99	99	100	99	97	97	98	99	99	99	100
Maio	99	98	99	99	100	98	98	98	99	99	99	99
Jun	96	96	96	96	97	100	99	99	98	97	97	97
Jul	96	96	96	96	97	99	100	99	97	97	97	96
Ago	97	96	96	97	97	99	99	100	98	97	97	97
Set	98	98	98	98	99	98	98	99	100	99	99	99
Out	99	99	99	99	99	98	98	98	99	100	100	99
Nov	99	99	99	99	99	98	98	98	99	100	100	99
Dez	99	99	99	100	99	97	97	98	99	99	99	100

Fonte: a autora (2026).

Tabela 30. Coincidência entre modelos mensais Regressão Logística – Variáveis segundo a bibliografia (%)

	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Jan	100	100	99	99	98	98	98	98	98	98	98	98
Fev	100	100	99	98	98	98	98	98	98	98	98	98
Mar	99	99	100	99	99	98	98	98	99	99	99	99
Abr	99	99	99	100	100	99	99	99	100	100	100	100
Mai	99	99	99	100	100	99	99	99	100	100	100	100
Jun	98	98	99	99	99	100	100	99	99	99	99	99
Jul	98	98	98	99	99	100	100	100	99	99	99	99
Ago	98	98	99	99	99	99	100	100	99	99	99	99
Set	99	99	99	100	100	99	99	99	100	100	100	100
Out	99	99	99	100	100	99	99	99	100	100	100	100
Nov	99	99	99	100	100	99	99	99	100	100	100	100
Dez	99	99	99	100	100	99	99	99	100	100	100	100

Fonte: a autora (2026).

Com relação às variáveis que mais contribuem para os modelos, na Figura 46 novamente, observa-se que a altitude é o fator que mais expressivo para os modelos mensais que consideram as variáveis de acordo com a bibliografia. A altitude contribuiu para a faixa de 32%. Essa participação é seguida pela distância a partir de cursos hídricos, com uma participação média de 21% e distância a partir de áreas urbanizadas, com uma contribuição média de 17%.

A geologia alcançou a segunda menor participação nos modelos de regressão logística com as variáveis da bibliografia, com 1% de participação média, atrás somente do TRI. Destaca-se ainda mais o comportamento da geologia neste modelo, devido ao fato de que nos modelos análogos com o Maxent, a geologia retornou a maior participação entre as variáveis segundo a bibliografia.

Ainda, para a altitude, novamente, observam-se áreas mais elevadas com suscetibilidade de ocorrência de huaicos, mas a configuração dessas ocorrências está associada as formas do relevo, os leitos de rios, sejam eles perenes ou efêmeros. A distância a partir dos cursos hídricos, as áreas com suscetibilidade de ocorrência se concentram a uma distância de até 4.000 metros. Para a distância a partir de áreas urbanizadas, há a concentração das áreas com suscetibilidade de ocorrência de huaicos associadas a áreas de maior concentração de urbanização, alcançando uma distância máxima de 10.000 metros de áreas urbanizadas. Mais uma vez, o comportamento da precipitação é bastante errático, especialmente em pontos próximos a áreas com suscetibilidade de ocorrência observam-se precipitações baixas.

Das variáveis consideradas como deflagrações dos huaicos pelo INGEMMET, nos modelos que consideram as variáveis pela bibliografia, a altitude possui a maior contribuição e a geologia uma das menores. A declividade e a geomorfologia possuem baixa contribuição, na ordem de 3%. A precipitação média mensal retornou valores na ordem de 7% de contribuição.

Com relação ao comportamento comparado a distribuição normal, o gráfico segue o mesmo padrão de algoritmos como floresta aleatória e redes neurais artificiais: nos dados que mais contribuíram para os modelos, observa-se a maior sensibilidade associada a introdução dos dados de precipitação média mensal. A inserção desses dados mensais modifica as respostas das demais variáveis em termos de contribuição e variação na distribuição normal.

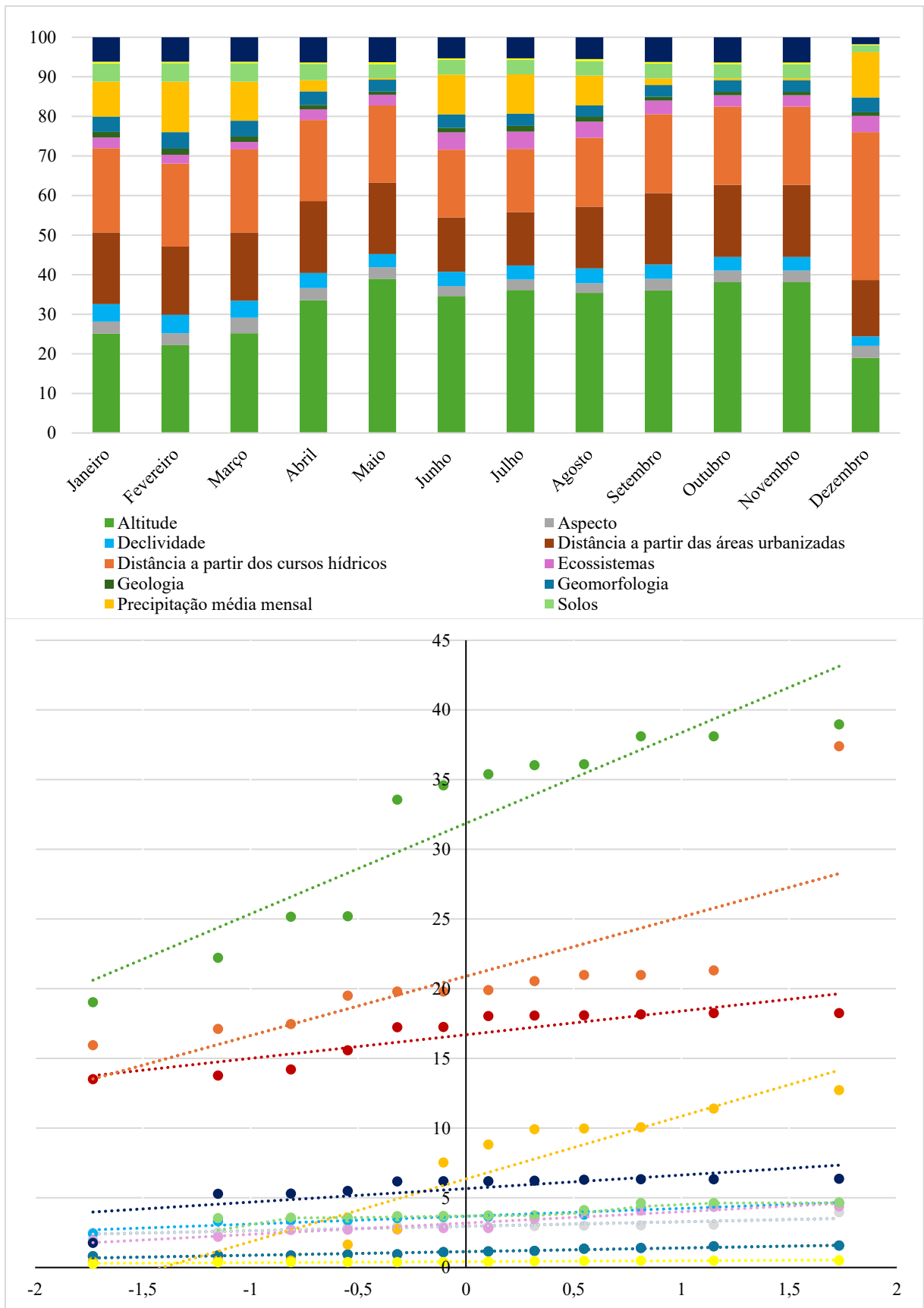


Figura 46. Contribuição das variáveis para cada modelo mensal Regressão Logística considerando as variáveis ambientais segundo a bibliografia e o comportamento de cada dado com relação a uma distribuição normal.
Fonte: a autora (2026).

Nos modelos mensais que consideram apenas as variáveis indicadas pelo INGEMMET como relevantes para a ocorrência de huaicos, diferentemente dos demais modelos, a precipitação média mensal foi a variável mais significativa. Os meses de janeiro, fevereiro e março apresentaram os maiores valores de participação mensal, com 55%, 56 e 53%. Meses como junho, agosto e setembro apresentaram participações da precipitação média mensal na faixa entre 47% a 48%. Os meses de outubro e novembro retornaram participações na ordem de 32%. Embora os meses de janeiro, fevereiro e março coincidam com as maiores precipitações e participações, não há uma relação linear entre as demais participações e meses. Os demais fatores têm menor influência e contribuições semelhantes, mas, ainda assim, se destaca a participação da geomorfologia, com uma média de 26% (Figura 47).

A respeito do comportamento pela distribuição normal, a precipitação média mensal e a geomorfologia predominam nos modelos mensais. Mas nestes modelos a predominância da precipitação média mensal causou perturbações em basicamente todas as demais variáveis.

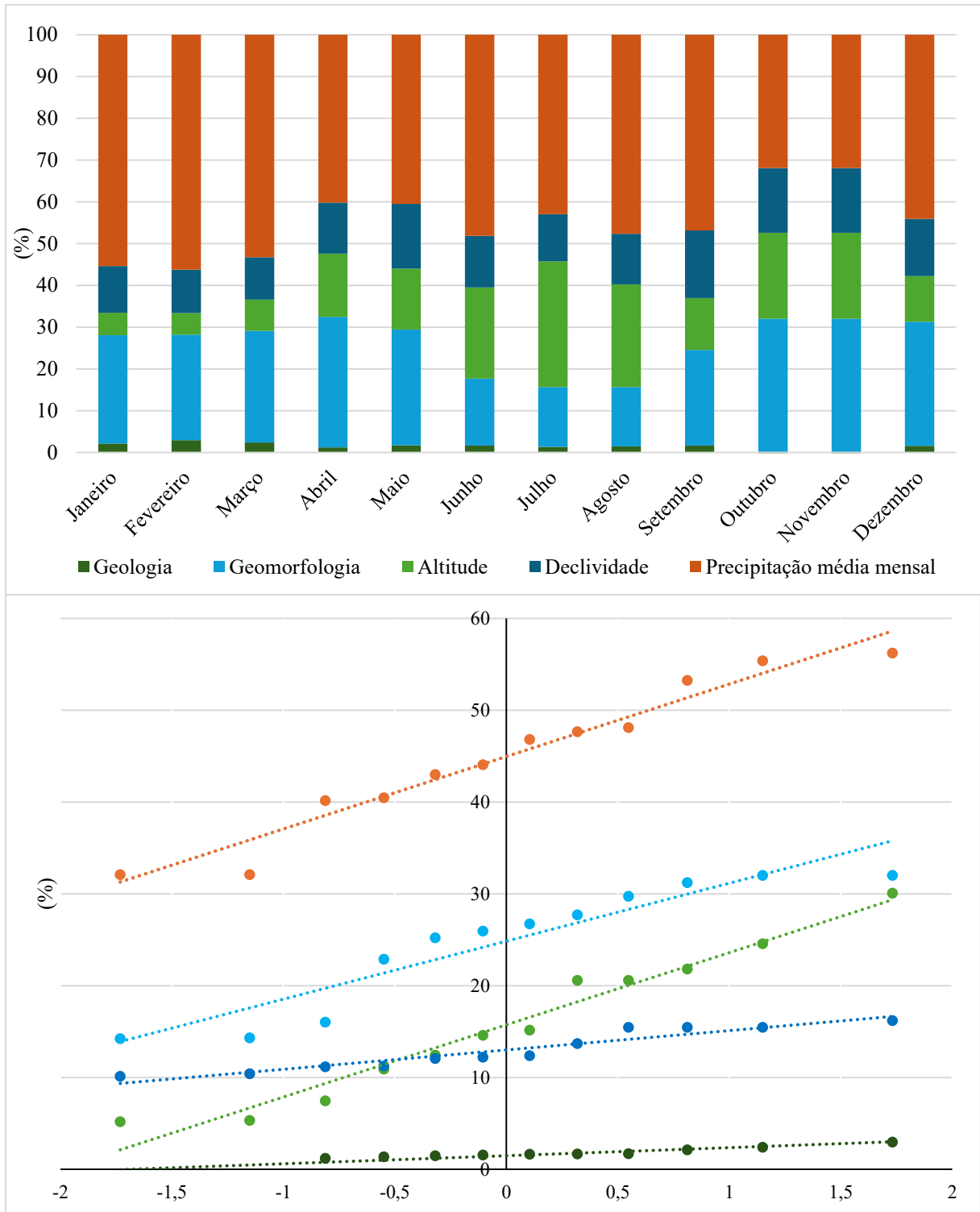


Figura 47. Contribuição das variáveis para cada modelo mensal de Regressão Logística, considerando as variáveis ambientais indicadas pelo INGEMMET (a) e o comportamento de cada variável em relação a uma distribuição normal (b). Fonte: a autora (2026).

Gu et al. (2023) descrevem que os modelos com máquinas de vetores de suporte e regressão logística apresentaram resultados semelhantes e, de modo geral, as áreas de alta e muito alta suscetibilidade nesses dois métodos são significativamente menores do que nos métodos tradicionais de atribuição de pesos. Para este trabalho, destaca-se que, embora os

resultados obtidos na aplicação deste algoritmo tenham sido satisfatórios, dentre os algoritmos, o algoritmo de regressão logística apresentou o pior desempenho, com relação à validação pelo índice Kappa.

Na aplicação de Ahmad et al. (2021), os fatores mais influentes no modelo foram a geologia, a altitude, a declividade. Ainda, possuem elevada participação no modelo: a precipitação e o uso e a cobertura do solo. Nesta pesquisa, o comportamento do algoritmo mostrou-se altamente complexo e instável, alterando radicalmente a ordem de importância das variáveis dependendo do conjunto de dados fornecido, embora altitude e precipitação tenham se estabelecido como variáveis estruturais importantes em ambos os cenários avaliados.

4.2.6 Considerações a respeito das modelagens

É importante ressaltar que os próprios registros de ocorrência de huaicos introduzem incerteza no modelo. Como os dados de ocorrência utilizados como entrada para o modelo estavam no formato de pontos – assim como os dados disponíveis – não existem registros das áreas discretas afetadas pelos huaicos, e não há como inserir áreas como dados de entrada no MaxEnt e para os demais algoritmos. Isso limita a geração de um modelo que reflita a realidade com maior precisão. Além disso, não há garantia de que os pontos identificados como áreas de ocorrência de huaicos tenham sido localizados corretamente.

Adicionalmente, os demais dados de entrada apresentam escalas de elaboração distintas. Os dados de solos, em particular, estão em escala 1:2.000.000 e apresentam informações genéricas quanto à espacialidade. Isso não impede a avaliação de uma grande área, como o Departamento de Lima, mas, nos modelos em que esses dados são inseridos, pode reduzir a qualidade dos resultados. Por esse mesmo motivo, nos modelos que consideram apenas as variáveis indicadas como influentes pelo INGEMMET, a ausência de dados de solo reduz a incerteza.

Ao avaliar cada conjunto de modelos separadamente, embora todos tenham atingido níveis satisfatórios tanto na qualidade do modelo quanto no índice Kappa, os melhores resultados foram obtidos com o Maxent, que apresentou resultados aproximadamente perfeitos tanto na qualidade do modelo quanto no índice Kappa. O conjunto de modelos com as variáveis do INGEMMET, com os algoritmos máquina de vetores de suporte e redes neurais artificiais também resultou nos segundos melhores resultados, quando avaliados os índices Kappa.

Ambos os modelos (máquina de vetores de suporte e redes neurais artificiais) retornaram valores semelhantes quanto à participação das variáveis no modelo, com a altitude

como variável mais importante, seguida da precipitação média mensal. As áreas resultantes em termos de espacialidade, seguem particularmente os leitos dos rios. E em termos de quantidade, em média as áreas dos modelos de máquina de vetores de suporte são 7% superiores que as áreas resultantes como com suscetibilidade de ocorrência de huaicos dos modelos de redes neurais artificiais.

Além disso, essas apresentam comportamentos distintos dos resultados obtidos com o algoritmo Maxent. Em termos de qualidade do modelo, os que consideram as variáveis conforme a bibliografia, são ligeiramente melhores. Com relação aos índices Kappa, são praticamente iguais, já que, em alguns meses, os modelos com as variáveis do INGEMMET são superiores e, em outros, inferiores. Ainda assim, comparando os modelos com variáveis segundo o INGEMMET elaborados com o Maxent e com os algoritmos máquina de vetores de suporte e redes neurais artificiais, os modelos com o Maxent possuem uma correspondência com os dados reservados para validação superior em 5 a 9%.

Observando o comportamento apenas dos modelos com o Maxent, os bons resultados do conjunto de modelos com as variáveis da bibliografia evidenciam que 40% das áreas do departamento de Lima são apontadas como tendo suscetibilidade de ocorrência de huaicos. Os demais modelos, mesmo com índices Kappa inferiores, apontaram uma área máxima de suscetibilidade de 29%. E os próprios modelos do Maxent, com as variáveis do INGEMMET, indicaram uma área média de 26% no departamento de Lima. Assim, entende-se que o modelo do Maxent que considera as variáveis segundo a bibliografia, superestimou as áreas com suscetibilidade de ocorrência de huaicos, devido ao fato do aumento de variáveis, o que dificulta a identificação das variáveis mais influentes.

Porém, a partir dos resultados de todos os modelos, independentemente do conjunto de variáveis e do algoritmo, observou-se que não houve padrão no comportamento da precipitação média mensal. Em meses onde a precipitação era mínima a participação nos modelos era elevada, em meses onde a precipitação era elevada, a participação nos modelos era pequena. Ainda assim, não se observou esse comportamento como verdadeiro ao longo de todos os meses.

Além disso, a presença da precipitação média mensal impactou as demais variáveis no modelo de forma diferente, a cada mês, em cada conjunto de dados e em cada algoritmo. Portanto, sugere-se uma avaliação futura sobre como se comportam os conjuntos de dados após a remoção do dado de precipitação. O comportamento da precipitação pode estar atrelado ao fato de que mínimos níveis já podem ser suficientes para desencadear huaicos.

Giráldez et al. (2022) avaliaram os eventos climáticos extremos de precipitação em

Lima Metropolitana entre 1965 e 2013. Os autores registraram 16 mm de precipitação em 1970, valor suficiente para alagar ruas, causar falhas no fornecimento de energia elétrica, danificar as instalações do Aeroporto Internacional Jorge Chávez, destruir cerca de 2.000 casas e provocar deslizamentos e transbordamentos de rios.

Além desses, observou-se que algumas variáveis não se mostraram relevantes em nenhuma das avaliações. O aspecto, os ecossistemas, os tipos de solos, o TRI (*terrain ruggedness index*), o TWI (*topographic wetness index*) e a própria declividade não alteram significativamente os resultados entre os conjuntos de dados. Ademais, as variáveis que contribuíram com os modelos que consideram as variáveis segundo a bibliografia, distância a partir de cursos hídricos e a partir de áreas urbanizadas, não contribuem diretamente na formação do fenômeno, mas sim estão associadas as características de áreas associadas ao registro de ocorrência de huaicos. Ainda assim, ambas as variáveis se revelaram relevantes nos modelos, exceto os do Maxent. Resume-se no Quadro 9 os principais resultados das aplicações dos algoritmos.

Quadro 9. Resumo dos resultados obtidos

Algoritmo	Conjunto de variáveis	AUC médio	Índice Kappa médio	Área média com suscetibilidade	Variável mais relevante
Floresta aleatória	INGEMMET	0,86	0,59	14%	Altitude
	Bibliografia	0,92	0,77	11%	Distância a partir de cursos hídricos
Máquina de vetores de suporte	INGEMMET	0,85	0,87	25%	Altitude
	Bibliografia	0,88	0,76	29%	Altitude
Maxent	INGEMMET	0,91	0,96	26%	Geologia
	Bibliografia	0,93	0,95	39%	Geologia
Redes Neurais Artificiais	INGEMMET	0,85	0,91	24%	Altitude
	Bibliografia	0,89	0,75	25%	Altitude
Regressão logística	INGEMMET	0,84	0,77	32%	Precipitação média mensal
	Bibliografia	0,88	0,71	28%	Altitude

Fonte: a autora (2026).

Em avaliações futuras, sugere-se a remoção dessas variáveis dos modelos. No entanto, considerando os conjuntos de dados elaborados neste trabalho, entende-se que todos os algoritmos têm potencial para a modelagem de perigos naturais. Ainda que para a situação aqui proposta, o Maxent, na condição de avaliação das variáveis segundo o INGEMMET apresentou os melhores resultados. Mesmo com a variação mensal decorrente dos diferentes níveis de precipitação histórica, aproximadamente 95% das áreas resultantes desses modelos coincidem. Recomenda-se, portanto, para futuras replicações dos resultados, o uso do algoritmo Maxent, considerando as variáveis apontadas pelo INGEMMET para identificação de áreas com suscetibilidade de ocorrência de huaicos.

Inevitavelmente, as mudanças climáticas nessa região indicam um aumento da temperatura, o que provavelmente levará a mais eventos de precipitação, aumentando o número de huaicos e de áreas afetadas. Ou seja, não há expectativa razoável de reduzir essas áreas suscetíveis. Essas áreas poderiam ser zoneadas por meio de políticas públicas, nas quais se adotariam ações de preparação para desastres. Além disso, essas áreas poderiam ser avaliadas para evitar ocupações ou até mesmo desocupadas. No que se refere às políticas públicas, os moradores devem ser continuamente orientados sobre os riscos dos huaicos, e medidas estruturais, como sistemas de contenção do fluxo, podem ser implementadas para minimizar os impactos desses eventos.

5 CONCLUSÕES

Entende-se que este trabalho atingiu o objetivo geral de desenvolver e validar um modelo de machine learning para identificar áreas com suscetibilidade de huaicos no departamento de Lima, Peru. A partir da análise dos algoritmos disponíveis e trabalhos que fazem uso desses para a modelagem de suscetibilidade de ocorrência fenômenos naturais nas mais diversas áreas do planeta, foi possível identificar e mapear primeiramente, quais os fatores com potencial de contribuição os modelos. Foram elaborados conjuntos de modelos para os algoritmos floresta aleatória, máquina de vetores de suporte, maxent, redes neurais artificiais e regressão logística, considerando variáveis relacionadas ao relevo, a formação natural, ao uso e cobertura do solo e ao clima.

Todos os modelos, para cada algoritmo, para os conjuntos de dados de entrada conforme o INGEMMET e conforme a bibliografia e ainda, com a alteração da precipitação média mensal, atenderam os critérios de qualidade dos modelos, a área abaixo da curva (AUC) e, para além disso, em resultados em índices Kappa satisfatórios. Ainda, destaca-se que os resultados variam desde uma porcentagem de área do departamento de Lima com suscetibilidade de ocorrência de huaicos correspondente a um mínimo de 14% até 39%, sendo o algoritmo com os melhores resultados retornando que 26% da área de estudo possui suscetibilidade de ocorrência de huaicos.

Detalha-se que o comportamento das variáveis segundo a bibliografia favorece a distribuição espacial dos dados, enquanto as variáveis apontadas pelo INGEMMET são variáveis relacionadas aos fatores deflagradores do fenômeno.

Ainda, partir dos resultados de todos os modelos, independentemente do conjunto de variáveis e do algoritmo, observou-se que não houve padrão no comportamento da precipitação média mensal. Em meses em que a precipitação era mínima, a participação nos modelos era elevada; em meses em que a precipitação era elevada, a participação nos modelos era pequena. Ainda assim, não se observou esse comportamento como verdadeiro ao longo de todos os meses.

Além disso, a presença da precipitação média mensal impactou as demais variáveis no modelo de forma diferente, a cada mês, a cada conjunto de dados, em cada algoritmo. Portanto, sugere-se uma avaliação futura sobre como se comportam os conjuntos de dados após a remoção do dado de precipitação. O comportamento da precipitação pode estar atrelado ao fato de que mínimos níveis já podem ser suficientes para desencadear huaicos.

Tendo em vista a temática aqui apresentada, reforça-se a importância de elaborar um

mapeamento para identificar áreas com suscetibilidade de ocorrência de *huaicos* no Peru. O país é atingido pelos mais diversos fenômenos naturais, mas possui em recorrência de *huaicos* que impactam significativamente a população.

REFERÊNCIAS

- ADO, M.; AMITAB, K.; MAJI, A. K.; JASIŃSKA, E.; GONO, R.; LEONOWICZ, Z.; JASIŃSKI, M. Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey. **Remote Sensing**, v. 14, n. 13, p. 3029, 2022.
- ADSUARA, J. E.; PEREZ-SUAY, A.; MUNOZ-MARI, J.; MATEO-SANCHIS, A.; PILES, M.; CAMPS-VALLS, G. J. I. T. G. Nonlinear Distribution Regression for Remote Sensing Applications. **IEEE Transactions on Geoscience and Remote Sensing**, v. 57, n. 12, p. 10025–10035, 2019.
- AHMAD, H.; NINGSHENG, C.; RAHMAN, M.; ISLAM, M. M.; POURGHASEMI, H. R.; HUSSAIN, S. F.; HABUMUGISHA, J. M.; LIU, E.; ZHENG, H.; NI, H.; DEWAN, A. Geohazards Susceptibility Assessment along the Upper Indus Basin Using Four Machine Learning and Statistical Models. **ISPRS International Journal of Geo-Information**, v. 10, n. 5, p. 315, 2021.
- AKEL, A. J. N.; HOVSTADB, J. S.; RUTHB, M. S.; PARMEGGIANIC, S.; PATRIARCAA, R.; PALTRINIERIB, N. A Machine Learning Approach to Analyzing Natural-Hazard Accident Scenarios. **Chemical Engineering Transactions**, v. 91, 2022.
- ALCÂNTARA, E.; BAIÃO, C. F.; GUIMARÃES, Y. C.; MANTOVANI, J. R.; MARENGO, J. A. Machine learning approaches for mapping and predicting landslide-prone areas in São Sebastião (Southeast Brazil). **Natural Hazards Research**, 2024.
- ALKHATIB, R.; SAHWAN, W.; ALKHATIEB, A.; SCHÜTT, B. A Brief Review of Machine Learning Algorithms in Forest Fires Science. **Applied Sciences**, v. 13, n. 14, p. 8275, 2023.
- ALNAIM, A.; SUN, Z.; TONG, D. Evaluating Machine Learning and Remote Sensing in Monitoring NO₂ Emission of Power Plants. **Remote Sensing**, v. 14, n. 3, p. 729, 2022.
- AMIN, M. F. Confusion matrix in three-class classification problems: a step-by-step tutorial. **Journal of Engineering Research**, v. 7, n. 1, 2023.
- ASWM - ASSOCIATION OF STATE WETLAND MANAGERS. **Report on State Definitions, Jurisdiction and Mitigation Requirements in State Programs for Ephemeral, Intermittent and Perennial Streams in the United States**. 2014. 175 p.
- BILSKIE, M. V.; ANGEL, D. DEL; YOSKOWITZ, D.; HAGEN, S. C. Northern Gulf of Mexico Flood Risk Under Sea Level Rise. **Earth's Future**, v.10, n.4, 2020.
- AUDEBERT, N.; LE SAUX, B. LEFEVRE, S. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. **IEEE Geoscience and Remote Sensing Magazine**, v. 7, n. 2, p. 159–173, 2019.
- AVAND, M.; MORADI, H.; LASBOYEE, M. R. Spatial modeling of flood probability using geo-environmental variables and machine learning models, case study: Tajan watershed, Iran. **Advances in Space Research**, v. 67, n. 10, p. 3169–3186, 2021.
- BELGIU, M.; DRĂGUȚ, L. Random forest in remote sensing: A review of applications and

future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 114, p. 24–31, 2016.

BERHANE, T. LANE, C. WU, Q.; AUTREY, B. C.; ANENKHONOV, O. A.; CHEPINOGA, V. V.; LIU, H. J. R. Decision-Tree, Rule-Based, and Random Forest Classification of High-Resolution Multispectral Imagery for Wetland Mapping and Inventory. **Remote Sensing**, v. 10, n. 4, p. 580, 2018.

BOGNER, C.; SEO, B.; ROHNER, D.; REINEKING, B. J. P. Classification of rare land cover types: Distinguishing annual and perennial crops in an agricultural catchment in South Korea. **PLOS ONE**, v. 13, n. 1, p. e0190476, 2018.

CHANG, K.; MERGHADI, A.; YUNUS, A. P.; PHAM, B. T.; DOU, J. Evaluating scale effects of topographic variables in landslide susceptibility models using GIS-based machine learning techniques. **Scientific Reports**, v. 9, n. 1, 2019.

CHAUHAN, V.; GUPTA, L.; DIXIT, J. Machine learning and GIS-based multi-hazard risk modeling for Uttarakhand: Integrating seismic, landslide, and flood susceptibility with socioeconomic vulnerability. **Environmental and Sustainability Indicators**, v. 26, p. 100664, 2025.

CHEN, W.; POURGHASEMI, H.; R.; KORNEJADY, A.; ZHANG, N. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. **Geoderma**, v. 305, p. 314-327, 2017.

CHEN, Y.; QIN, S.; QIAO, S.; DOU, Q.; CHE, W.; SU, G.; YAO, J.; NNANWUBA, U. E. Spatial Predictions of Debris Flow Susceptibility Mapping Using Convolutional Neural Networks in Jilin Province, China. **Water**, v. 12, n. 8, p. 2079, 2020.

CHEN, Y.; ZHAO, C.; ZHI, H. Analysis of ENSO Event Intensity Changes and Time-Frequency Characteristic Since 1875. **Atmosphere**, v. 15, n. 12, p. 1428, 2024.

CPTEC/INPE - CENTRO DE PREVISÃO DE TEMPO E ESTUDOS CLIMÁTICOS/ INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Monitoramento do El Niño durante NDJ-2019/2020**. Disponível em: <http://enos.cptec.inpe.br>. 2022.

DALAPICOLLA, J. Tutorial de Modelos de Distribuição de Espécies: Guia Prático Usando o Maxent e o ArcGIS 10. In *Laboratório de Mastozoologia e Biogeografia*; Universidade Federal do Espírito Santo: Vitória, Brazil, 2016.

DARABI, H.; CHOUBIN, B.; RAHMATI, O.; HAGHIGHI, A. T.; PRADHAN, B.; KLØVE, B. Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. **Journal of Hydrology**, v. 569, p. 142-154, 2019.

ELBES, M.; ALZU'BI, S.; KANAN, T. Deep Learning-Based Earthquake Prediction Technique Using Seismic Data. In: **2023 International Conference on Multimedia Computing, Networking and Applications (MCNA)**. IEEE, 2023, p. 103–108.

FAO – Food and Agriculture Organization of the United Nations. Why are drylands so important? Disponível em: <https://www.fao.org/dryland-assessment/en/>.

FIELDING, A. H.; BELL, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. **Environmental Conservation**, v. 24, n. 1, p. 38–49, 1997.

FRENCH, A.; MECHLER, R.; ARESTEGUI, M.; MACCLUNE, K.; CISNEROS, A. Root causes of recurrent catastrophe: The political ecology of El Niño-related disasters in Peru. *International Journal of Disaster Risk Reduction*, v. 47, 2020.

GAO, S.; LI, X. Machine learning models for predicting urban flood susceptibility using remote sensing data. **Science of The Total Environment**, v. 747, 2020.

GANJIRAD, M.; DELAVAR, M. R. Flood risk mapping using random forest and support vector machine. **ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. X-4/W1-2022, p. 201–208, 2023.

GIRÁLDEZ, L.; SILVA, Y.; FLORES-ROJAS, J.L.; TRASMONTE, G. Diagnosis of the extreme climate events of temperature and precipitation in metropolitan lima during 1965–2013. **Climate**, v. 10, n. 8, p. 112, 2022.

GHOSH, S. S.; DEY, S.; BHOGAPURAPU, N.; N., HOMAYOUNI, S.; BHATTACHARYA, A.; MCNAIRN, H. J. R. S. Gaussian Process Regression Model for Crop Biophysical Parameter Retrieval from Multi-Polarized C-Band SAR Data. **Remote Sensing**, v. 14, n. 4, p. 934, 2022.

GONZALES, E.; INGOL, E. Determination of a New Coastal ENSO Oceanic Index for Northern Peru. **Climate**, v. 9, n. 5, p. 71, 2021.

GRAFF, C. A.; COFFIELD, S. R.; CHEN, Y.; FOUFOULA-GEORGIOU, E.; RANDERSON, J. T.; SMYTH, P. J. I. T. G. Forecasting Daily Wildfire Activity Using Poisson Regression. **IEEE Transactions on Geoscience and Remote Sensing**, v. 58, n. 7, p. 4837–4851, 2020.

GU, F.; CHEN, J.; SUN, X.; LI, Y.; ZHANG, Y.; WANG, Q. Comparison of machine learning and traditional statistical methods in debris flow susceptibility assessment: A case study of Changping District, Beijing. **Water**, v. 15, n. 4, p. 705, 2023.

GUYON, X.; YAO, J. On the Underfitting and Overfitting Sets of Models Chosen by Order Selection Criteria. **Journal of Multivariate Analysis**, v. 70, n. 2, p. 221–249, 1999.

HIEN, D. M.; VAN HOANG, N.; LE DUNG, M.; HUU DUNG, L.; THI THUY, N.; THI HANG, V. Large-scale Mapping of Landslide and Debris Flow using Flowr Model with Statistical and Machine Learning Methods. **VNU Journal of Science: Earth and Environmental Sciences**, v. 38, n. 4, 2022.

HSIEH, W. W. Improving Predictions by Nonlinear Regression Models from Outlying Input Data. **Journal of Environmental Informatics**, 2023.

HU, J.; PANG, A.; DENG, C. Urban flood risk analysis using presence-only machine learning approach: an integrated MaxEnt-cloud model framework in Harbin, China. **Natural Hazards**,

v. 121, n. 14, p. 16827–16856, 2025.

HUANG, H.; WANG, Y.; LI, Y.; ZHOU, Y.; ZENG, Z. Debris-Flow Susceptibility Assessment in China: A Comparison between Traditional Statistical and Machine Learning Methods. **Remote Sensing**, v. 14, n. 18, p. 4475, 2022.

HUYEN, C. *Projetando sistemas de Machine Learning: processo interativo para aplicações prontas para produção*. Rio de Janeiro: Editora Alta Books, 2024. Ebook. ISBN 9788550819648.

IFRC - INTERNATIONAL FEDERATION OF RED CROSS AND RED CRESCENT SOCIETIES. What is a disaster? 2022. Disponível em: <https://www.ifrc.org/our-work/disasters-climate-and-crises/what-disaster>.

INDECI - INSTITUTO NACIONAL DE DEFENSA CIVIL. **Compendio Estadístico del Instituto Nacional de Defensa Civil**. INDECI: Lima, Peru, 2021.

INDECI - INSTITUTO NACIONAL DE DEFENSA CIVIL. **Compendio Estadístico del Instituto Nacional de Defensa Civil**. INDECI: Lima, Peru, 2019.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Carta Geológica Nacional a escala 1:50000**. INGEMMET: Lima, Peru, 2022.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación de peligros de las zonas propuesta para reubicación del centro poblado Borogueña**. Informe Técnico n.º A7037. Peru, 2020.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación de riesgos en el barrio de Chua Baio**. Informe Técnico n.º A7315. Peru, 2022.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación de peligros geológicos por flujo de detritos en la quebrada Chaupimayo**. Informe Técnico n.º A7143. Peru, 2021.

INGEMMET – INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación de peligros geológicos en los sectores de Aplao, Casquina y Caspani**. Informe Técnico n.º A6876. 2019.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación de peligros geológicos por flujo de detritos en la quebrada Huarequeque, sector Huayan**. Informe Técnico n.º A7193. Peru, 2021.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Evaluación Geológica-Geodinámica del centro poblado Santa María de Otopongo**. Informe Técnico n.º A7053. Peru, 2020.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Inspección geológica en el asentamiento humano Sol Naciente - III etapa**. Informe Técnico n.º A6856. Peru, 2018.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. *Mapa de Inventario de Peligros Geológicos*. INGEMMET: Lima, Perú, 2019.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Mapa Geomorfológico del Perú**. INGEMMET: Lima, Peru, 2016.

INGEMMET - INSTITUTO GEOLÓGICO, MINERO Y METALÚRGICO. **Peligros geológicos en el Sector de Venado Muerto**. Informe Técnico n.º A6654. Peru, 2014.

INNOCENTI, C.; BATTAGLINI, L.; D'ANGELO, S.; FIORENTINO, A. Submarine landslides: mapping the susceptibility in European seas. **Quarterly Journal of Engineering Geology and Hydrogeology**, v. 54, n. 1, 2021.

JAVIDAN, N.; KAVIAN, A.; POURGHASEMI, A. R.; CONOSCENTI, C.; JAFARIAN, Z.; COMINO, J. R. Evaluation of multi-hazard map produced using MaxEnt machine learning technique. **Scientific Reports**, v. 11, 2021.

KALMAR, T. M.; DÎRJA, M.; RĂDULESCU, A. T. G. M.; MĂRAN, P. D.; RĂDULESCU, G. M. T.; NYCHVYD, M.; KALYNYCH, I.; RĂDULESCU, V. M. G. M.; ZAHARIA, G.; DANKU, G. Geospatial technologies for landslide monitoring: a case study of Sighetu Marmăției, Romania. **Environmental Earth Sciences**, v. 83, n. 10, 2024.

KIM, H. G.; PARK, C. Y. Landslide susceptibility analysis of photovoltaic power stations in Gangwon-do, Republic of Korea. **Geomatics, Natural Hazards and Risk**, v. 12, n.1, p. 2328-2351, 2021.

KORNEJADY, A.; OWNEGH, M.; BAHREMAND, A. Landslide susceptibility assessment using maximum entropy model with two different data sampling methods. **Catena**, v. 152, p. 144-162, 2017.

KOTSIANTIS, S. B.; KANELLOPOULOS, D.; PINTELAS, P. Data preprocessing for supervised learning. **International Journal of Computer Science**, v. 1, n. 1, p. 111–117, 2006.

KUMAR, C.; WALTON, G.; SANTI, P.; LUZA, C. An Ensemble Approach of Feature Selection and Machine Learning Models for Regional Landslide Susceptibility Mapping in the Arid Mountainous Terrain of Southern Peru. **Remote Sensing**, v. 15, n. 5, p. 1376, 2023.

LA MULA. **Así quedó Trujillo luego del huaico del domingo**: Fotos. Disponível em: <https://redaccion.lamula.pe/2017/03/20/trujillo-huanchaco-lluvias-huaicos/victorliza/>. Acesso em: 12 abr. 2021.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, v. 33, n. 1, p. 159-174, 1977.

LI, L.; ZHANG, Z.; ZHAO, D.; QIANG, Y.; NI, B.; WU, H.; HU, S.; LIN, H. Debris flow scale prediction based on correlation analysis and improved support vector machine. **Water**, v. 15, n. 23, p. 4161, 2023.

LIANG, D.; FREDERICK, D. A.; LLEDO, E. E.; ROSENFELD, N.; BERARDI, V.;

LINSTEAD, E.; MAOZ, U. J. B. I. Examining the utility of nonlinear machine learning approaches versus linear regression for predicting body image outcomes: The U.S. Body Project I. **Body Image**, v. 41, p. 32–45, 2022.

LIU, Y.; PEIHUA XU, P.; CHEN CAO, C.; BO SHAN, B.; ZHU, K.; MA, Q.; ZHANG, Z.; YIN, H. A comparative evaluation of machine learning algorithms and an improved optimal model for landslide susceptibility: a case study. **Geomatics, Natural Hazards and Risk**, v. 12, n.1, p. 1973–2001, 2021.

LIU, Y.; XU, P.; CAO, C.; SHAN, B.; ZHU, K.; MA, Q.; ZHANG, Z.; YIN, H. A comparative evaluation of machine learning algorithms and an improved optimal model for landslide susceptibility: a case study. **Geomatics, Natural Hazards and Risk**, v. 12, n. 1, p. 1973–2001, 2021.

LIU, Y.; ZHAO, L.; BAO, A.; LI, J.; YAN, X. Chinese High Resolution Satellite Data and GIS-Based Assessment of Landslide Susceptibility along Highway G30 in Guozigou Valley Using Logistic Regression and MaxEnt Model. **Remote Sensing**, v. 14, 2022.

LIU, F.; ZHANG, T.; DENG, Y.; QIAN, F.; YANG, N. **Landslide Susceptibility Prediction Based on Landform Predisposing Indexes - an Example from the Beiluo River Basin**. [s.l.]: Elsevier BV, 2023. Acesso em: 12 Apr. 2026.

LOKESH, P.; MADHESH, C.; ANEESH, M.; SHEKAR, P. R. Machine learning and deep learning-based landslide susceptibility mapping using geospatial techniques in Wayanad, Kerala state, India. **HydroResearch**, v. 8, p. 113–126, 2025.

LUMEN LEARNING. **Types of Mass Movement**. 2021. Disponível em: <https://courses.lumenlearning.com/wmopen-geology/chapter/outcome-types-of-mass-movement/>. Acesso em: 12 abr. 2021.

MAHARJAN, M.; TIMILSINA, S.; AYER, S.; SINGH, B.; MANANDHAR, B.; SEDHAIN, A. **Flood Risk Assessment Using Machine Learning Approach in Mohana-Khutiya River of Nepal**. [s.l.]: Elsevier BV, 2023. Acesso em: 12 Apr. 2026.

MANSOURI, E.; FEIZI, F.; JAFARI RAD, A.; ARIAN, M. J. S. E. Remote-sensing data processing with the multivariate regression analysis method for iron mineral resource potential mapping: a case study in the Sarvian area, central Iran. **Solid Earth**, v. 9, n. 2, p. 373–384, 2018.

MINAM - MINISTERIO DEL AMBIENTE. Mapa nacional de ecosistemas del Perú. 1:200.000. 2019. Disponível em: <https://sinia.minam.gob.pe/mapas/mapa-nacional-ecosistemas-peru>. Acesso em: 05 abr. 2021.

MINAM - MINISTERIO DEL AMBIENTE. Mapa Nacional de Suelos del Perú. INGEMMET: Lima, Peru, 2009.

MOKHTARI, M.; ABEDIAN, S. Spatial prediction of landslide susceptibility in Taleghan basin, Iran. **Stochastic Environmental Research and Risk Assessment**, v. 33, p. 1297–1325, 2019.

MOLINA-SCHILLER, D.; ROSALES, S. A.; FREITAS, T. O. Oceanographic conditions off coastal South America in relation to the distribution of Burmeister's Porpoise, *Phocoena spinipinnis*. *LAJAM*, v. 4, n. 2, p. 141-156, 2005.

NORALLAHI, M., KABOLI, H. S. Urban flood hazard mapping using machine learning models: GARP, RF, MaxEnt, and NB. **Natural Hazards**, v. 106, p. 119–137, 2021.

OREFICE, S.; INNOCENTI, C. Regional assessment of coastal landslide susceptibility in Liguria, Northern Italy, using MaxEnt. **Natural Hazards**, v. 121, n. 3, p. 2613–2639, 2024.

PANNAKKONG, W.; THIWA-ANONT, K.; SINGTHONG, K.; PARTHANADEE, P.; BUDDHAKULSOMSIRI, J. J. M. P. E. Hyperparameter Tuning of Machine Learning Algorithms Using Response Surface Methodology: A Case Study of ANN, SVM, and DBN. **Mathematical Problems in Engineering**, v. 2022, p. 1–17, 2022.

PASSUNI, S. Quebradas. INDECI (Instituto Nacional de Defensa Civil). 2019.

PCM - PRESIDENCIA DEL CONSEJO DE MINISTROS. Decreto Supremo n° 034 de 13 de mayo de 2014. Aprueba el Plan Nacional de Gestión del Riesgo de Desastres (PLANAGERD) 2014-2021. Lima, 2014. Disponible em: https://cdn.www.gob.pe/uploads/document/file/1229859/DS-N_034-2014-PCM-PLANAGERD20200812-2906259-1x7m674.pdf. Acceso Em: 11 mar. 2021.

PCM - PRESIDENCIA DEL CONSEJO DE MINISTROS. Decreto Supremo n° 048 de 26 de mayo de 2011. Aprueba el Reglamento de la Ley n° 29664, que crea el Sistema Nacional de Gestión del Riesgo de Desastres (SINAGERD). Lima, 2011. Disponible em: <https://www.geoidep.gob.pe/images/descargas/DS-048-2011-PCM.pdf>. Acceso em: 11 mar. 2021.

PCM - PRESIDENCIA DEL CONSEJO DE MINISTROS. Decreto Supremo n° 087 de 23 de diciembre de 2004. Reglamento de Zonificación Ecológica y Económica. Lima, 2004. Disponible em: <https://sinia.minam.gob.pe/normas/aprueban-reglamento-zonificacion-ecologica-economica-zee>. Acceso em: 09 abr. 2020.

PCM - PRESIDENCIA DEL CONSEJO DE MINISTROS. Lei n° 29.664 de 19 de febrero del 2011. Ley que crea el Sistema Nacional de Gestión del Riesgo de Desastres (SINAGERD). Lima, 2011. Disponible em: <https://busquedas.elperuano.pe/normaslegales/ley-que-crea-el-sistema-nacional-de-gestion-del-riesgo-de-de-ley-n-29664-605077-1/>. Acceso em: 11 mar. 2020.

PCM - PRESIDENCIA DEL CONSEJO DE MINISTROS. Mapas de Cuencas Hidrográficas del Perú. 2003.

PEI, T.; QIU, T. Debris flow susceptibility mapping in Colorado Front Range, USA: A comparison of physics-based and data-driven approaches. **E3S Web of Conferences**, v. 415, p. 01018, 2023.

PHILLIPS, S. J.; ANDERSON, R. P.; DUDÍK, M.; SCHAPIRE, R. E.; BLAIR, M. E. Opening the black box: an open-source release of MaxEnt. **Ecography**, v. 40, n.7, p. 887-893, 2017.

PRASAD, A. S.; FRANCESCUTTI, L. H. Natural Disasters. **International Encyclopedia of Public Health**, 2nd ed., v. 5, p. 215-222, 2017.

QING, F.; ZHAO, Y.; MENG, X.; SU, X.; QI, T.; YUE, D. Application of machine learning to debris flow susceptibility mapping along the China–Pakistan Karakoram Highway—**Remote Sensing**, v. 12, n. 18, p. 2933, 2020.

RAJABI, A.; SHABANLOU, S.; YOSEFVAND, F.; KIANI, A. Exploring the sample size and replications scenarios' effect on spatial prediction of food, using MARS and MaxEnt methods, case study: Saliantape catchment, Golestan, Iran. Ahmad. **Natural Hazards**, v. 109, n. 1, p. 871-901, 2021.

REAL ACADEMIA ESPAÑOLA. Huaico. Disponível em: <https://dle.rae.es/huaico>. 2022a.

REAL ACADEMIA ESPAÑOLA. Quebrada. Disponível em: <https://dle.rae.es/quebrada>. 2022b.

RIVERA, M.; VÍLCHEZ, M. & VELA, J. Peligros por huaicos en la ciudad de Arequipa. Taller Internacional: Fortalecimiento de capacidades para mitigar los impactos de los huaicos en Perú, Lima y Arequipa. 2018: Libro de Resúmenes. Arequipa: INGEMMET, p. 45-49.

RUSK, J.; MAHARJAN, A.; TIWARI, P.; CHEN, T. K.; SHNEIDERMAN, S.; TURIN, M.; SETO, K. C. Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya. **Science of The Total Environment**, v. 804, 2022.

SANTOS, G. M. dos. Análise das condicionantes ambientais da bacia hidrográfica de San Idelfonso – província de Trujillo, Peru. 2021. Dissertação (Mestrado em Engenharia e Ciências Ambientais) – Universidade de Caxias do Sul. Caxias do Sul, 2022.

SANTOS, G. M.; SCHNEIDER, V. E.; CEMIN, G.; POLETTO, M. Identifying the Areas at Risk of Huaico Occurrences in the Department of Lima, Peru. **Climate**, v. 13, n. 1, p. 11, 2025.

SATARZADEH, E.; SARRAF, A.; HAJIKANDI, H.; SADEGHIAN, M. S. Flood hazard mapping in western Iran: assessment of deep learning vis-à-vis machine learning models. **Natural Hazards**, v. 111, n. 2, p. 1355–1373, 2021.

SENAMHI – SERVICIO NACIONAL DE METEOROLOGÍA E HIDROGRAFÍA DEL PERU. **El Fenômeno El Niño en el Perú**: Plan Nacional de Gestión del Riesgo de Desastres – PLANAGERD 2014 - 2021 (Objetivo Nacional y Objetivo Estratégico 1). Lima: Impresiones y Servicios Generales TAWA, 2014. Disponível em: https://www.minam.gob.pe/wp-content/uploads/2014/07/Dossier-El-Ni%C3%B1o-Final_web.pdf. Acesso em: 11 jul. 2021.

SENAMHI – SERVICIO NACIONAL DE METEOROLOGÍA E HIDROGRAFÍA DEL PERU. **Fenômeno El Niño**: Anomalía de la temperatura superficial del mar. 2021. Disponível em: <https://www.senamhi.gob.pe/?p=fenomeno-el-nino>.

SHAHZAD, N; DING, X.; ABBAS, S. A Comparative Assessment of Machine Learning Models for Landslide Susceptibility Mapping in the Rugged Terrain of Northern Pakistan.

Applied Sciences, v. 12, n. 5, p. 2280, 2022.

SHEIKH, V.; KORNEJADY, A.; OWNEGH, M. Application of the coupled TOPSIS–Mahalanobis distance for multi-hazard-based management of the target districts of the Golestan Province, Iran. **Natural Hazards**, v. 96, p. 1335-1365, 2019.

SHI, X.; LI, Y.; ZHAO, Q. Flexible Hierarchical Gaussian Mixture Model for High-Resolution Remote Sensing Image Segmentation. **Remote Sensing**, v. 12, n. 7, p. 1219, 2020.

SHI, Y.; ZHANG, Z.; XUE, C.; FENG, Y. Machine Learning Prediction of Co-Seismic Landslide with Distance and Azimuth Instead of Peak Ground Acceleration. **Sustainability**, v. 16, n. 19, p. 8332, 2024.

SI, A.; ZHANG, J.; ZHANG, Y.; KAZUVA, E., DONG, Z., BAO, Y., & RONG, G. Debris Flow Susceptibility Assessment Using the Integrated Random Forest-Based Steady-State Infinite Slope Method: A Case Study in Changbai Mountain, China. **Water**, v. 12, n. 7, p. 2057, 2020.

SUDALAIMUTHU, K.; SUDALAYANDI, K. Development of a linear regression model to predict ground elevation from satellite elevation–statistical approach. *In: AIP Conference Proceedings*. [s.l.]: AIP Publishing, 2019, v. 2112, p. 020031.

SUHERMAT, M.; SUGIANTI, K.; YUNARTO, Y.; KUMORO, Y.; NUR, W. H.; SUKRISTİYANTI, S.; LESTIANA, H. Effectiveness of landslide susceptibility mapping using the maximum entropy model and weights of evidence modelling in the Kuningan Regency, West Java, Indonesia. **Rudarsko-geološko-naftni zbornik**, v. 39, n. 3, p. 27–42, 2024.

SUSILUOTO, J.; SPANTINI, A.; HAARIO, H.; HÄRKÖNEN, T.; MARZOUK, Y. Efficient multi-scale Gaussian process regression for massive remote sensing data with satGP v0.1.2. **Geoscientific Model Development**, v. 13, n. 7, p. 3439–3463, 2020.

ULLAH, K.; WANG, Y.; FANG, Z.; WANG, L.; RAHMAN, M. (2022). Multi-hazard susceptibility mapping based on Convolutional Neural Networks. **Geoscience Frontiers**, v. 13, n. 5, p. 101425, 2022.

UN - UNITED NATIONS. Sendai Framework for Disaster Risk Reduction 2015-2030. 2015. Disponível em: <https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030>.

UNDRR – UNITED NATIONS OFFICE FOR DISASTER RISK REDUCTION. Terminology. Disponível em: <https://www.undrr.org/terminology>.

UPRETI, A. **Machine Learning Application in G.I.S. and Remote Sensing: An Overview**. MDPI AG, 2022.

XAVIER, K. Rios perenes, efêmeros e intermitentes. 2020. Disponível em: <http://www.karinexavier.com.br/2019/01/rios-perenes-efemeros-e-intermitentes.html>.

XIONG, K.; ADHIKARI, B. R.; STAMATOPOULOS, C. A.; ZHAN, Y.; WU, S.; DONG,

Z.; DI, B. Comparison of Different Machine Learning Methods for Debris Flow Susceptibility Mapping: A Case Study in the Sichuan Province, China. **Remote Sensing**, v. 12, n. 2, p. 295, 2020.

YANG, L.; SHAMI, A. On hyperparameter optimization of machine learning algorithms: Theory and practice—Neurocomputing, v. 415, p. 295–316, 2020.

YESILNACAR, E.K. The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey. Ph.D Thesis, Department of Geomatics, University of Melbourne, Parkville, VIC, Australia, 2005; p. 423.

YUSRI, S.; RETNOWATI, E.; PRASTOWO, M.; IDRIS; FAKHRURROZI. Combining Participatory Mapping, Cloud Computing and Machine Learning for Mapping Climate Induced Landslide Susceptibility in Lembeh Island, North Sulawesi. **IOP Conference Series: Earth and Environmental Science**, v. 363, 2020.

ZHU, X.; GUO, H.; HUANG, J. J. Urban flood susceptibility mapping using remote sensing, social sensing, and an ensemble machine learning model. **Sustainable Cities and Society**, v. 108, p. 105508, 2024.

**APÊNDICE A – IDENTIFYING AREAS AT RISK OF HUAICOS OCCURRENCE IN
THE DEPARTMENT OF LIMA, PERU**

Article

Identifying the Areas at Risk of Huaico Occurrences in the Department of Lima, Peru

Geise Macedo dos Santos ¹, Vania Elisabete Schneider ^{2,3}, Gisele Cemin ⁴ and Matheus Poletto ^{1,*} 

¹ Postgraduate Program in Process Engineering and Technologies, University of Caxias do Sul, Caxias do Sul 95070-560, RS, Brazil; gmsantos5@ucs.br

² Graduate School, Trujillo Catholic University (UCT), Moche 13001, Peru; veschnei@ucs.br

³ Postgraduate Program in Civil Engineering, Federal University of Sergipe, Aracaju 49060-108, SE, Brazil

⁴ Postgraduate Program in Environmental Engineering and Sciences, University of Caxias do Sul, Caxias do Sul 95070-560, RS, Brazil; gcemin3@ucs.br

* Correspondence: mpolett1@ucs.br

Abstract: Because of local climate, a phenomenon called huaico occurs in the coastal regions of Peru, configured by an alluvial flow of surface runoff caused by precipitation and accompanied by the transport of solid particles. A total of 24% of the huaicos recorded in Peru from 2003 to 2019 were concentrated in the Department of Lima alone and affected 38,000 people. Thus, the aim of this study was to use Maxent to identify the areas at risk of huaicos in this department. To this end, a georeferenced database was created that included the locations of these events for modeling. We used variables suggested by Peru's Geological, Mining, and Metallurgical Institute (INGEMMET)—geology, geomorphology, DEM, slope, and precipitation—which returned extremely high kappa coefficients. Approximately 42% of Lima's area is likely to have a huaico occurrence. The most crucial variable for the models was the geomorphological classification characterized by the accumulation of mobilized material, as was the case in previous huaico models. In addition, the monthly approach should have been more effective at determining the differences in the precipitation levels. Thus, new models for the coastal departments of Peru using Maxent algorithms should take a new approach related to precipitation, although the use of Maxent proved satisfactory.

Keywords: Maxent; debris flow; maximum monthly precipitation; monthly approach



Academic Editor: Alessandro Pezzoli

Received: 16 November 2024

Revised: 17 December 2024

Accepted: 24 December 2024

Published: 5 January 2025

Citation: Santos, G.M.d.; Schneider, V.E.; Cemin, G.; Poletto, M. Identifying the Areas at Risk of Huaico Occurrences in the Department of Lima, Peru. *Climate* **2025**, *13*, 11. <https://doi.org/10.3390/cli13010011>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The coastal formation of Peru is basically a strip of plain formed by aeolian and alluvial deposits dating from the Cenozoic Era and Quaternary Period. Further inland are volcanic and sedimentary mountain formations with pyroclastic characteristics from the Mesozoic Era of the Cretaceous Period [1–3]. Vegetation density is low or nonexistent in areas covered by woodlands and coastal Andean deserts, respectively, exposing the soil or fragmented rock formations. This contributes to the transport of these materials down the mountains [4]. This configuration favors a specific type of natural disaster locally called a huaico, which corresponds to alluvial flows. A huaico is characterized by the surface runoff flows caused by precipitation in areas with steep slopes and is accompanied by the erosion and transport of solid particles from poorly cohesive soils or the fall and landslide of rocks [5,6]. The risk associated with huaicos lies in the location of the occurrence and the affected areas. Usually, huaicos reach the urbanized areas in the coastal zone of Peru, a region where around 65% of the country's population is concentrated [7].

Between 2003 and the first quarter of 2021 alone, 2421 huaicos were recorded in the country [8], affecting approximately 450,000 people, damaging approximately 37,000 homes, and causing 141 deaths. Although these are recurrent events, their location varies significantly over time. In the first quarter of 2021 alone, the department with the highest number of huaicos was the Department of Cusco, with 41 recorded. In the first quarter of 2019, 83 huaicos were recorded in the Lima Department, the region with the highest recurrence that year [8,9]. Despite the spatial unpredictability of occurrences, the Peruvian department most affected by huaicos is the Department of Lima, where the country's capital is located, which registered 24% of the huaicos between 2003 and 2021. During this period, 584 huaicos affected 38,100 people [8]. Figure 1 presents some examples of the areas impacted by huaicos in the Department of Lima.



Figure 1. Photographic records of huaicos in the Department of Lima. (a) Huaico that occurred in 2023 in an unspecified area of Lima [10]. (b) Huaico that also occurred in 2023 in Chaclayo [11]. (c) Huaico that occurred in 2023 in Punta Hermosa [12]. (d) Huaico that occurred in 2023 in an unspecified area of Lima [13].

The occurrence of huaicos is associated primarily with precipitation, which in the El Niño years reaches unusual levels in the country. During the recent coastal El Niño event in 2017, 570 huaicos were recorded—around 24% of the occurrences between 2003 and 2021. In 2019, another El Niño year, 492 huaicos were recorded in Peru [9].

With the recurrence of natural disasters, policies and practices must be adopted to reduce the risks resulting from natural hazards. The Peru National Disaster Risk Management Plan, valid from 2022 to 2030, identifies the country as vulnerable to huaicos. Furthermore, in adopting the guidelines of the Sendai Framework for Disaster Risk Reduction 2015–2030, which has, as its first guiding action, an obligation to minimize disaster risks by increasing our understanding of them [14], the plan also points out a deficiency in the country's familiarity with all the dimensions of disaster risks. One of the issues highlighted is the need for specific investigations with the potential for practical applications. Inseparable from the risks is the human occupation of the areas subject to natural hazards. In addition to their identification, structural and structuring actions must be followed to minimize the vulnerability of the resident population [6].

In recent years, machine learning methods have effectively mapped areas susceptible to natural hazards. In the literature, some reviews have categorized the use of machine

learning by hazard typology, such as landslides [15–17], floods [18,19], natural fires [20,21], and earthquakes [22]. For Peru, studies have evaluated only fires in the Amazon region [23], earthquake risk in Pisco [24], and susceptibility to landslides [25]. The machine learning algorithms most frequently used in these studies are the artificial neural network (ANN) and multilayer perceptron (MLP), both approaches using neurons, support vector machine (SVM), and random forest (RF).

Although Maxent was initially designed to model species distributions, the machine learning technique called maximum entropy modeling [26] has recently been used to map natural hazards. According to Phillips [26], the program calculates the probability of an event occurring by considering the input data characterizing the area and the occurrence records from regions with characteristics compatible with those that have undergone similar events. Javidan et al. [27] describe Maxent as a machine learning technique that enables investigating the relationship between a dependent variable (such as a landslide, flood, and others) and several independent variables (driving factors/geoenvironmental factors).

In risk mapping, Maxent determines the hazard distribution by finding the distribution with the highest degree of randomness within the assigned constraints, defined by the relationship between hazard circumstances at the input and their environmental covariates [28].

It is noteworthy that most studies using Maxent for mapping risk are in the Middle East and Asia. In these studies, as seen in Table 1, the recurrence of relief elements generated from a digital elevation model (DEM), such as slope, aspect, height above the drainage network, terrain ruggedness index (TRI), and topographic wetness index (TWI), in addition to the DEM itself, is observed. Geology and information related to the anthropic use of the areas, such as distance from roads and land use, are also recurring elements in the models. The climate is also addressed, especially when assessing fires, landslides, and floods.

Since a huaico results from specific climatic and geomorphological variables, identifying the environmental variables that are most likely to drive this type of hazard is necessary to build a model. Identifying areas where these events are likely to occur on a large scale is an essential planning tool. Such a model can be replicated for Peru and other Latin countries affected by huaicos, such as Colombia, Ecuador, and Venezuela.

Table 1. Articles that address the use of Maxent, with a focus on natural hazards risk mapping.

Hazard Type	Study Area	Most Used Variables
Landslides [27–38].	Iran (Gorganrood Watershed), Hindu Kush Himalaya, Iran (Ziarat watershed), China (Wanyuan), Iran (Taleghan Basin), Iran (Golestan Province), North Sulawesi (Lembeh Island), European Seas, Republic of Korea (Gangwon-do), China (Xulong Gully), China (Guozigou Valley), North Pakistan	Altitude, aspect, slope, distance from permanent watercourses, road distance, geology, curvature profile, curvature plane, precipitation, land use, topographic wetness index (TWI), and terrain ruggedness index (TRI)
Erosion [32,39–42].	Iran (Golestan am Basin), Iran (Fars Province), United States (Fort Cobb watershed), Italy (Oltrepo Pavese)	Altitude, height above nearest drainage, aspect, slope, distance from roads, permanent watercourses, geology, precipitation, temperature, and land use

Table 1. *Cont.*

Hazard Type	Study Area	Most Used Variables
Fires [28,43–46].	Hindu Kush Himalaya, Italy (Aosta Valley), Iran (Mazandaran Province), Northeast Spain, South Korea	Altitude, aspect, slope, precipitation (average, minimum), temperature (average, maximum), land use, and wind
Floods [27,28,32,47–49].	Iran (Gorganrood Watershed), Hindu Kush Himalaya, Iran (Kermanshah City), Iran (Sari City), Iran (Golestan Province), Iran (Saliantapeh Catchment)	Altitude, aspect, slope, CN (curve number), distance from permanent watercourses, precipitation, and land use

Thus, this work aims to use the Maxent machine learning algorithm, which is already used to identify potential natural hazard areas, to identify areas at risk of huaicos in the Department of Lima. Lima has been heavily affected by huaicos in recent years, making it ideal for testing the potential of using this algorithm. In addition, it is essential to evaluate the changes in the models related to the different average maximum precipitation levels. Thus, an approach using monthly models was adopted. A set of models was created that considered the driving variables described in the literature and other recurrent variables to identify their correlation with the occurrence of huaicos.

The use of machine learning in general and Maxent specifically to model huaicos is novel, as it is a monthly approach for hazardous area identification. There are no studies related to modeling huaico-prone areas and no specifications for precipitation data.

2. Materials and Methods

This section identifies the study area and describes the input data and the procedures applied to develop the models in Maxent.

2.1. Study Area

The Department of Lima, where Peru's capital is located, experienced 24% of the huaicos registered between 2003 and 2019 [8]. In addition to the study area location, the map in Figure 1 shows 1650 huaico locations, as inventoried by Peru's Geological, Mining, and Metallurgical Institute (INGEMMET) [50], during the period from 2003 to 2021.

The Department of Lima is classified as Hot Desert Climate by the Köppen–Geiger classification [51]. The National Meteorological and Hydrological Service of Peru (SENAMHI) sub-characterizes five different climates in the Department of Lima, varying from arid with moisture deficiency throughout the year in the coastal zone to rainy with a dry autumn and winter in the most continental area of the department. In the arid climate, precipitation is 5 mm annually; in the rainy climate, precipitation can reach 1200 mm. The average maximum temperature varies from 19 °C to 31 °C [52].

Figure 2 illustrates that the data are not randomly allocated, with the concentration of huaicos near the capital, Lima. This work identifies the environmental variables coinciding with historical records.

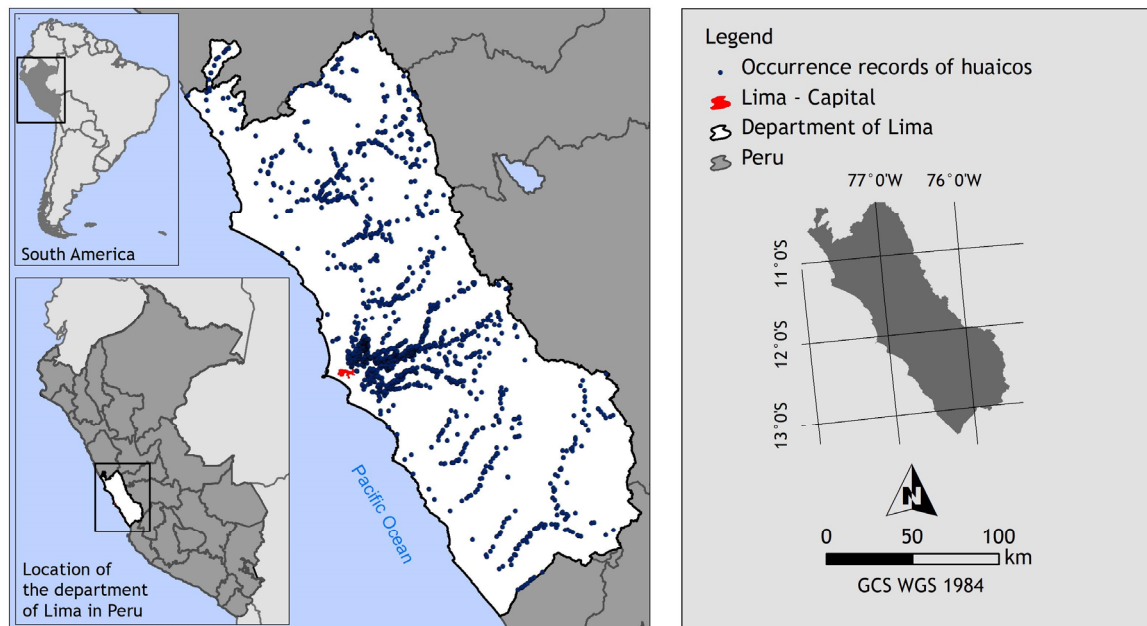


Figure 2. Map of the study area and locations of huaico occurrences.

2.2. Input Data for the Algorithm

2.2.1. Data from the Inventory of Geological Hazards in Peru

The Geological, Mining, and Metallurgical Institute (INGEMMET) maintains metadata catalog records of geological hazards in Peru, including inventories of falls, landslides, flows, complex movements, overturns, and others. This inventory identifies huaicos as debris flows (INGEMMET, 2023) [50]. The records are in point format. However, there is no description of how the points were selected when the phenomenon described occurs over a large area.

The geological hazards file was filtered to encompass the Department of Lima using QGIS software 3.30.2. Then, using the WGS 1984 system, the geographic coordinates of the occurrences were extracted and converted into CSV format (comma-separated values) to be input as data into Maxent. Note that in any case, Maxent only accepts points as input data.

It is worth noting that the Maxent algorithm works only in the presence of an event, and it is unnecessary to indicate areas where the event did not occur. However, the model must consider some pixels as a practical absence to calculate model statistics [30,48].

Maxent relies on a machine learning response that makes predictions from incomplete data. This results in a Maxent output in ASCII format, which is a continuous prediction of specific presence ranging from 0 to 1 [27,30].

2.2.2. Environmental Variables

The characterization of areas that experience huaicos in the Department of Lima, derived from INGEMMET data, is based on the driving factors for the occurrence of these hazards. These include the precipitation in areas where ephemeral water courses form, the geological and geomorphological formation itself, and the relief of the region, both the altitudinal conformation of the terrain and its slope [53–57].

Therefore, these characteristics were expected to be the most prominent in the models for identifying areas at risk of huaicos occurrence. However, other environmental variables already cited in the literature were also added to the modeling, as described in the following sections.

It was also essential to evaluate the correlation between environmental variables before applying them to the models. This evaluation was carried out using a specific tool within the QGIS environment.

Digital Elevation Model

Kornejady, Ownegh, and Bahreman [29] treat altitude as the most straightforward relief feature, and they agree with Pourghasemi et al. [39] regarding its importance in landslides. In general, topography contributes to the formation of hazard-prone environments. Topographic factors extend beyond the altitude itself to slope and aspect. The topographic wetness index (TWI) quantifies the impact of topography on hydrological processes and is related to landslides [38]. The terrain roughness index (TRI) measures the terrain's roughness, and is therefore another index associated with landslides.

Although the literature does not refer to huaicos, a phenomenon that could be understood as debris flows, but to landslides and floods, Rusk et al. [24] determined that topography and its covariates influence soil stability and water movement, among other factors.

This work used the Shuttle Radar Topography Mission (STRM) DEM with 30 m resolution pixels, dating from 2000, obtained through the United States Geological Survey (USGS).

The digital elevation model of the study area in QGIS served as input data for modeling. From the DEM, slope data in degrees were also generated in QGIS using the slope tool. The aspect measuring the orientation of each cell relative to north; the topographic wetness index (TWI), which indicates the tendency of the terrain to accumulate water; and the terrain roughness index (TRI) were used. Specific tools in QGIS also generated the aspect data and the terrain roughness index (TRI). The topographic wetness index corresponds to the natural logarithm of the ratio between the runoff accumulation area that drains into a pixel and the tangent of the pixel's slope. The TWI was obtained through QGIS's mathematical operations with the DEM.

Climate Data

Sheikh, Kornejady, and Ownegh [32] used the average annual precipitation to assess landslide and flood risk. Arabameri et al. [41] also used yearly average precipitation to assess erosion risk. In assessing fire risk, Adab et al. [44] considered the average annual rainfall and the average precipitation of the driest month. For this study, using monthly precipitation rather than yearly precipitation was more credible owing to prior knowledge of the area, which distinguished between rainy months (especially in El Niño years) and dry months. For this reason, since precipitation is usually low in the region, historical maximum values were used because an average value would underestimate precipitation volumes. Again, according to the climate classification created by SENAMHI in the Department of Lima, precipitation averages 5 mm annually and in the rainy climate. Although more continental areas of the department can reach precipitations of 1200 mm, it is possible to associate the occurrence of huaicos with arid regions [52]. Therefore, in the arid climate, all months are virtually the driest, while in the rainy environment, winter and autumn are dry [52]. Nevertheless, the average monthly precipitation was computed to compare the average and maximum data.

The National Meteorological and Hydrological Service of Peru (SENAMHI) provides historical rainfall data for Peru. The Department of Lima has 46 rainfall stations, 36 of which have at least 40 years of historical data. Information on maximum and historical monthly averages for each station was generated using electronic spreadsheets.

Subsequently, the maximums were spatially distributed through Thiessen polygons in a GIS environment, generating a set of precipitation coverage polygons per season for each month. The use of Thiessen polygons was presented as the first evaluation option.

It is worth noting that other approaches to spatialization of precipitation can be tested and compared.

Geomorphological Data

The literature frequently asserts that the area's geology has a remarkable influence on the stability of rock formations when assessing the risk of landslides and erosion. In Peru, INGEMMET makes geological data available at a scale of 1:50,000 and was updated in 2022 [58].

Although geomorphological features are related to the occurrence of huaicos, geomorphological and hydrogeological data have not been used in the literature. Hydrogeology, in turn, was evaluated because it is related to soil saturation, which can also influence soil/rock stability. Geomorphological data at a scale of 1:250,000, dating from 2016, and hydrogeological data at a scale of 1:2,000,000, also from 2016, were obtained from INGEMMET [59].

Soil typology was obtained from the Ministry of the Environment (MINAM) at a scale of 1:2,000,000 and dates from 2009 [60]. Soil properties, such as shear strength and permeability, are related to the propensity for sliding and water movement in the surface portion [28].

Using QGIS software, all data were filtered to include only the Lima Department.

Land Use Data

Files were prepared to characterize land use, showing distances from urbanized areas, watercourses, and ecosystems. The hydrography was obtained from the Geo Peru National Georeferenced Data Platform at a scale of 1:100,000, prepared by the National Geographic Institute.

Urbanized areas were obtained by classifying Landsat 2023 images provided by the United States Geological Survey (USGS). The classification was conducted in QGIS using an unsupervised method, pixel by pixel.

The linear distance of each cell from the point of interest was then generated in QGIS for the hydrography and urbanized areas. Although huaicos occur due to the generation of ephemeral surface flows and in natural environments, the distances between urbanized areas and hydrography can suggest a possible indirect relationship between these factors. Although huaicos are characterized by the generation of ephemeral flows, these occur both in basins completely devoid of surface flows and in basins with perennial watercourses. If this region has the potential for huaicos to occur, the flow generated by the huaicos will contribute to the watercourse. It is worth noting that it is during this gravitational movement that the flows generated by the huaicos reach the urbanized areas that are in their path.

The 2019 MINAM ecosystem information was used on a scale of 2:200,000. Although at first glance it is not directly related to the occurrence of huaicos, this information may allow us to identify a correlation.

2.3. Maxent

To insert the data into Maxent, all environmental variables were transformed into a matrix format (Geotiff), ensuring that all data had the same size, number of pixels, and positioning. The data compatibility process was executed in QGIS software, where the initial vector format data were transformed into matrix data. The matrix data underwent reprojection of the coordinate system and datum (WGS UTM 18S) and were filtered to display only the Department of Lima.

Of the records of huaico occurrences in the Department of Lima, 70% were randomly selected, with the remaining 30% reserved for model validation, according to the methodology proposed by Arabameri et al. [41] and Javidan et al. [27].

Maxent has a user-friendly interface where it is simple to input and configure data. The samples (70% of the records) comprising the data on the occurrence of huaicos in the Department of Lima were input into CSV format.

After inserting the environmental variables into ASCII format, each ecological variable must be considered as either continuous or categorical data. The remaining information corresponds to the model configuration. The model configuration was performed according to the guidelines established by Dalapicolla [61], suggesting the adoption of a random test percentage of 30%, bootstrap model replica-running, 15 replicas, and a maximum number of 10,000 iterations.

The data insertion process in Maxent followed the steps shown in Figure 3.

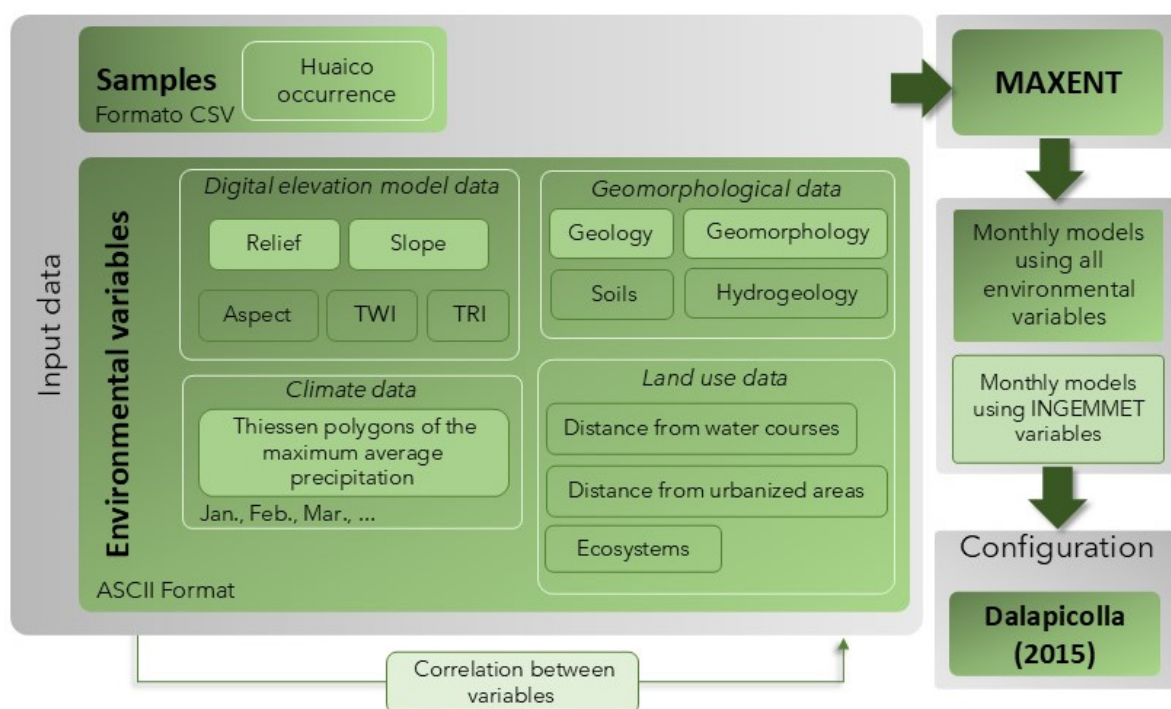


Figure 3. Flowchart of the modeling process.

A set of monthly models was organized, that is, one model for each month, considering all the environmental variables (DEM, slope, aspect, TRI, TWI, geology, geomorphology, hydrogeology, soils, distance from water courses, distance from urbanized areas, ecosystems, and monthly precipitation) and another set of monthly models (one model for each month) for only the variables described by INGEMMET (DEM, slope, geology, geomorphology, and precipitation) as drivers in the huaico formation process.

The data from Maxent were analyzed with the kappa index for each model and the overlap of each area with the risk of occurrence in each monthly model for the other months. The continuous probability model (0 to 1) was reclassified based on the logistic threshold of a training presence of 10 percentiles.

3. Results and Discussion

3.1. Environmental Variables

The results obtained from the correlation between environmental variables in a GIS environment showed the maximum correlation for each variable related to itself, resulting

in the value of 1. The only high correlations, above 0.7 [29], were between the average maximum precipitations. The correlations did not exceed 0.4 among the other variables, indicating acceptable correlation levels below 0.7. A model containing more than one mean maximum precipitation will include a high correlation between these variables, which, according to Kornejady, Ownegh, and Bahremand [29], should be avoided to minimize the complexity of the model. In any case, Maxent remains minimally affected by the correlation of variables, even with small samples [48].

Regarding the physical characteristics of the study area (Figure 4), the altitude, according to the digital elevation model of the Department of Lima, ranges from 0 to 6401 m. The variation from 0 to 1000 m is located in the coastal region of the department, with approximately 30% of the Department of Lima within this altitude range. Moving inland, the altitude of the department increases. However, a closer look at the department's relief reveals the “quebradas”, which are micro basins where significant variations in altitude occur on a small scale and where superficial flows are formed. These are found along the coastal region and the valleys of the perennial rivers of the department. The slope is also steeper in these regions, becoming flatter at the bottoms of the valleys and in the coastal region, which can be called a floodplain. Throughout the department, the slope varies from 0° to 89.95° , with 21% of the areas having a slope up to 10° and 83% having a slope up to 50° . The aspect ratio in the Department of Lima varies from 0° to 360° , with visual orientation patterns in the “quebradas”. The TRI (terrain ruggedness index) varies from 0 to 454, being lower in the lower and flatter areas. This TRI ranges from level to moderately rough, with the index reaching levels of extreme ruggedness. The TWI (topographic wetness index) varies from 0 to 20.01, with the highest values in the lower areas and less steep portions of the terrain, thus indicating that these areas are the most suitable for water accumulation and possible flooding.

The Department of Lima has 38 geological classifications, which are grouped into 14 macro-classifications in Figure 5.

On the other hand, the Department of Lima has 40 hydrogeological categories. Similar to the geology, these categories were grouped into classifications to facilitate visualization. Regional geomorphology has 45 classifications. It should be noted that the department's 38 geological classifications, 40 hydrogeological classifications, and 45 geomorphological classifications were used for modeling. There are only seven pedological classifications in the Department of Lima.

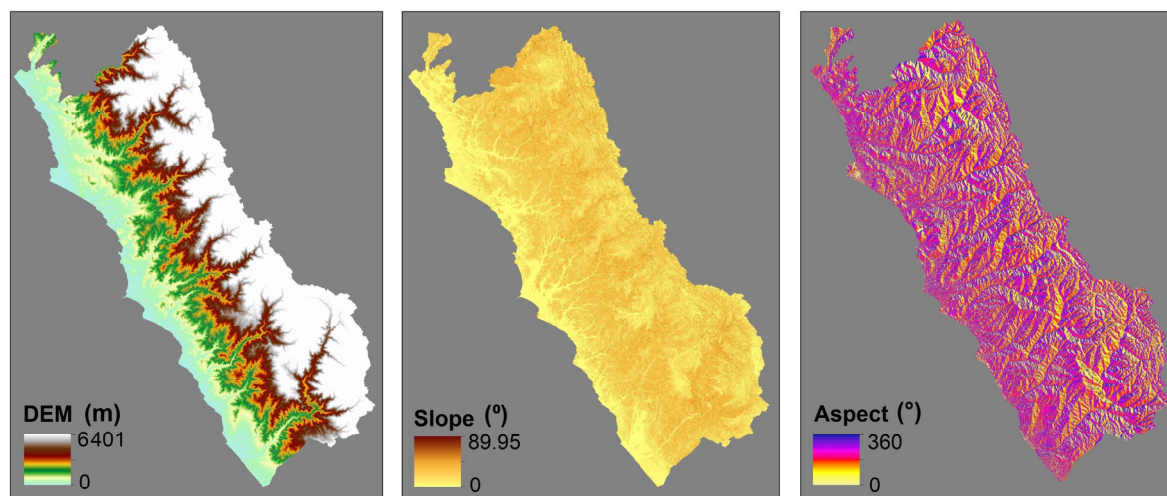


Figure 4. Cont.

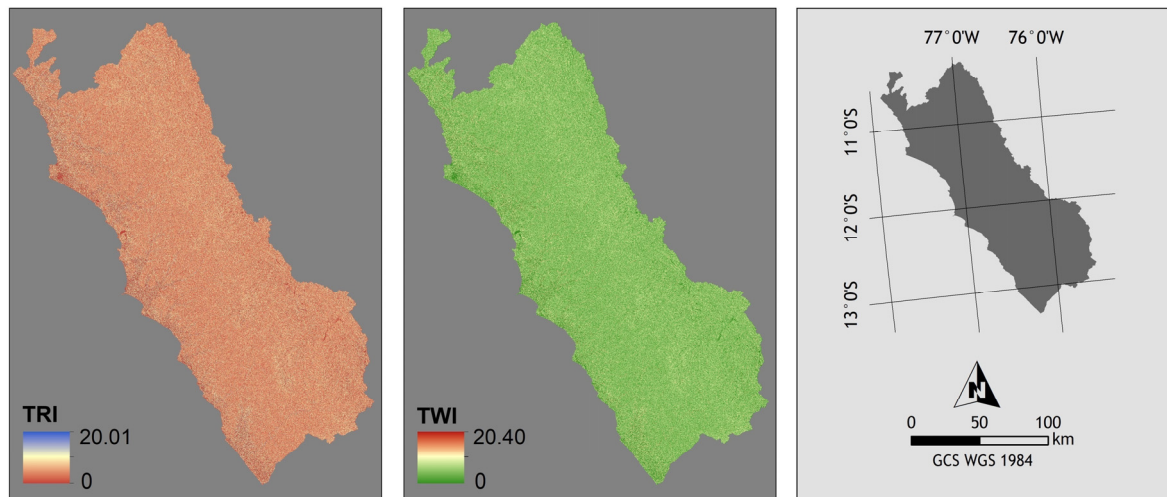


Figure 4. Environmental relief variables.

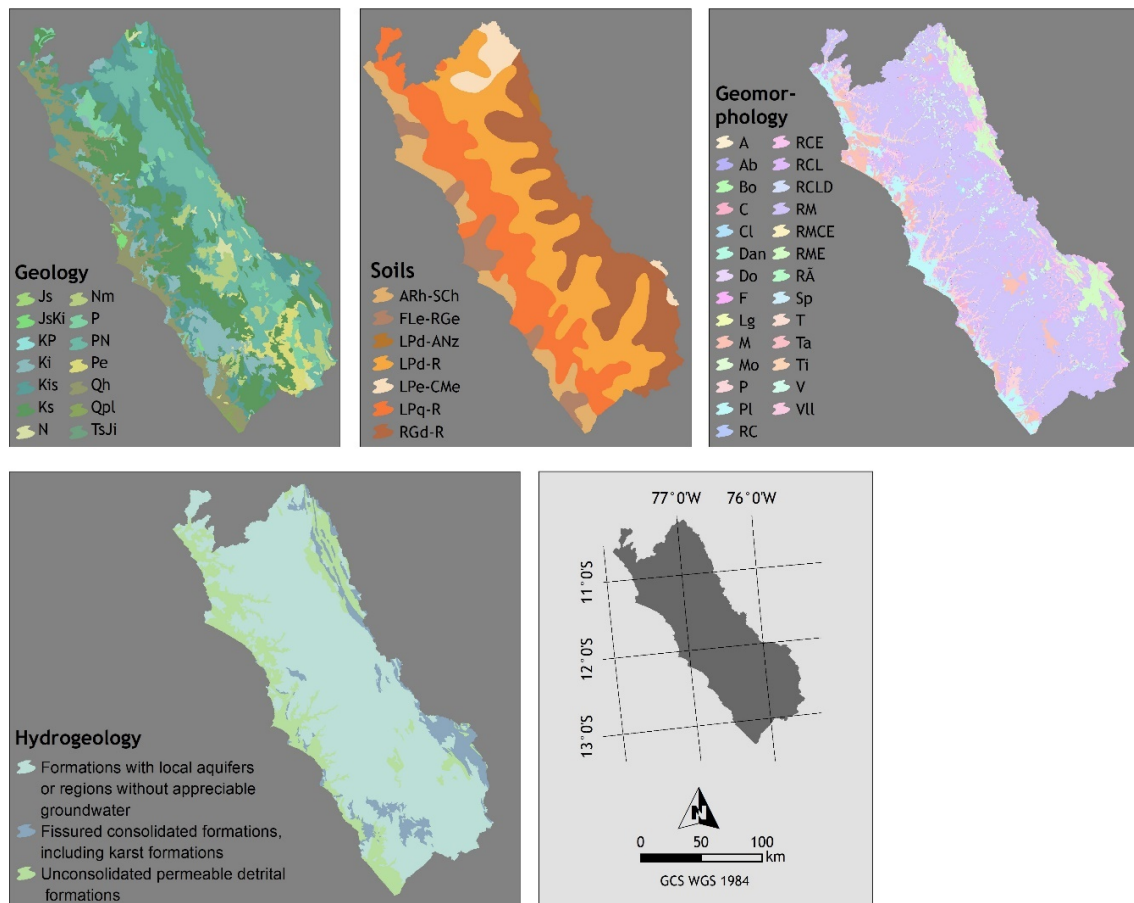


Figure 5. Geomorphological environmental variables. Geology map legend: Js—Upper Jurassic marine, JsKi—Upper Jurassic Lower Cretaceous, volcanic sediments, KP—Cretaceous Paleocene, tonalite, granodiorite, Ki—Lower Cretaceous, KG, Kis—Cretaceous, marine and volcanic sediments, Ks—Cretaceous, super units, N—Neogene, Nm—Neogene Miocene, P—Paleozoic, PN—Paleogene Neocene, Pe—Paleogene Eocene, Qh—Quaternary Holocene, Qpl—Quaternary Pleistocene and TsJi—Upper Triassic, Lower Jurassic. Soil map legend: ARh-SCh—Arenosol haplic—Solonchak haplic, FLe-RGe—Fluvisol eutric—Regosol eutric, LPd-ANz—Leptosol dystic—Andosol vitreous, LPd-R—Leptosol districo—Lithic outcrop, LPe-CMe—Leptosol eutric—Cambisol eutric, LPq-R—Lithic Leptosol—Lithic outcrop and RGd-R—Regosol districo—Lithic outcrop. Geomorphology map legend: A—lagoon, Ab—mountain foothill fan, Bo—bofedales (wetlands), C—dune field and coast, Cl—coastal ridge, Dan—anthropogenic

deposit, Do—salt dome, F—coastal strip, Lg—water bodies, M—volcanic plateau, Mo—glacial debris, P—mountain slope or foothill, Pl—floodplain, RC—rock hill, RCE—structural hill, RCL—hill and ridge, RCLD—dissected hill and ridge, RM—mountain, RMCE—mountain and structural hill, RME—structural mountain, RĀ—riverbed, Sp—swamp system, T—marine terrace, Ta—alluvial terrace, Ti—undifferentiated terrace, V—slope, and VII—valley.

The caption in Figure 5 presents the correspondence between the code for each class and the class denomination as a whole.

Figure 6 illustrates that the Euclidean distance from urbanized areas varies from 0 to 35.7 km, and the perennial watercourses in the department vary from 0 to 25.6 km. Regarding ecosystems, the Peruvian Pacific Coastal Desert classifications are observed. These are characterized by a hyper-arid climate devoid of vegetation, consisting of sandy soils and rocky outcrops. In addition, the classes of Serranía Esteparia (mountain range steppe) and Puna stand out for their herbaceous vegetation.

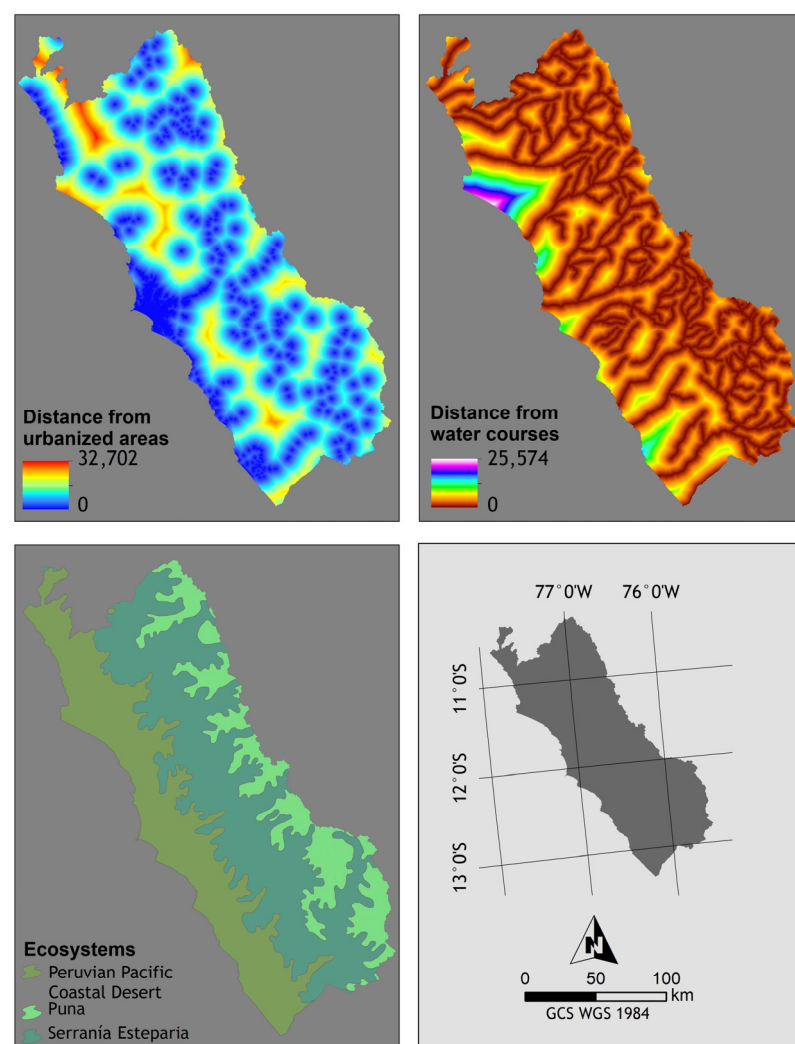


Figure 6. Environmental variables of land use and land cover.

Figure 7 shows that some meteorological stations experienced historical precipitation maximums of as little as 1 mm in June, July, and August. In contrast, there are regions in the department where, during these same months, the average maximums reached 308 mm. In any case, February has historically had the highest average maximums for the stations with the lowest average maximums (52 mm). For the stations with the highest average maximums, precipitation during February was much larger (974 mm). March, October, and

December also have quite pronounced historical average maximums. These observations are validated by the data shown in Figure 8, which indicates the average values for the 36 rainfall stations analyzed. They highlight the comparison between the averages and the maximums, where the average levels represent a maximum of 44% of the maximum historical levels in March, reaching only 13% of the maximum levels historically recorded in the department during May.

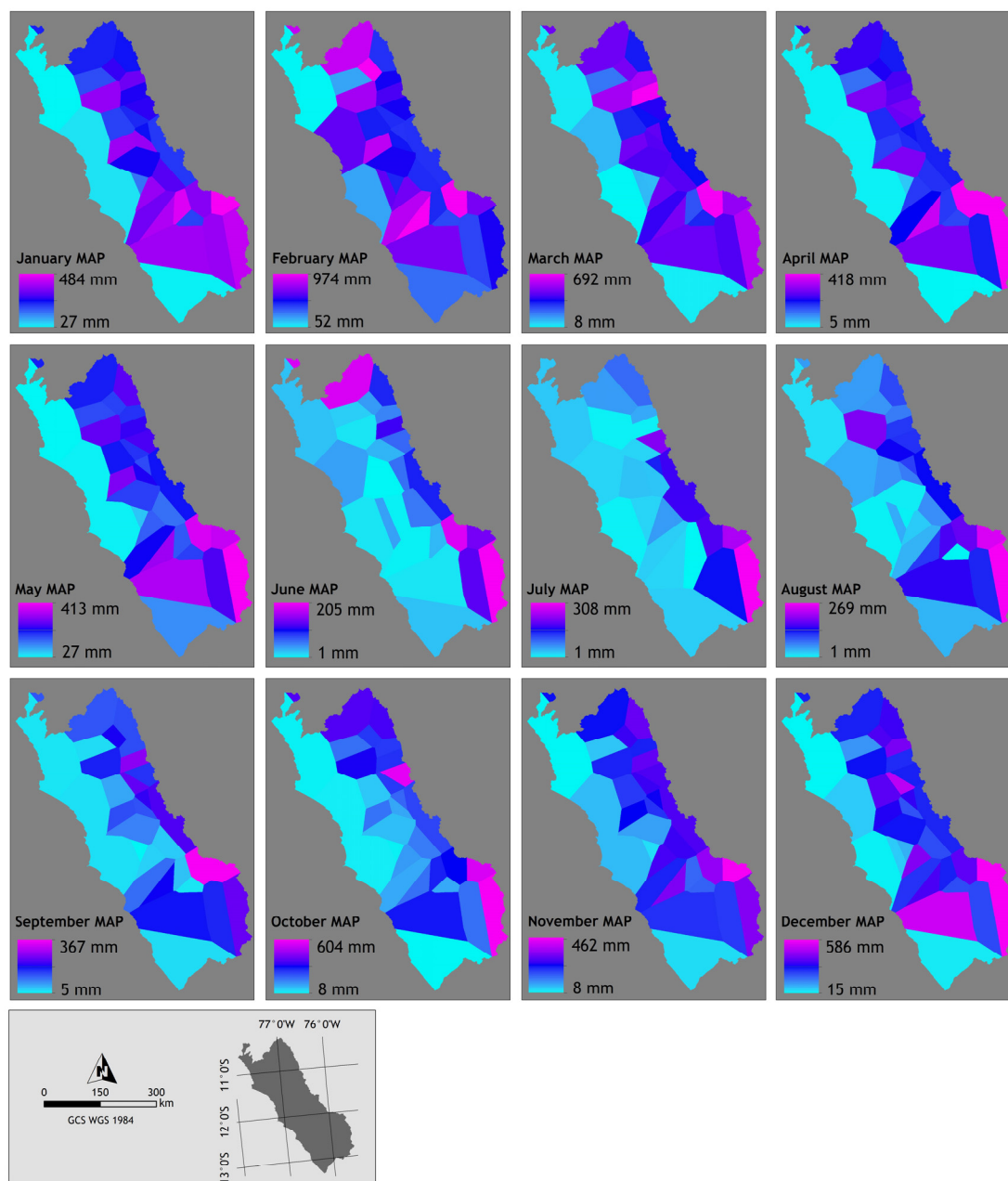


Figure 7. Monthly maximum average precipitation (MAP).

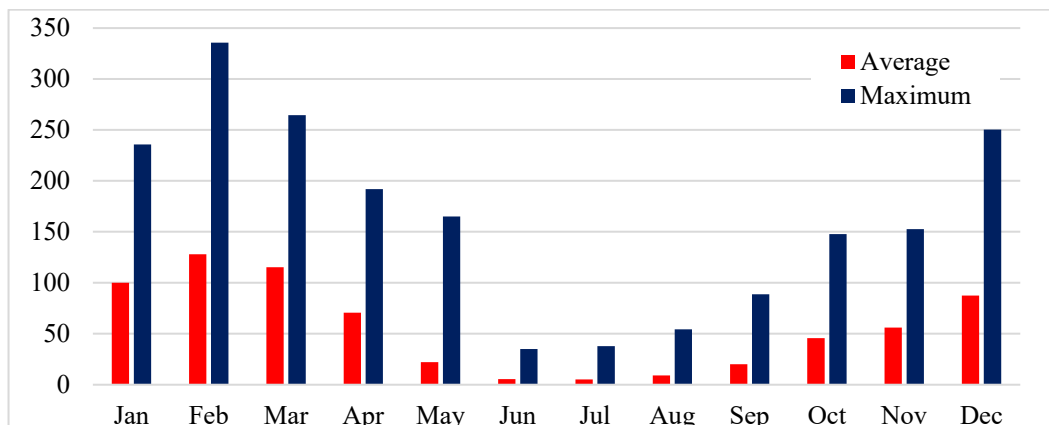


Figure 8. Average precipitation behavior at the rain gauge stations in the Department of Lima.

Regarding the increase in precipitation levels in Peru, the El Niño phenomenon responsible for warming the Pacific waters stands out [62]. For example, in the cycles of 1982–1983, 1997–1998, and 2017, precipitation levels were extraordinary, exceeding 1000 mm over a three-month cycle, triggering huaicos and floods that led to considerable economic losses and even deaths [62].

Gonzales and Ingol [41] observed the particular formation of the 2017 Coastal El Niño that directly and powerfully struck Peru’s coastal region. The authors concluded that it was a succession of events. In their own words, the Pacific Decade Oscillation in its positive phase gave rise to coupled ocean–atmosphere processes generating the ENSO Modoki and La Niña Modoki, and the coastal ENSO was a sub-process of the La Niña Modoki, the lower part of 2017’s ENSO Modoki. The Modoki event represents a unique pattern of tripolar pressure at sea level [63]. Strong El Niño events result in higher levels of precipitation and areas affected by precipitation, which consequently increase the occurrence and recurrence of huaicos.

Events such as the 2017 Coastal El Niño are difficult to predict and prepare for. Including the data in the modeling allows for an evaluation of how the Coastal El Niño intensity influences the occurrence of the huaicos.

3.2. Modeling in Maxent

Table 2 shows the quality of information of the Maxent models. All models have adequate quality, both in the “Test AUC”, which must be greater than 0.7, and in the “10 percentile training presence test omission”, which should not exceed 0.2. The area under the curve ranges from 0 to 1, with 1 indicating the highest accuracy. The AUC ranges of 0.5–0.6, 0.6–0.7, 0.7–0.8, 0.8–0.9, and 0.9–1 classify the models as poor, average, good, very good, and excellent, respectively, according to Yesilnacar [64], with our models classified as very good.

The kappa coefficient represents five classes— $k < 0.4$, $0.4 < k < 0.55$, $0.55 < k < 0.85$, $0.85 < k < 0.99$, respectively—with poor, average, good, excellent, and very excellent validation [65]. In the kappa test, the models generated with all variables performed excellently. However, the models generated were based only on the environmental variables described by INGEMMET as influential in the formation of huaicos, and the kappa index was almost maximum, classified as very excellent.

Figures 9 and 10 present the results of the models, reclassified to represent the areas with a probability of huaico occurrence.

Table 2. Quality of models.

Month	Environmental Variables	AUC Test	Ten Percentile Training Presence Test Omission	Kappa Coefficient	Environmental Variables	AUC Test	Ten Percentile Training Presence Test Omission	Kappa Coefficient
Jan	DEM, slope, aspect, TWI, TRI, geology, geomorphology, hydrogeology, soils, ecosystems, distance from watercourses, distance from urbanized areas, precipitation	0.8802	0.1291	0.79	DEM, slope, geology, geomorphology, precipitation	0.8289	0.1142	0.98
Feb		0.8802	0.1291	0.80		0.8221	0.1140	0.98
Mar		0.8787	0.1302	0.79		0.8257	0.1187	0.97
Apr		0.8792	0.1282	0.79		0.8181	0.1072	0.97
May		0.8794	0.1271	0.80		0.8225	0.1111	0.98
Jun		0.8958	0.1270	0.80		0.8156	0.1228	0.98
Jul		0.8773	0.1407	0.79		0.8229	0.1214	0.98
Aug		0.8861	0.1184	0.79		0.8262	0.1130	0.98
Sep		0.8787	0.128	0.79		0.8187	0.114	0.98
Oct		0.8809	0.1336	0.79		0.8304	0.0848	0.98
Nov		0.8793	0.1269	0.79		0.8197	0.1244	0.98
Dec		0.8798	0.1299	0.79		0.8287	0.1195	0.98

On average, 17% of the areas in the Department of Lima have a probability of huaico occurrence when considering the models using all the variables. July and August had the lowest probability of huaicos occurring, while June and September had the highest likelihood of huaicos occurring. However, these higher areas in June and September seem inconsistent with the average maximum precipitation of these months, which is lower than those with the highest precipitation.

When analyzing the models generated using the INGEMMET variables identified as drivers of huaico occurrence, the models' quality was better, and the areas with a probability of huaico occurrence were higher. On average, 42% of the Department of Lima has a likelihood of occurrence of huaicos. Using variables that do not effectively contribute to the occurrence of huaicos made the models complex, causing the algorithm to search for nonexistent relations between the sample data and the environmental variables. This resulted in lower model quality than those using influential contributing variables.

Regarding the coincidence between the monthly models, Table 3 shows the coincidence matrix of the areas with a probability of occurrence for each model, taking the months horizontally as a reference. On average, there is a coincidence of 84% between the monthly models. The greater the correspondence between the models throughout the year, the greater the chance that even with precipitation variation, the areas with a probability of occurrence will be basically the same.

Regarding the variables that contribute the most to the models, Figure 11 shows that the distance from water courses is the factor that contributes most to the monthly models, followed by geomorphology, digital elevation model, distance from urbanized areas, slope, and geology.

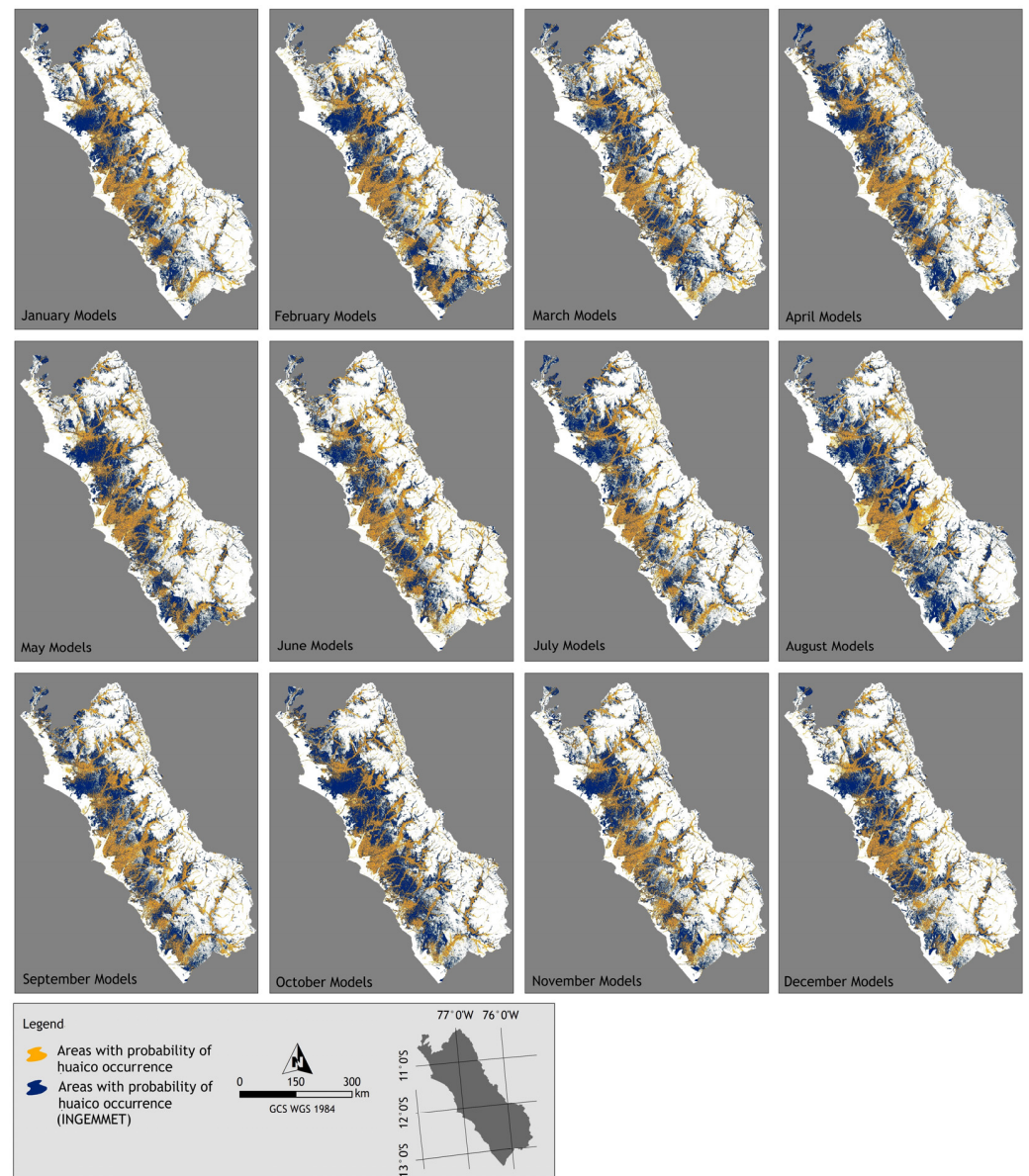


Figure 9. Probability results considering monthly models.

Table 3. Coincidence between monthly models (%).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Jan	100	89	84	89	88	67	84	66	91	86	90	87
Feb	85	100	83	85	86	65	83	64	92	83	86	85
Mar	87	91	100	82	86	98	98	68	88	80	84	79
Apr	89	89	83	100	88	66	84	66	93	87	88	86
May	87	90	87	88	100	68	82	67	92	84	86	87
Jun	85	86	127	84	87	100	84	91	87	84	85	84
Jul	90	93	108	90	88	71	100	69	94	90	70	89
Aug	84	85	89	85	86	92	83	100	86	83	84	84
Sep	84	88	83	85	85	63	80	62	100	83	86	83
Oct	89	90	84	90	87	68	87	67	94	100	90	88
Nov	89	90	84	87	86	66	65	65	93	86	100	87
Dec	90	91	83	89	90	68	86	67	93	88	91	100

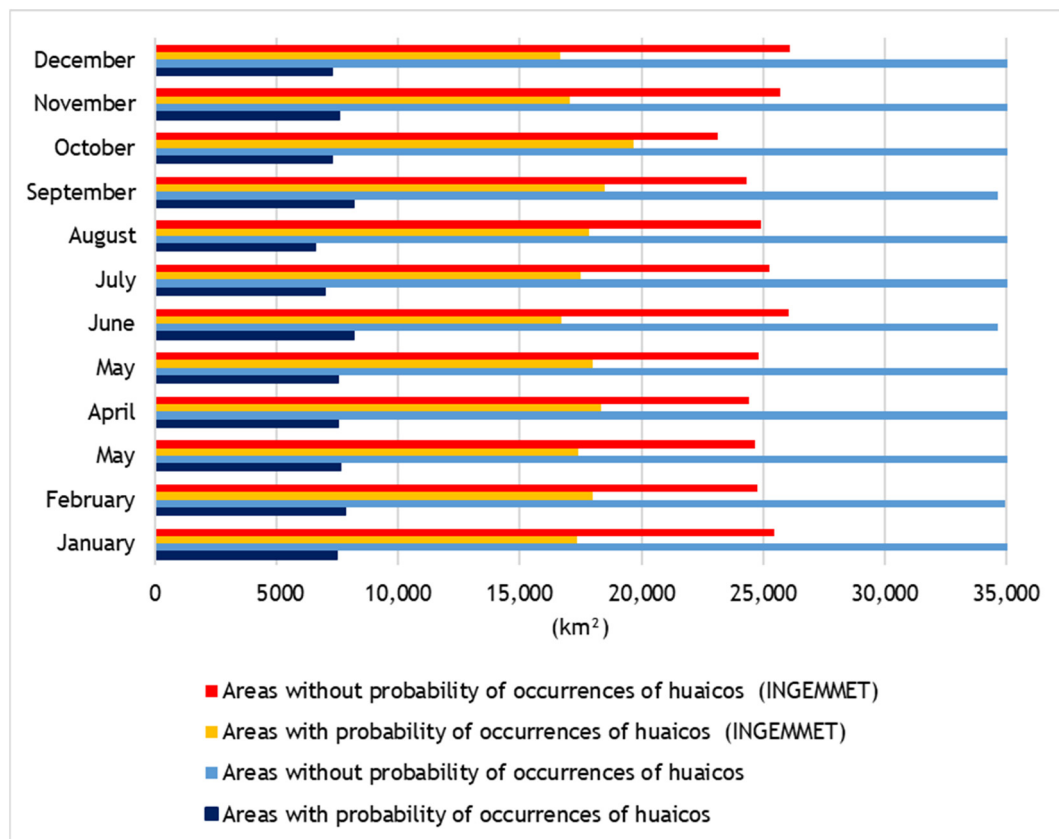


Figure 10. Areas with and without the probability of huaico occurrences.

In general, there is a predominance of the probability of occurrence of huaicos up to 5 km from watercourses. The distance from urbanized areas coincides with the distance from rivers, as they are located primarily in the bottoms of valleys, close to waterways. Thus, the occurrence of huaicos is associated mainly with a distance of up to 2.5 km from urbanized areas.

The geomorphology representing the probability of occurrence includes mountains in intrusive rocks, mountains in volcanic rocks, mountains in volcanic-sedimentary rocks, and alluvial–torrential slopes or foothills.

Mountains in intrusive rock correspond to outcrops of intrusive rocks, igneous rocks with high degrees of fracturing due to tectonic processes that, together with the steep slopes associated with this formation, increase rockfalls [54]. Mountains in volcanic rocks are outcrops of volcanic rocks with steep slopes susceptible to rockfalls and collapses, and that may present intense fracturing and weathering [66]. Mountains in volcanic-sedimentary rocks also have steep slopes associated with large landslides, huaicos, and rock avalanches [54]. The geomorphological classification of alluvial–torrential slope or foothill is characterized by the accumulation of material mobilized by previous huaicos, which locally modify the direction of river courses and are located at the mouths of streams toward the main rivers [55].

Regarding the relief, the areas where huaicos are likely to occur are in the valleys, which corroborates precisely with the shorter distances from water courses. Regarding slope, the occurrences are related to inclinations of up to 30 degrees, which in themselves represent extremely high slopes of approximately 65%.

However, it is important to note that the variable contributing the most also has the greatest distance from a normal distribution.

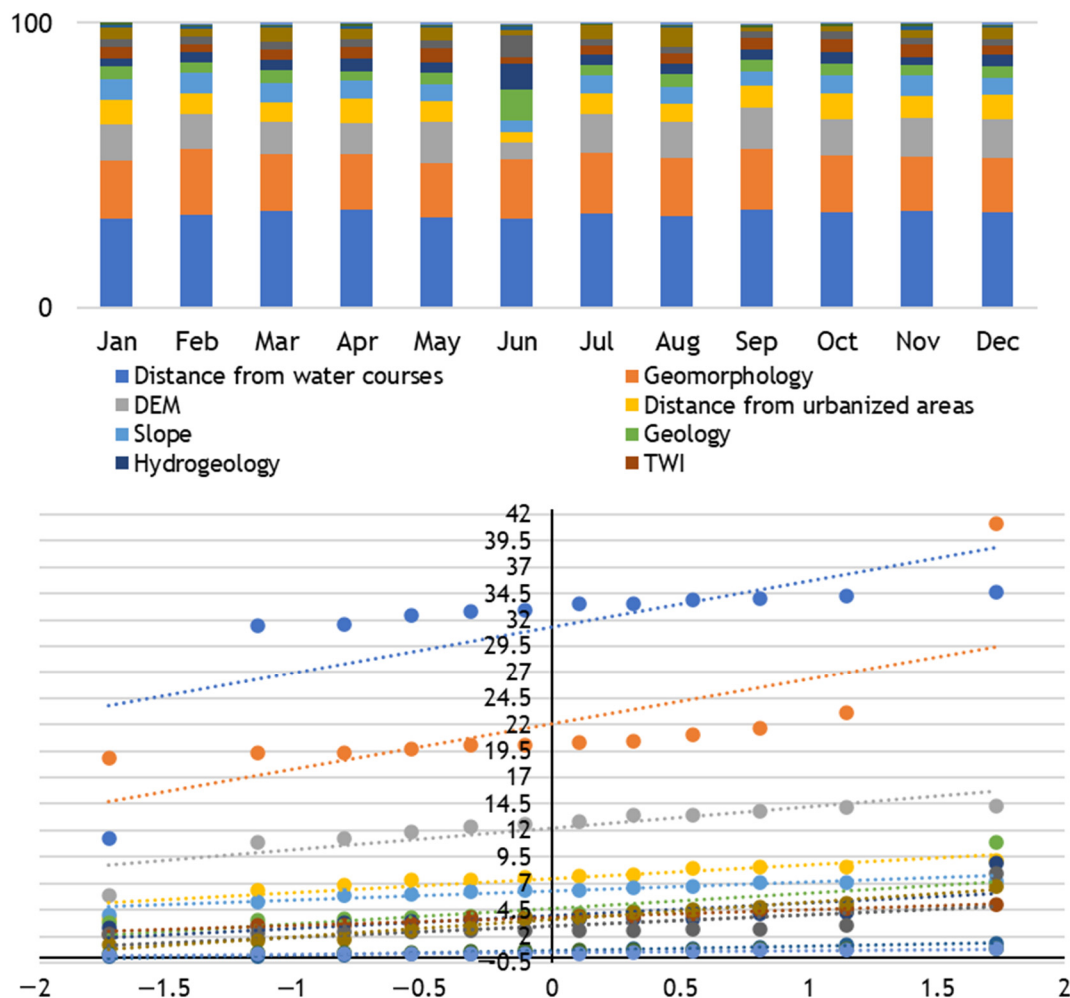


Figure 11. Contribution of variables for each monthly model considering all environmental variables and the behavior of the data compared to normal distribution.

In the monthly models that consider only the variables specified by INGEMMET as relevant to the occurrence of huaicos, all months presented a similar distribution of contributions from environmental variables, with the most significant contribution from geomorphology in the classes described above. The second most important contributing factor is relief, which also follows the ranges described in the characterization of the results of the previous model. The other factors, geology, slope, and average monthly precipitation, have the least influence in the models and similar influences (Figure 12).

Unlike the models using all variables, in the model using only the variables considered influential in the occurrence of huaicos according to INGEMMET, all variables have a normal distribution, contributing to their better quality and representativeness.

Kornejady, Ownegh, and Bahremand [29] and Phillips [26] warn against interpreting the contributions of factors in models, since Maxent can find the best solution through different paths, which may result in various factor contributions. In other words, the contribution cannot be considered an indicator of the importance of the environmental variable. The jackknife test provides more significant information regarding the importance of variables. The test requires removing a factor from the modeling, using a variable alone, and using all variables [29]. The jackknife tests responded almost equally to the contributions of the variables in all the models. Thus, in this work, it is safe to regard the relations of the variable contributions as a measure of the importance of the variables in predicting the occurrence of huaicos.

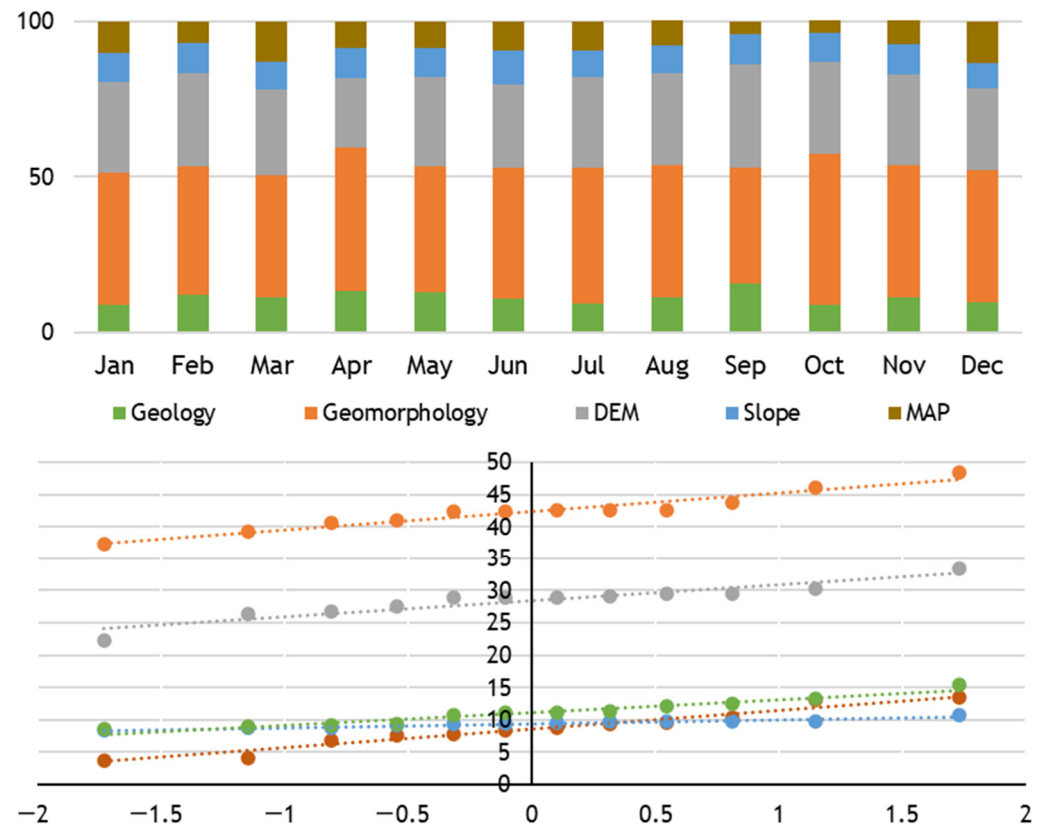


Figure 12. Contribution of variables for each monthly model, considering the environmental variables indicated by INGEMMET and the behavior of the data compared to normal distribution.

Although identifying areas where huaicos occur does not correspond to identifying areas where landslides and floods occur, Rusk et al. [28] evaluated both hazards and adopted variables similar to those in this work. For landslides, factors related to relief and geomorphology had a high contribution. However, unlike this work, for Rusk et al. [28], precipitation contributed significantly to the model for both types of hazards: approximately 35% for floods and 25% for landslides. The results for landslides by Javidan et al. [27] point in the same direction, where relevant geomorphological factors, such as slope and altitude alone, did not interfere with the model.

Although the occurrence of huaicos does not directly depend on the distance from watercourses, there is a concentrated probability of huaico occurrence in the vicinity of rivers in the models that used all variables, unlike the results of Javidan et al. [28]. They found the distance from watercourses to be an essential predictive factor due to its importance in the speed and magnitude of flooding, given the high flow concentration in the drainage network, increasing the occurrence of floods near streams. However, the relation is comprehensible because some “quebradas” belong to perennial water course watersheds.

Chen et al. [30] also indicated a relationship between landslides and the proximity of watercourses and roads in more significant portions of the terrain. The authors also highlighted the distance from roads in their models, an anthropic factor that modifies the natural stability of the studied areas and threatens the infrastructure, significantly decreasing the quality of the model. Anthropogenic factors do not influence huaico occurrences. This is due to the fact that the formation of huaicos does not occur in inhabited areas. The areas where huaicos form are areas of steep slope, difficult access, few inputs, that is, of low general attractiveness for the spread of urbanization and crops. In short, they are areas with desert characteristics. Still, the impacts are intensified by the anthropic factors of urbanized areas and infrastructure. Since, the expansion of urbanization in the flatter areas

corresponds precisely to the areas where the flows generated by the huaicos are directed by gravity.

Similar to the results of Javidan et al. [27], there is a lower incidence of huaicos in high-altitude locations. This finding was based on the results of the elevation response curves, which demonstrate that huaicos, like floods, occur in low-lying areas and plains. It should be noted that the runoff from the basins that form the huaicos converges typically into a single channel, where it can either reach an urbanized area or follow a floodplain to the ocean and then reach an urbanized area.

Javidan et al. [27] found that higher precipitation levels increase landslides (650–690 mm annually) and lower rainfall levels cause erosion. For huaicos, an increase in precipitation should result in a larger area with a probability of huaicos occurring. However, the rise in precipitation did not respond linearly in the models. In general, precipitation did not contribute significantly to this phenomenon. Therefore, the model constitution does not permit us to infer the levels of rainfall that trigger the huaicos.

Giráldez et al. [67] evaluated the extreme climate events of precipitation in metropolitan Lima from 1965 to 2013. The authors found that 16 mm of precipitation occurred in 1970, which was sufficiently strong to flood the streets, causing power failures, damaging the Jorge Chávez International Airport facilities, destroying about 2000 homes, and triggering landslides and rivers to overflow. Although there are no specific data on huaicos, their occurrence during this extreme event is almost certain. This could mean that almost all the precipitation data used in this work would be enough to trigger a huaico occurrence.

In contrast, Rusk et al. [28] point out that climate change can alter model responses to increase the potential for disasters. A further step in taking action to prevent and mitigate the impacts of disasters is identifying the social vulnerability of the affected populations.

3.3. Uncertainty

It is important to note that the records of huaico occurrences themselves represent a degree of uncertainty in the model. Since the occurrence data used as input for the model were in point format, as were the available occurrence data, there are no records of the discrete areas affected by huaicos, and there is no way to insert areas as input data in Maxent. This limits the generation of a model that reflects reality as closely as possible. There is also no guarantee that the points identified as areas of huaico occurrence were located correctly.

Furthermore, the other input data have different scales of elaboration. The hydrogeology and soil data, in particular, at scales of 1:2,000,000, present generic data about the spatiality of this information. This does not prevent us from evaluating a large area, such as the Department of Lima, but in the models where these data are inserted, they can decrease the quality of the models. For this same reason, in the models considering only the variables indicated as influential by INGEMMET in the occurrence of huaicos, the absence of hydrogeology and soil values reduces the uncertainty in these models. The existence of more detailed data could alter both the final response of the models and the relevance of these data within the analysis. This could improve the quality levels of the models, although, it could have no effect on the overall quality of these as well.

4. Conclusions

Using variables not considered to be drivers of huaicos made the models complex and underestimated the probability of occurrence. Moreover, the monthly approach was ineffective when distinguishing differences among the precipitation levels. In general, precipitation did not contribute significantly to this phenomenon. Thus, the constitution of these models does not permit inference of the precipitation levels that trigger the huaicos.

Therefore, new models for the coastal departments of Peru, in Maxent algorithms and other machine learning algorithms, will only use the INGEMMET driver variables, and a single precipitation datum will be used. The question remains as to whether the annual precipitation data can fit the models better than a month's average data. Another alternative yet to be tested is using other machine learning algorithms in huaico occurrence susceptibility.

Inevitably, climate change in this region points toward an increase in temperature that will likely lead to an increase in precipitation events, thereby increasing the number of huaicos and affected areas, although the models in this work were not yet able to identify the contribution of precipitation in the huaico occurrences. Because these models use past data, a priori predictions of future changes would not alter the construction of the model. However, future work could incorporate projected precipitation data as an input variable for comparison with those based on historical data. Firstly, the use of precipitation data should be adjusted as suggested previously.

Maxent returned extremely high levels of correspondence between the models and the actual registered occurrence data. In this panorama, almost 42% of the area of the Department of Lima exhibits a probability of huaico occurrence. There is no reasonable expectation of reducing the areas with the potential for huaicos to occur. These areas could be zoned according to public policies in which disaster preparations could be adopted. Furthermore, these areas could be evaluated to avoid occupying or even vacate them. Regarding public policies, residents should be continuously educated about the risks of huaicos, and structural measures, such as systems to dam the flow, can be implemented to minimize the impacts of huaicos.

Author Contributions: G.M.d.S. was responsible for conceptualization, methodology, data development, formal analysis, and writing—original draft preparation. V.E.S. and G.C. were responsible for formal analysis and writing—review and editing. M.P. was responsible for writing—review and editing and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Moncayo, O.P. *Rasgos Generales Sobre la Geología del Perú*; INGEMMET: Lima, Peru, 1998.
2. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Estudio de Los Recursos Minerales del Perú*; Franja n° 4: Lima, Peru, 2005.
3. Medina, L.; Villacorta, S.; Zavala, B.; Vílchez, M.; Núñez, S.; Luque, G.; Calderón, E. Caracterización geomorfológica del Norte peruano. In *Boletín de la Sociedad Geológica del Perú*; INGEMMET: Lima, Peru, 2015; Volume 110, pp. 205–208.
4. Ministerio del Ambiente del Gobierno del Perú—MINAM. *Mapa Nacional de Ecosistemas del Perú*; Memoria: Louisville, KY, USA, 2019.
5. Rivera, M.; Vilchez, M.; Vela, J. Peligros por huaicos en la ciudad de Arequipa. In *Fortalecimiento de Capacidades Para Mitigar Los Impactos de Huaicos en Perú*; Porras, M.R., Phillips, J.C., Armijos Burneo, M.T., Contreras, R.A., Eds.; INGEMMET: Lima, Peru, 2018; pp. 45–49.
6. Perú. *Plan Nacional de Gestión del Riesgo de Desastres (PLANAGERD) 2022–2030*; Instituto Nacional de Defensa Civil—INDECI: Lima, Peru, 2022.
7. French, A.; Mechler, R.; Arestegui, M.; Macclune, K.; Cisneros, A. Root causes of recurrent catastrophe: The political ecology of El Niño-related disasters in Peru. *Int. J. Disaster Risk Reduct.* **2020**, *47*, 101539. [[CrossRef](#)]
8. Instituto Nacional de Defensa Civil—INDECI. *Compendio Estadístico del Instituto Nacional de Defensa Civil*; INDECI: Lima, Peru, 2021.
9. Instituto Nacional de Defensa Civil—INDECI. *Compendio Estadístico del Instituto Nacional de Defensa Civil*; INDECI: Lima, Peru, 2019.
10. Federación de Periodistas del Perú. Lluvias y Huaicos: Al Menos 13 Distritos de Lima Sufren de Inundaciones. Available online: https://fpp.org.pe/_trashed-91/ (accessed on 10 December 2024).
11. Diario La Otra Cara. Lima Este Inundada por los Huaicos y las Intensas Lluvias. Available online: <https://diariolaotracara.com/lima-este-inundada-por-los-huaicos-y-las-intensas-lluvias/> (accessed on 10 December 2024).

12. Willax. Huaicos en Lima: Alcalde de Punta de Hermosa Señala que “No hay Presupuesto” para Evitar el Colapso de Casa de Playa. Available online: https://willax.pe/actualidad/huaicos-en-lima-alcalde-de-punta-de-hermosa-senala-que-no-hay-presupuesto-para-evitar-el-colapso-de-casa-de-playa#google_vignette (accessed on 10 December 2024).
13. CanalB pe. Lima Fueron los Más Golpeados Nuevamente por Huaicos. Available online: <https://canalb.pe/noticias/actualidad/distritos-del-este-y-el-sur-de-lima-fueron-los-mas-golpeados-nuevamente-por-huaicos> (accessed on 10 December 2024).
14. United Nations—UN. *Sendai Framework for Disaster Risk Reduction 2015–2030*; United Nations Office for Disaster Risk Reduction: Geneva, Switzerland, 2015.
15. Tehrani, F.S.; Calvello, M.; Liu, Z.; Zhang, L.; Lacasse, S. Machine learning and landslide studies: Recent advances and applications. *Nat. Hazards* **2022**, *114*, 1197–1245. [\[CrossRef\]](#)
16. Ado, M.; Amitab, K.; Maji, A.K.; Jasińska, E.; Gono, R.; Leonowicz, Z.; Jasiński, M. Landslide susceptibility mapping using machine learning: A literature survey. *Remote Sens.* **2022**, *14*, 3029. [\[CrossRef\]](#)
17. Bao, H.; Zeng, C.; Peng, Y.; Wu, S. The use of digital technologies for landslide disaster risk research and disaster risk management: Progress and prospects. *Environ. Earth Sci.* **2022**, *81*, 446. [\[CrossRef\]](#)
18. Mosavi, A.; Ozturk, P.; Chau, K. Flood prediction using machine learning models: Review. *Water* **2018**, *10*, 1536. [\[CrossRef\]](#)
19. Li, J.; Fang, Z.; Zhang, J.; Huang, O.; He, C. Mapping basin-scale supply-demand dynamics of flood regulation service—A case study in the Baiyangdian Lake Basin, China. *Ecol. Indic.* **2022**, *139*, 108902. [\[CrossRef\]](#)
20. Bot, K.; Borges, J.G. A systematic review of applications of machine learning techniques for wildfire management decision support. *Inventions* **2022**, *7*, 15. [\[CrossRef\]](#)
21. Alkhatib, R.; Sahwan, W.; Alkhatieb, A.; Schütt, B. A brief review of machine learning algorithms in forest fire science. *Appl. Sci.* **2023**, *13*, 8275. [\[CrossRef\]](#)
22. Ridzwan, N.S.M.; Yusoff, S.H.M. Machine learning for earthquake prediction: A review (2017–2021). *Earth Sci. Inform.* **2023**, *16*, 1133–1149. [\[CrossRef\]](#)
23. Ferreira, I.J.M.; Campanharo, W.A.; Barbosa, M.L.F.; da Silva, S.S.; Selaya, G.; Aragão, L.E.O.C.; Anderson, L.O. Assessment of fire hazard in Southwestern Amazon. *Front. For. Glob. Change* **2023**, *6*, 1107417. [\[CrossRef\]](#)
24. Izquierdo- Horna, L.; Zevallos, J.; Yopez, Y. An integrated approach to seismic risk assessment using random forest and hierarchical analysis: Pisco, Peru. *Heliyon* **2022**, *8*, e10926. [\[CrossRef\]](#)
25. Kumar, C.; Walton, G.; Santi, P.; Luza, C. An Ensemble Approach of Feature Selection and Machine Learning Models for Regional Landslide Susceptibility Mapping in the Arid Mountainous Terrain of Southern Peru. *Remote Sens.* **2023**, *15*, 1376. [\[CrossRef\]](#)
26. Phillips, S.J.; Anderson, R.P.; Dudík, M.; Schapire, R.E.; Blair, M. Opening the black box: An open-source release of Maxent. *Ecography* **2017**, *40*, 887–893. [\[CrossRef\]](#)
27. Javidan, N.; Kavian, A.; Pourghasemi, H.R.; Conoscenti, C.; Jafarian, Z.; Rodrigo-Comino, J. Evaluation of multi-hazard map produced using MaxEnt machine learning technique. *Sci. Rep.* **2021**, *11*, 6496. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Rusk, J.; Maharjan, A.; Tiwari, P.; Chen, T.H.K.; Shneiderman, S.; Turin, M.; Seto, K.C. Multi-hazard susceptibility and exposure assessment of the Hindu Kush Himalaya. *Sci. Total Environ.* **2022**, *804*, 150039. [\[CrossRef\]](#)
29. Kornejady, A.; Ownegh, M.; Bahreman, A. Landslide susceptibility assessment uses a maximum entropy model with two data sampling methods. *Catena* **2017**, *152*, 144–162. [\[CrossRef\]](#)
30. Chen, W.; Pourghasemi, H.R.; Kornejady, A.; Zhang, N. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma* **2017**, *305*, 314–327. [\[CrossRef\]](#)
31. Mokhtari, M.; Abedian, S. Spatial prediction of landslide susceptibility in Taleghan basin, Iran. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1297–1325. [\[CrossRef\]](#)
32. Sheikh, V.; Kornejady, A.; Ownegh, M. Application of the coupled TOPSIS–Mahalanobis distance for multi-hazard-based management of the target districts of the Golestan Province, Iran. *Nat. Hazards* **2019**, *96*, 1335–1365. [\[CrossRef\]](#)
33. Yusri, S.; Retnowati, E.; Prastowo, M. Combining Participatory Mapping, Cloud Computing and Machine Learning for Mapping Climate Induced Landslide Susceptibility in Lembeh Island, North Sulawesi. In *IOP Conference Series; Earth and Environmental Science*; New York, NY, USA, 2019; Volume 363. [\[CrossRef\]](#)
34. Innocenti, C.; Battaglini, L.; D’Angelo, S.; Fiorentino, A. Submarine landslides: Mapping the susceptibility in European seas. *Q. J. Eng. Geol. Hydrogeol.* **2020**, *54*, qjeh2020–qjeh2027. [\[CrossRef\]](#)
35. Kim, H.G.; Park, C.Y. Landslide susceptibility analysis of photovoltaic power stations in Gangwon-do, Republic of Korea. *Geomat. Nat. Hazards Risk* **2021**, *12*, 2328–2351. [\[CrossRef\]](#)
36. Liu, Y.; Xu, P.; Cao, C.; Shan, B.; Zhu, K.; Ma, Q.; Zhang, Z.; Yin, H. A comparative evaluation of machine learning algorithms and an improved optimal model for landslide susceptibility: A case study. *Geomat. Nat. Hazards Risk* **2021**, *12*, 1973–2001. [\[CrossRef\]](#)
37. Liu, Y.; Zhao, L.; Bao, A.; Li, J.; Yan, X. Chinese high-resolution satellite data and gis -based assessment of landslide susceptibility along Highway G30 in Guozigou Valley using logistic regression and maxent model. *Remote Sens.* **2022**, *14*, 3620. [\[CrossRef\]](#)

38. Shahzad, N.; Ding, X.; Abbas, S. A comparative assessment of machine learning models for landslide susceptibility mapping in the rugged terrain of northern Pakistan. *Appl. Sci.* **2022**, *12*, 2280. [CrossRef]
39. Pourghasemi, H.R.; Yousefi, S.; Kornejady, A.; Cerdà, A. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci. Total Environ.* **2017**, *609*, 764–775. [CrossRef] [PubMed]
40. Botero-Acosta, A.; Chu, M.L.; Guzman, J.A.; Starks, P.J.; Moriasi, D.N. Riparian erosion vulnerability model based on environmental characteristics. *J. Environ. Manag.* **2017**, *203*, 592–602. [CrossRef] [PubMed]
41. Arabameri, A.; Asadi Nalivan, O.; Saha, S.; Roy, J.; Pradhan, B.; Tiefenbacher, J.P.; Thi Ngo, P.T. Novel ensemble approaches of machine learning techniques in modeling the gully erosion susceptibility. *Remote Sens.* **2020**, *12*, 1890. [CrossRef]
42. Bosino, A.; Giordani, P.; Quénéhervé, G.; Maerker, M. Assessment of calanchi and rill–interrill erosion susceptibilities using terrain analysis and geostochastics: A case study in the Oltrepo Pavese, Northern Apennines, Italy. *Earth Surf. Process. Landf.* **2020**, *45*, 3025–3041. [CrossRef]
43. Vacchiano, G.; Foderi, C.; Berretti, R.; Marchi, E.; Motta, R. Modeling anthropogenic and natural fire ignitions in an inner-alpine valley. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 935–948. [CrossRef]
44. Adab, H.; Atabati, A.; Oliveira, S.; Gheshlagh, A.M. Assessing fire hazard potential and its main drivers in Mazandaran province, Iran: A data-driven approach. *Environ. Monit. Assess.* **2018**, *190*, 670. [CrossRef] [PubMed]
45. Martín, Y.; Zúñiga-Antón, M.; Mimbrero, M.R. Modeling temporal variation of fire occurrence towards the dynamic prediction of human wildfire ignition danger in northeast Spain. *Geomat. Nat. Hazards Risk* **2018**, *10*, 385–411. [CrossRef]
46. Lim, C.H.; Kim, Y.S.; Won, M.; Kim, S.J.; Lee, W.K. Can satellite-based data substitute for surveyed data to predict the spatial probability of forest fire? A geostatistical approach to forest fire in the Republic of Korea. *Geomat. Nat. Hazards Risk* **2019**, *10*, 719–739. [CrossRef]
47. Darabi, H.; Choubin, B.; Rahmati, O.; Haghighi, A.T.; Pradhan, B.; Kløve, B. Urban flood risk mapping using the GARP and QUEST models: A comparative study of machine learning techniques. *J. Hydrol.* **2019**, *569*, 142–154. [CrossRef]
48. Norollahi, M.; Kaboli, H.S. Urban flood hazard mapping using machine learning models: GARP, RF, MaxEnt, and NB. *Nat. Hazards* **2021**, *106*, 119–137. [CrossRef]
49. Rajabi, A.; Shabanlou, S.; Yosefvand, F.; Kiani, A. Exploring the sample size and replication scenarios effect on spatial prediction of flood, using MARS and MaxEnt methods case study: Saliantape catchment, Golestan, Iran. *Nat. Hazards* **2021**, *109*, 871–901. [CrossRef]
50. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Mapa de Inventario de Peligros Geológicos*; INGEMMET: Lima, Peru, 2019.
51. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]
52. Servicio Nacional de Meteorología e Hidrología del Perú—SENAMHI. Mapa Climático del Perú. Available online: <https://www.senamhi.gob.pe/?p=mapa-climatico-del-peru> (accessed on 15 December 2024).
53. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Zonas Críticas Por Peligros Geológicos en la Cuenca Hancay-Lambayeque*; INGEMMET: Lima, Peru, 2005.
54. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Peligros Por Movimientos en Masa en la Quebrada Tinaja (Dist. Pachacamac-Cieneguilla, Provincia y Región Lima)*; INGEMMET: Lima, Peru, 2012.
55. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Evaluación de Peligros Geológicos Por Movimientos en Masa en el Anexo de Colpa*; INGEMMET: Lima, Peru, 2019.
56. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Inspección Geológica Realizada en la Quebrada San Lázaro—Arequipa*; INGEMMET: Lima, Peru, 2020.
57. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Evaluación de Peligros Geológicos Por Flujo de Detritos (Huaicos) en la Quebrada Zaparo*; INGEMMET: Lima, Peru, 2021.
58. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Carta Geológica Nacional a Escala 1:50000*; INGEMMET: Lima, Peru, 2022.
59. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Mapa Geomorfológico del Perú*; INGEMMET: Lima, Peru, 2016.
60. Ministerio del Ambiente del Gobierno del Perú—MINAM. *Mapa Nacional de Suelos del Perú*; INGEMMET: Lima, Peru, 2009.
61. Dalapicolla, J. Tutorial de Modelos de Distribuição de Espécies: Guia Prático Usando o Maxent e o Arcgis 10. In *Laboratório de Mastozoologia e Biogeografia*; Universidade Federal do Espírito Santo: Vitória, Brazil, 2016.
62. Servicio Nacional de Meteorología e Hidrología del Perú—SENAMHI. Ministerio del Ambiente *El Niño: Histórico de Lluvias*; Lima, Peru. 2021. Available online: <https://www.senamhi.gob.pe/?p=escenarios-lluvia> (accessed on 17 September 2024).
63. Gonzales, E.; Ingol, E. Determination of a New Coastal ENSO Oceanic Index for Northern Peru. *Climate* **2021**, *9*, 71. [CrossRef]
64. Yesilnacar, E.K. The Application of Computational Intelligence to Landslide Susceptibility Mapping in Turkey. Ph.D Thesis, Department of Geomatics, University of Melbourne, Parkville, VIC, Australia, 2005; p. 423.
65. Monserud, R.A.; Leemans, R. Comparing global vegetation maps with the Kappa statistic. *Ecol. Model.* **1992**, *62*, 275–293. [CrossRef]

66. Instituto Geológico, Minero y Metalúrgico—INGEMMET. *Evaluación de Peligros Geológicos en el Sector de Octuyo-Manás*; INGEMMET: Lima, Peru, 2021.
67. Giráldez, L.; Silva, Y.; Flores-Rojas, J.L.; Trasmonte, G. Diagnosis of the Extreme Climate Events of Temperature and Precipitation in Metropolitan Lima during 1965–2013. *Climate* **2022**, *10*, 112. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

**APÊNDICE B – OCORRÊNCIA DE HUAICOS NO PERU: UMA ANÁLISE DOS
FATORES DESENCADEANTES**

Ocorrência de huaicos no Peru: Uma análise dos fatores desencadeantes

Geise Macedo dos Santos

Programa de Pós-Graduação em Engenharia de
Processos e Tecnologias Universidade de Caxias do
Sul, Brasil. gmsantos5@ucs.br

Vania Elisabete Schneider

Professor Visitante da Universidade Católica de
Trujillo, Trujillo, Peru & Programa de Pós-
Graduação em Engenharia Civil, Universidade
Federal de Sergipe, Brasil. veschnei@ucs.br

Gisele Cemin

Programa de Pós-Graduação em Engenharia e
Ciências Ambientais, Universidade de Caxias do
Sul, Brasil. gcemin3@ucs.br

Matheus Poletto

Programa de Pós-Graduação em Engenharia de
Processos e Tecnologias, Universidade de Caxias do
Sul, Brasil. mpolett1@ucs.br

Palavras-chave: Relevo; geologia; geomorfologia.

Abstract

In Peru, a phenomenon occurs triggered, among other factors, by El Niño through increased precipitation. These are debris flows, locally called huaicos, associated with large socioeconomic losses over the years. In addition to precipitation, other physical factors may be related to the occurrence of huaicos. Understanding the interrelationship between factors can assist in natural disaster risk prevention policies. Therefore, in this work we evaluated the relationship between the occurrences of huaicos inventoried in Peru and factors identified in the literature as influencing the occurrence of the phenomenon. For this purpose, geographic information systems were used. As a result, it was observed that the slope of the form presented was not suitable for the evaluation and the altitude concentrates 56% of the huaicos up to 2,000 meters. Regarding geology, 13% of occurrences are related to the Quaternary Holocene Continental classification. From geomorphology, the Alluvial-Torrential Slope class stands out with 24% of the total occurrences and a further 16% of occurrences are related to the geomorphological classification of Mountain in sedimentary rock. Although there are indications of physical factors that influence the occurrence of huaicos, the relationship between the inventory of huaico occurrences and these factors proved to be a major challenge in validating a pattern. The occurrence of huaicos is a phenomenon that requires the consideration of more triggering factors, such as precipitation. Furthermore, it requires more robust tools and methodologies to find patterns of occurrence that can be used to identify other areas of occurrence, different from those already recorded.

Introdução

Os rios efêmeros correspondem a cursos hídricos que possuem fluxo superficial somente em decorrência de precipitações. Por vezes, essa vazão atípica pode ocasionar o transbordamento de outros corpos hídricos ou ainda atingir ao longo do próprio leito efêmero áreas de uso antrópico, como de agricultura ou urbanização.

Um caso especial são os fluxos de lama ou corridas de lama, também chamados de huaicos, termo que caracteriza a massa de lama e pedras desprendidas das montanhas em decorrências de chuvas torrenciais que ao alcançar os cursos d'água causam inundações (REAL ACADEMIA ESPAÑOLA, 2022a) que ocorrem de forma recorrente no Peru, em decorrência do aumento da precipitação ocasionada pelo fenômeno El Niño. Nessa situação, os fluxos formados em áreas nomeadas de quebradas (bacias hidrográficas), que caracteriza as fendas na montanha, ou talvegues, ou seja, a área de formação dos cursos hídricos na montanha, geram inundações (REAL ACADEMIA ESPAÑOLA, 2022b) que carregam materiais de variadas granulometrias em direção ao mar atravessando em alguns casos áreas urbanizadas, causando prejuízos tanto de natureza ambiental, quanto econômicos e sociais.

Somente entre 2003 e o primeiro trimestre de 2021 foram registradas 2.421 ocorrências de huaicos no país (INDECI, 2021) atingindo aproximadamente 450.000 pessoas e danificando aproximadamente 37.000 habitações, além de 141 óbitos. Embora sejam eventos recorrentes a localização destes é bastante variável ao longo do tempo. Apesar da recorrência desses desastres, a urbanização na região costeira do Peru vem se desenvolvendo de forma desordenada, sem atentar para os riscos de novos impactos decorrentes dessas corridas de lama.

No entanto, já existem, a nível mundial, federal e provincial políticas e planos para gestão e prevenção de riscos a desastres, sendo necessária a compatibilização ao fator local. No âmbito nacional no Peru, existem outros instrumentos reguladores sobre a gestão de risco: A Lei nº 29.664 (PCM, 2011a) instituiu o Sistema Nacional de Riscos de Desastres e o validado pelo Decreto Supremo nº 048/2011 (PCM, 2011b) que aprova os regimentos da Lei nº 29.664. Nesses documentos são apontadas as entidades responsáveis pela gestão de risco e elaboração do plano e desempenho de atividades, dentre outras disposições.

Destaca-se ainda, o Plano Nacional de Gestão de Riscos, o qual contém uma estimativa do risco, a prevenção, a redução, a preparação, a resposta, a reabilitação e reconstrução do risco. Além do plano nacional geral existem planos nos demais níveis governamentais, dentre os quais se destacam Planos de prevenção e redução de risco de desastres, de preparação, de operação de emergência, de educação comunitária, de reabilitação e de contingência. A última versão do Plano Nacional data do período de 2014 a 2021, conforme o Decreto Supremo nº 034 de 2014 (PCM, 2014).

A Estrutura Sendai para Redução do Risco de Desastres 2015–2030 foi adotada na Terceira Conferência Mundial das Nações Unidas sobre Redução do Risco de Desastres, realizada em 2015 em Sendai, Japão (UN, 2015). A Conferência representou uma oportunidade para os países participantes adotarem uma estrutura pós 2015 concisa, focada, voltada para o futuro com vistas à ação para redução de risco de desastres; avaliar a experiência adquirida através das estratégias nacionais e regionais para redução do risco de desastres e identificar modalidades de cooperação para implementar posterior a 2015 para redução de risco de

desastres (UN, 2015). As ações guia para alcançar minimizar os riscos de desastres segundo UN (2015) passam por: (I) Compreender o risco de desastres; (II) Fortalecer a governança do risco de desastres para gerenciar o risco de desastres; (III) Investir na redução de desastres para resiliência; (IV) Melhorar a preparação para desastres para uma resposta eficaz e “reconstruir melhor” a recuperação, reabilitação e reconstrução.

Nesse contexto, torna-se importante primeiramente compreender o risco, mapear as zonas e riscos relacionados às bacias das quebradas, tanto para ações específicas para essas áreas, quanto para direcionar a urbanização futura da região, com vistas a minimizar e mitigar os impactos dos *huaicos* ocasionadas pelos corpos hídricos efêmeros no Peru, em decorrência do aumento de precipitação durante fenômenos de *El Niño*.

Este trabalho tem como objetivo identificar os fatores ambientais que mais influenciam a ocorrência de *huaicos* no Peru.

Metodologia

Para identificar os fatores que mais influenciam na ocorrência de *huaicos* no Peru foi realizada uma pesquisa bibliográfica a respeito dos *huaicos* no país. Além disso, foram identificadas as áreas de ocorrência de *huaicos* registradas no país através dos dados do Inventário de Perigos Geológicos no Peru, com registros desde 1998 até 2018 (INGEMMET, 2023). Uma base de fatores ambientais relacionados a ocorrência de *huaicos* foi elaborada, a partir dos resultados da pesquisa bibliográfica, para os quais foram identificadas as características relacionadas a ocorrência de *huaicos*. Isso se deu pelo cruzamento das áreas de ocorrência com os fatores ambientais, por meio de software de sistema de informações geográficas QGIS.

Resultados e discussões

Na Figura 1 pode ser observada a área de estudo, o Peru, localizado na América do Sul, banhado pelo oceano Pacífico, sob influência direta do fenômeno *El Niño*. Na mesma figura podem ser observadas as áreas de ocorrência de *huaicos* registradas pelo Inventário de Perigos Geológicos no país (INGEMMET, 2023). São 7.736 registros de ocorrências de *huaicos*.

As caracterizações de áreas com ocorrência de *huaicos* na província de Lima, realizadas pelo INGEMMET, apontam como fatores preponderantes para a ocorrências desses desastres a precipitação em áreas de formação de cursos hídricos efêmeros, a própria formação geológica e geomorfológica, além do relevo da região, tanto com relação a conformação altitudinal do terreno, como a declividade desse (INGEMMET, 2005; 2012; 2019a; 2020a; 2020b; 2021). Além desses, há a própria influência do fenômeno *El Niño* aumentando o nível de precipitação no país (SENAMHI, 2021).

A informação de relevo obtida a partir do SRTM (*Shuttle Radar Topography Mission*) com 30 metros de pixel revela que a altitude no Peru varia de 0 metros no nível do mar até 6.125 metros na porção central do país. Identifica-se uma cadeia montanhosa dividindo o país conforme avança-se da costa em direção ao continente. Observa-se que as ocorrências de *hwaicos* estão limitadas

desde a costa até esta elevação. Na porção do país após a região elevada, nos departamentos de Loreto, Ucayali e Madre de Dios praticamente não existem ocorrências de *hwaicos*.

Cruzando as informações de ocorrências e altitude, 36% dos registros de ocorrências de *hwaicos* ocorreram até 1.000 metros de altitude. Considerando os registros até uma altitude de 2.000 metros, 56% dos *hwaicos* estão compreendidos nessa altitude. Na Figura 2 podem ser observadas as demais faixas de altitude relacionadas as ocorrências de *hwaicos*.

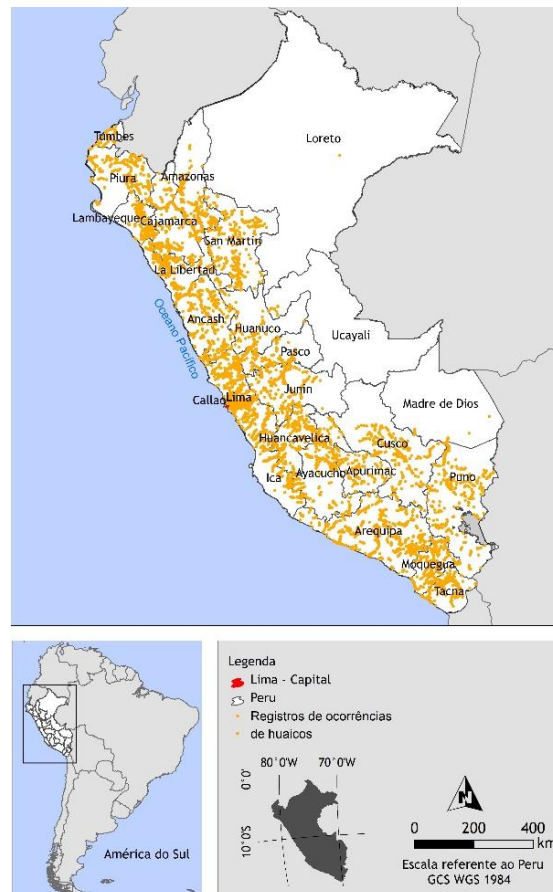


Figura 1: Área de estudo (Peru) e ocorrências de *hwaicos*

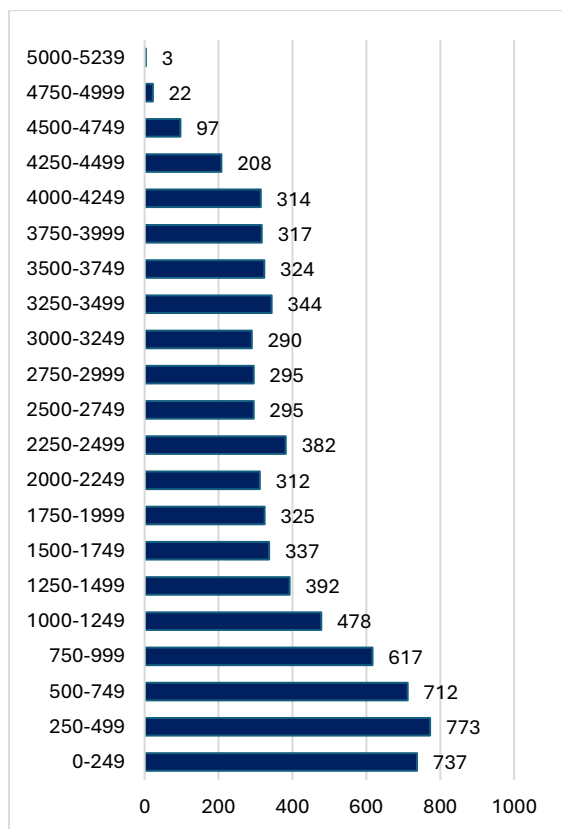


Figura 2: Ocorrências de huaicos com relação à altitude

A declividade, outro fator influente na ocorrência de huaicos foi gerada no QGIS partir do Modelo Digital de Elevação SRTM. Esta varia de 0° à 49,58° de inclinação, correspondente à 110% de inclinação, considerando 45° como 100% de inclinação.

As ocorrências de huaicos acontecem especialmente em áreas de inclinação de até 5°, com 71% das ocorrências nesse intervalo (Figura 3). Isso indica que as ocorrências foram registradas nas porções de acúmulo do fluxo. Entende-se que a geração do fluxo ocorre em áreas de maior inclinação, no entanto, seria necessária a avaliação de cada microbacia isoladamente para compreender o gradiente de declividade relacionado a cada ocorrência de huaicos.

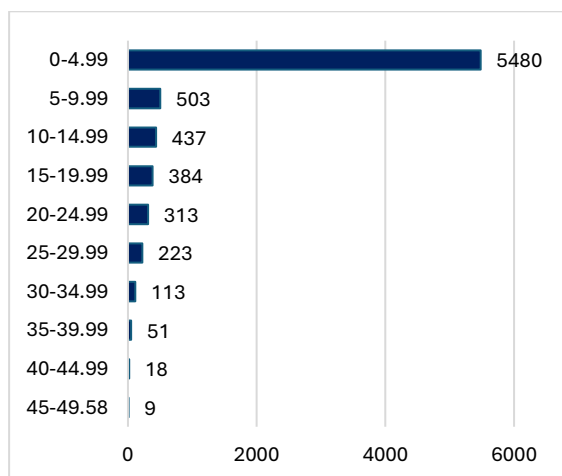


Figura 3: Ocorrências de huaicos com relação a declividade em graus (°)

Dos dados geológicos disponibilizados pelo INGEMMET (2019b), observam-se mapeadas 112 classes geológicas no país. Dessas 95 estão relacionadas as ocorrências de huaicos. 13% das ocorrências estão relacionadas a classificação, Quaternário Holoceno Continental (1.015 ocorrências). Além disso, 52% das ocorrências estão associadas à outras 10 classificações, conforme a Figura 4.

A classificação Quaternário Holoceno Continental diz respeito ao período geológico ainda em vigência. A época do Holoceno teve início há 11,7 mil anos, compondo então a formação geológica mais recente e estratigraficamente localizada acima das demais formações (INGEMMET, 2018).

São identificadas 129 classes geomorfológicas, segundo (INGEMMET, 2016). Das ocorrências de huaicos cruzadas com a geomorfologia destacam a relação dessas com 82 classificações. Dessas, destaca-se a classe geomorfológica de Vertente Aluvio-Torrencial, com 1.820 ocorrências de huaicos, 24% do total e 16% das ocorrências estão relacionadas a classificação geomorfológica de Montanha em rocha sedimentar. Somadas, essas classes representam 40% das ocorrências. As outras classes geomorfológicas com recorrência de ocorrências de huaicos estão apontadas na Figura 5.

A classe de Vertente Aluvio-Torrencial é uma planície inclinada que se estende ao pé dos contrafortes ou sistemas montanhosos andinos. São o resultado do acúmulo de materiais mobilizados pelos huaicos (INGEMMET, 2021; 2023).

A classe de Montanha em rocha sedimentar é litologicamente composta por sequências sedimentares de argilitos siltosos, arenitos, argilitos carbonáceos e sequências calcárias que se caracterizam por quedas de rochas, deslizamentos, deslizamentos e avalanches de rochas (INGEMMET, 2021).

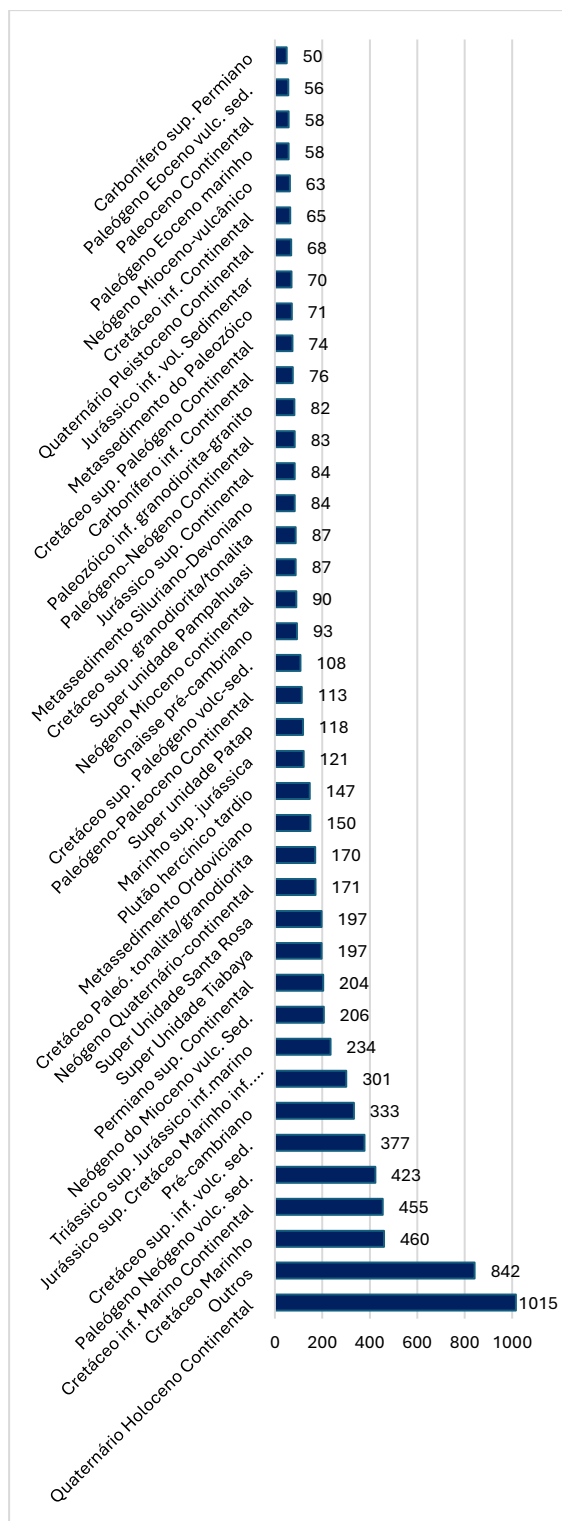


Figura 4: Ocorrências de huaicos com relação a geologia

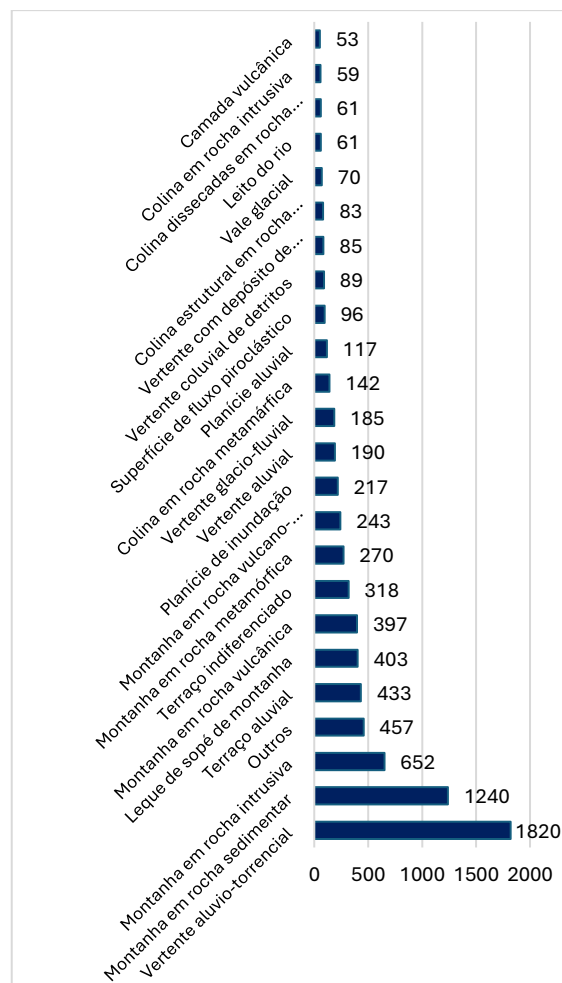


Figura 5: Ocorrências de huaicos com relação a geomorfologia

Observa-se ainda a relação entre a classe geomorfológica de maior recorrência Vertente Aluvio-Torrencial com as classificações geológicas na Figura 6. O resultado do cruzamento aponta que com exceção de classes com baixas recorrências, a classe geológica Quaternário Holoceno Continental representa a maior coincidência com a geomorfologia, sendo esta também a classificação de maior recorrência geológica no geral.

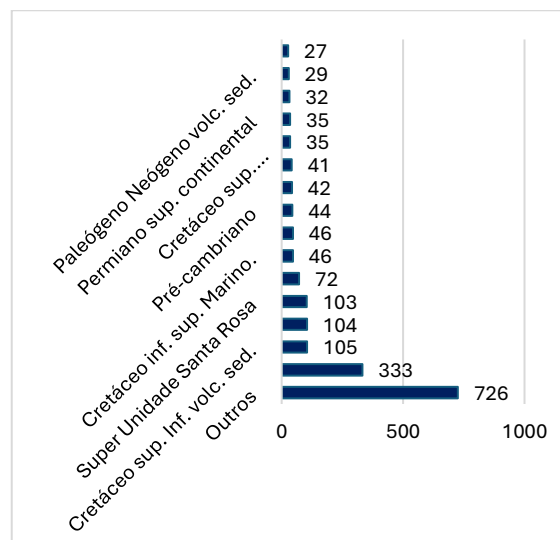


Figura 6: Correspondência entre a classe geomorfológica Vertente Aluvio-Torrencial e a geologia

A avaliação realizada com base na classe geológica de maior recorrência na ocorrência de huaicos Quaternário Holoceno Continental, novamente se destaca a relação com a classe geomorfológica de Vertente Aluvio-Torrencial, como a maior correspondência, conforme pode ser observado na Figura 7.

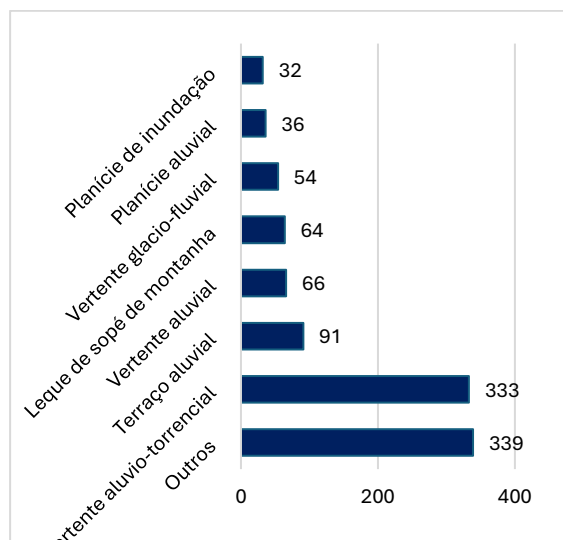


Figura 7: Correspondência entre a classe geológica Quaternário Holoceno Continental e a geomorfologia

No entanto, há somente 33% de correspondência, tendo como referência à classe geológica Quaternário Holoceno Continental e a classe geomorfológica de Vertente Aluvio-Torrencial. Quando tomada por referência a classificação geomorfológica, essa correspondência é menor ainda, de 18%.

Conclusões

A avaliação dos fatores desencadeantes da ocorrência de huaicos é extremamente complexa. Embora existam indicações de fatores físicos como relevo e formação geológica e geomorfológica que influenciem a ocorrência de huaicos, a relação entre o inventário de ocorrências de huaicos e a altitude, geomorfologia e geologia, embora tenham retornado alguns padrões, revelaram grandes desafios na validação de um padrão. Do total de classes geológicas encontradas no país, 85% são relacionadas com a ocorrência de huaicos. Das classes geomorfológicas existentes no país, 66% são relacionadas aos huaicos.

Ou seja, a ocorrência de huaicos é um fenômeno que exige a consideração de mais fatores desencadeantes, especialmente e fundamentalmente a precipitação, sendo este o fenômeno externo à formação geomorfológica diretamente relacionado aos huaicos considerando-se o fator de arraste de materiais pelos fluxos decorrentes do escoamento superficial. A análise dos fatores determinantes dos huaicos exige ferramentas e metodologias mais robustas para encontrar padrões de ocorrência que possam servir para identificar outras áreas de ocorrência, distintas das já registradas.

Referências bibliográfica

- INDECI – Instituto Nacional de Defensa Civil. (2021). Compendio Estadístico del Instituto Nacional de Defensa Civil. <https://www.gob.pe/institucion/indeci/informes-publicaciones/2702972-compendio-estadistico-2021>. (acessado em 05 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2019b). Carta Geológica Nacional a escala 1:50000. <https://portal.ingemmet.gob.pe/web/guest/carta-geologica-nacional-a-escala-1-50-000>. (acessado em 10 fev. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2091). Evaluación de peligros geológicos por movimientos en masa en el Anexo de Colpa. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/2330>. (acessado em 24 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2021). Evaluación de peligros geológicos en el sector de San Francisco de Panamá. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/3059>
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2021). Evaluación de peligros geológicos por flujo de detritos (huaicos) en la Quebrada Zaparo. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/2695>. (acessado em 24 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2021). Evaluación de peligros geológicos por flujo de detritos (huaicos) en los Anexos Ticapampa y Oconchay. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/3155>. (acessado em 24 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2023) Evaluación de peligro geológico por movimientos en masa en el centro poblado Ambasal. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/4566>. (acessado em 05 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2021). Geología del cuadrángulo de Calca (hojas 27s1, 27s2, 27s4). <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/3188>. (acessado em 17 jul. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2020). Inspección geológica realizada en la Quebrada San Lázaro – Arequipa. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/2793>. (acessado em 24 jun. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2016). Mapa geomorfológico del Perú. <http://geocatmin.ingemmet.gob.pe/geocatmin/>. (acessado em 08 mar. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2019). Mapa de Inventario de Peligros Geológicos. <https://portal.ingemmet.gob.pe/web/guest/mapa-de-inventario-de-peligros-geologicos>. (acessado em 08 mar. 2024).
- INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2012). Peligros por movimientos en masa en la quebrada Tinaja (Dist. Pachacamac-Cieneguilla, provincia y región

Lima). https://repositorio.ingemmet.gob.pe/bitstream/20.500.12544/1572/1/A6600-Peligros_movimientos_en_masa_Qda_Tinajas-Lima.pdf. (acessado em 24 jun. 2024).

INGEMMET - Instituto Geológico, Minero y Metalúrgico. (2005). Zonas críticas por peligros geológicos en la cuenca Hancay-Lambayeque. <https://repositorio.ingemmet.gob.pe/handle/20.500.12544/2008>. (acessado em 24 jun. 2024).

PCM - Presidencia del Consejo de Ministros. (2011b). Decreto Supremo n° 048 de 26 de mayo del 2011. Aprueba el Reglamento de la Ley n° 29664, que crea el Sistema Nacional de Gestión del Riesgo de Desastres (SINAGERD). Lima, 2011. <https://www.geoidep.gob.pe/images/descargas/DS-048-2011-PCM.pdf>. (acessado em 18 jun. 2024).

PCM - Presidencia del Consejo de Ministros. (2011a). Ley n° 29.664 de 19 de febrero del 2011. Ley que crea el Sistema Nacional de Gestión del Riesgo de Desastres (SINAGERD). Lima, 2011. <https://busquedas.elperuano.pe/normaslegales/ley-que-crea-el-sistema-nacional-de-gestion-del-riesgo-de-de-ley-n-29664-605077-1/>. (acessado em 1 jun. 2024).

PCM - Presidencia del Consejo de Ministros. (2014). Decreto Supremo n° 034 de 13 de mayo del 2014. Aprueba el Plan Nacional de Gestión del Riesgo de Desastres-PLANAGERD 2014-2021. Lima, 2014. https://cdn.www.gob.pe/uploads/document/file/1229859/DS-N_034-2014-PCM-PLANAGERD20200812-2906259-1x7m674.pdf. (acessado em 15 jun. 2024).

Real Academia Española. Huaico. 2022a. Disponível em: <https://dle.rae.es/huaico>. (acessado em 1 jun. 2024).

Real Academia Española. Quebrada. 2022b. Disponível em: <https://dle.rae.es/quebrada>. (acessado em 1 jun. 2024).

USGS – United States Geological Survey. SRTM - Shuttle Radar Topography Mission. <https://earthexplorer.usgs.gov/>. (acessado em 05 fev. 2024).

UN - United Nations. (2011). World population prospects: the 2011 revision. https://www.un.org/en/development/desa/population/publications/pdf/urbanization/WUP2011_Report.pdf. (acessado em 20 jun. 2024).