



UFRRJ
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM
MODELAGEM MATEMÁTICA E
COMPUTACIONAL

DISSERTAÇÃO

Predição de deslizamento de terra para auxiliar sistemas de alerta. Um estudo de caso contextualizando limiares de precipitação e umidade do solo para região urbana do município de São Bernardo do Campo

Antonio Carlos da Silva

2025



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
MATEMÁTICA E COMPUTACIONAL**

**PREDIÇÃO DE DESLIZAMENTO DE TERRA PARA
AUXILIAR SISTEMAS DE ALERTA. UM ESTUDO DE CASO
CONTEXTUALIZANDO LIMIARES DE PRECIPITAÇÃO E
UMIDADE DO SOLO PARA REGIÃO URBANA DO
MUNICÍPIO DE SÃO BERNARDO DO CAMPO**

ANTONIO CARLOS DA SILVA

Sob a orientação da professor

Marcos Bacis Ceddia

e Co-orientação do professor

Felipe Leite Coelho da Silva

Dissertação submetida como requisito parcial para obtenção o grau de Mestre em Ciências, no Curso de Pós-Graduação em Modelagem Matemática e Computacional, Área de Concentração em Modelagem Matemática e Computacional.

Seropédica, RJ
Abril de 2025

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

Silva, Antonio Carlos da, 1996-
S587p Predição de deslizamento de terra para auxiliar
sistemas de alerta. Um estudo de caso
contextualizando limiares de precipitação e umidade do
solo para região urbana do município de São Bernardo
do Campo / Antonio Carlos da Silva. - Pinheiral, 2025.
103 f.

Orientador: Marcos Bacis Ceddia.
Coorientador: Felipe Leite Coelho da Silva.
Dissertação(Mestrado). -- Universidade Federal
Rural do Rio de Janeiro, Programa Pós-Graduação em
Modelagem Matemática e Computacional, 2025.

1. Movimento de Massa. 2. Regressão Logística. 3.
Limiares Pluviométricos . 4. Umidade do Solo. 5.
Susceptibilidade. I. Ceddia, Marcos Bacis, 1968-,
orient. II. Silva, Felipe Leite Coelho da, 1981-,
coorient. III Universidade Federal Rural do Rio de
Janeiro. Programa Pós-Graduação em Modelagem Matemática
e Computacional. IV. Título.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL



Seropédica-RJ, 28 de abril de 2025.

ANTONIO CARLOS DA SILVA

Dissertação submetida como requisito parcial para a obtenção de grau de **Mestre**, no Programa de Pós-Graduação em Modelagem Matemática e Computacional PPGMMC, área de Concentração em Modelagem Matemática e Computacional.

DISSERTAÇÃO APROVADA EM 28/04/2025

Marcos Bacis Ceddia Drº (Orientador, Presidente da Banca-UFRRJ)

Rosane Ferreira de Oliveira Drª (membro interno-UFRRJ)

Diego Nunes Brandão Drº (membro externo à Instituição- CEFET-RJ)



ATA Nº ata/2025 - ICE (12.28.01.23)
(Nº do Documento: 1685)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 12/05/2025 12:31)
MARCOS BACIS CEDDIA
PROFESSOR DO MAGISTERIO SUPERIOR
DATS (11.39.00.35)
Matrícula: ####202#6

(Assinado digitalmente em 12/05/2025 15:06)
ROSANE FERREIRA DE OLIVEIRA
PROFESSOR DO MAGISTERIO SUPERIOR
DeptM (12.28.01.00.00.00.63)
Matrícula: ####76#9

(Assinado digitalmente em 09/05/2025 19:37)
DIEGO NUNES BRANDÃO
ASSINANTE EXTERNO
CPF: ####.###.947-##

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: **1685**, ano: **2025**,
tipo: **ATA**, data de emissão: **09/05/2025** e o código de verificação: **b6cce57e98**

Agradecimentos

Primeiramente, aos meus pais que, com muito carinho e apoio, não mediram esforços para que eu chegassem até esta etapa de minha vida, à eles todo o meu amor e gratidão.

Aos meus orientadores, Marcos Bacis e Felipe, que compartilharam comigo o conhecimento, bem mais precioso que se pode conquistar. Em especial ao professor Marcos Bacis, pela paciência na orientação e incentivo, sem os quais não seria possível a conclusão desta dissertação.

Agradeço também a todos os professores do PPGMMC e a equipe de laboratório LASA (Laboratório de Água e Solo em Agroecossistemas) que me acompanharam durante o mestrado e foram tão importantes na minha vida acadêmica e no desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Resumo da Dissertação apresentada ao PPGMMC/UFRRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

PREDIÇÃO DE DESLIZAMENTO DE TERRA PARA AUXILIAR SISTEMAS DE ALERTA. UM ESTUDO DE CASO CONTEXTUALIZANDO LIMIARES DE PRECIPITAÇÃO E UMIDADE DO SOLO PARA REGIÃO URBANA DO MUNICÍPIO DE SÃO BERNARDO DO CAMPO

Antonio Carlos da Silva

Abril/2025

Orientadores: Marcos Bacis Ceddia

Felipe Leite Coelho da Silva

Programa: Modelagem Matemática e Computacional

O aumento significativo de desastres naturais nas últimas décadas representa um grande desafio para as autoridades governamentais, especialmente os movimentos de massa, que possuem alta letalidade e capacidade de destruição. Sistemas de alerta antecipado são essenciais para a mitigação de riscos, mas muitos adotam limiares de precipitação empíricos sem consenso sobre critérios técnicos. Este estudo propõe um algoritmo para reconstrução objetiva e reproduzível de eventos de precipitação e infiltração, cujos resultados foram utilizados como entrada em modelos de limiar e regressão logística. O objetivo é avaliar a eficácia da inclusão de informações de umidade do solo e aspectos espaciais para aprimorar sistemas de alerta, reduzindo falsos positivos e aumentando a acurácia. A análise abrangeu eventos de movimento de massa em São Bernardo do Campo – SP, considerando dados topográficos, pedológicos, pluviométricos e de umidade do solo até 150 cm de profundidade. A introdução da umidade do solo reduziu os falsos positivos nos níveis "Alerta" e "Alerta Máximo" em 43,9% e 66,3%, respectivamente. No modelo de regressão logística, apenas a precipitação e a declividade foram variáveis significativas, resultando em um AUC de 0,94 e precisão de 86% para a classe de deslizamento, permitindo a predição da evolução espaço-temporal da suscetibilidade.

Abstract of Dissertation presented to PPGMMC/UFRRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

LANDSLIDE PREDICTION TO SUPPORT EARLY WARNING SYSTEMS: A CASE STUDY ON RAINFALL AND SOIL MOISTURE THRESHOLDS IN THE URBAN AREA OF SÃO BERNARDO DO CAMPO

Antonio Carlos da Silva

April/2025

Advisors: Marcos Bacis Ceddia

Felipe Leite Coelho da Silva

Department: Mathematical and Computational Modeling

The significant increase in natural disasters in recent decades poses a major challenge for government authorities, particularly mass movements due to their high lethality and destructive potential. Early warning systems are essential for risk mitigation; however, many rely on empirical rainfall thresholds without consensus on technical criteria. This study proposes an algorithm for the objective and reproducible reconstruction of rainfall and infiltration events, whose outputs were used as inputs in threshold-based and logistic regression models. The aim is to assess the effectiveness of incorporating soil moisture and spatial features to enhance warning systems by reducing false positives and improving accuracy. The analysis focused on mass movement events in São Bernardo do Campo, São Paulo, using topographic, pedological, pluviometric, and soil moisture data up to a depth of 150 cm. The inclusion of soil moisture data reduced false positives in the "Alert" and "Maximum Alert" levels by 43.9% and 66.3%, respectively. In the logistic regression model, only precipitation and slope were statistically significant variables, yielding an AUC of 0.94 and an accuracy of 86% for the landslide class, enabling the prediction of the spatiotemporal evolution of susceptibility.

Lista de Figuras

| | | |
|------|---|----|
| 3.1 | Principais tipos de movimentos de massa. | 6 |
| 3.2 | Destaque para situação da encosta com solo exposto, cicatriz de deslizamento, árvores inclinadas. | 10 |
| 3.3 | Deslizamento SBC. | 12 |
| 3.4 | Área de risco, SBC. | 12 |
| 3.5 | a) curva ajustada com 753 eventos resultantes em deslizamento. b) Curva de limiar com 1% e 5% de excedência. | 15 |
| 3.6 | Análise bayesiana bidimensional. (a) Gráfico de intensidade - duração. (b) Histograma da probabilidade de deslizamento de terra. | 21 |
| 3.7 | Classes indicando a facilidade de identificação dos eventos de precipitação. Exemplos de identificação manual (áreas cinzas) e automática (áreas traçadas). | 26 |
| 4.1 | Organograma das etapas do trabalho | 28 |
| 4.2 | Área de estudo. | 29 |
| 4.3 | Mapa de classes de solo de SBC. | 33 |
| 4.4 | Mapa de Uso de Solo de SBC. | 33 |
| 4.5 | Mapa de perigo de deslizamento de SBC. | 34 |
| 4.6 | Mapa de suscetibilidade de deslizamento de SBC. | 34 |
| 4.7 | Mapa de Altimetria de SBC. | 35 |
| 4.8 | Mapa de Declividade de SBC. | 35 |
| 4.9 | Mapa de orientação da vertente de SBC. | 36 |
| 4.10 | Mapa de curvatura vertical de SBC. | 36 |
| 4.11 | Mapa de curvatura horizontal de SBC. | 37 |
| 4.12 | Ocorrências relacionadas a deslizamento de terra. | 38 |
| 4.13 | Sensor de Umidade do Solo EnviroSCAN | 39 |
| 4.14 | Distribuição das estações geotécnicas e pluviométricas. | 39 |
| 5.1 | Exemplo do algoritmo detectando os eventos. | 42 |

| | | |
|------|--|----|
| 5.2 | Associação dos eventos dentro de um raio de busca pré-estabelecido. | 43 |
| 5.3 | Associação dos eventos com os deslizamentos. | 44 |
| 5.4 | Fluxograma com os critérios para selecionar estações aptas ao estudo. | 51 |
| 5.5 | Boxplot da umidade do solo nas seis profundidade da estação geotécnica Vila Baiana, situada no município de Guarujá - SP. | 53 |
| 5.6 | Boxplot da Precipitação da estação geotécnica Vila Baiana, situada no município de Guarujá - SP. | 53 |
| 5.7 | Correlação Pearson da estação Palmira Grassioto. | 54 |
| 5.8 | Correlação Pearson da estação Riviera. | 55 |
| 5.9 | Gráfico da umidade do solo e precipitação da estação Riviera. | 55 |
| 5.10 | Gráfico da precipitação e da umidade do solo nas profundidades 0 - 50 cm, 50 - 100 cm e 100 - 150 cm da estação Palmira Grassioto. | 56 |
| 5.11 | Gráfico da precipitação e da umidade do solo nas profundidades 0 - 50 cm, 50 - 100 cm e 100 - 150 cm da estação ETA Rio Grande. | 57 |
| 6.1 | Comparação do evento de infiltração calculado pelo algoritmo e manual. . | 62 |
| 6.2 | Distribuição de deslizamento em função das ocorrências em São Bernardo do Campo, SP | 63 |
| 6.3 | Eventos de precipitação que resultaram em deslizamentos de terra, no período de julho de 2021 a dezembro de 2022, com a curva das condições empíricas de Acumulado x Duração ajustada por mínimos quadrados. | 67 |
| 6.4 | Estimativa da função de densidade de probabilidade da distribuição gaussiana. | 68 |
| 6.5 | Curvas de limiar com os níveis de excedência. | 68 |
| 6.6 | Algoritmo de decisão usando precipitação e saturação do solo. | 69 |
| 6.7 | Algoritmo de decisão. | 70 |
| 6.8 | Limiar referente ao Alerta Máximo. | 72 |
| 6.9 | Matriz confusão para a primeira técnica de amostragem. | 74 |
| 6.10 | Matriz confusão para a segunda técnica de amostragem. | 76 |
| 6.11 | Curva ROC. | 77 |
| 6.12 | Importância das Variáveis do modelo. | 77 |
| 6.13 | Suscetibilidade prevista a deslizamentos de terra para os dias 18, 19, 20 e 23 de fevereiro de 2023. | 79 |
| 6.14 | Mapas de Suscetibilidade e perigo. | 80 |

Lista de Tabelas

| | | |
|-----|---|----|
| 3.1 | Principais tipos de movimentos de massa em encostas. | 7 |
| 3.2 | Fatores que Influenciam a Ocorrência de Movimento de Massa, adaptado de CRUDEN e VARNES (1996). | 8 |
| 3.3 | Principais ações e critérios dos níveis dos Planos Preventivos. | 11 |
| 4.1 | Codificação das variáveis estáticas condicionantes a escorregamento de terra. | 32 |
| 5.1 | Estações Geotécnica Descartadas. | 52 |
| 5.2 | Estações Geotécnica Descartadas. | 56 |
| 5.3 | Estações Aptas. | 56 |
| 6.1 | Parâmetros do algoritmo. | 60 |
| 6.2 | Diferentes raio de entrada para o algoritmo. | 61 |
| 6.3 | Frequência relativa das ocorrências de deslizamento de terra dentro das modalidades das variáveis condicionantes. | 66 |
| 6.4 | Resultados da previsão dos algoritmos de decisão | 71 |
| 6.5 | Resultados da previsão dos algoritmos de decisão. Fonte: Próprio Autor. | 73 |
| 6.6 | Evolução do valor p no processo de seleção Stepwise | 74 |
| 6.7 | Relatório de Classificação do Modelo | 74 |
| 6.8 | Evolução do valor p no processo de seleção Stepwise | 75 |
| 6.9 | Relatório de Classificação | 76 |

Sumário

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 2 | Justificativa | 3 |
| 2.1 | Organização da dissertação | 4 |
| 3 | Referencial Teórico | 5 |
| 3.1 | Movimento de Massa | 5 |
| 3.1.1 | Rastejo | 6 |
| 3.1.2 | Queda de Blocos | 6 |
| 3.1.3 | Escorregamento | 7 |
| 3.1.4 | Corrida de Massa | 8 |
| 3.2 | Fatores que Influenciam a Ocorrência de Movimento de Massa | 8 |
| 3.3 | Gestão de Risco de Deslizamentos em São Bernardo do Campo | 9 |
| 3.4 | Aviso Prévio de Deslizamento | 13 |
| 3.4.1 | Limiares Empíricos | 13 |
| 3.4.2 | Abordagens Estatísticas | 19 |
| 3.4.3 | Abordagens Físicas | 23 |
| 3.5 | Algoritmo de Reconstrução de Evento de Precipitação e Infiltração . . | 24 |
| 4 | Materiais | 28 |
| 4.1 | A Área de estudo | 29 |
| 4.2 | Banco de Dados | 29 |
| 4.2.1 | Variáveis Estáticas | 30 |
| 4.2.2 | Variáveis Dinâmicas | 37 |
| 5 | Método Adotado | 40 |
| 5.1 | O Algoritmo | 40 |
| 5.1.1 | Etapa 1: Normalização dos Dados de Umidade do Solo | 41 |
| 5.1.2 | Etapa 2: Detecção de Eventos | 41 |

| | | |
|----------|--|-----------|
| 5.1.3 | Etapa 3: Associação dos Eventos com Deslizamentos de Terra | 42 |
| 5.1.4 | Etapa 4: Cálculo das Métricas | 44 |
| 5.2 | Modelo 1: Limiar | 44 |
| 5.3 | Modelo 2: Regressão Logística | 45 |
| 5.3.1 | Método Stepwise | 47 |
| 5.3.2 | Amostragem | 48 |
| 5.3.3 | Dimensionamento dos Dados | 49 |
| 5.4 | Métricas de Ajuste | 50 |
| 5.5 | Pré Processamento e Organização dos Dados | 51 |
| 6 | Resultados e Discussão | 58 |
| 6.1 | Análise de Incerteza Associada aos Dados | 58 |
| 6.2 | Algoritmo | 59 |
| 6.3 | Análise Exploratória de dados | 63 |
| 6.4 | Limiares | 67 |
| 6.5 | Regressão Logística | 73 |
| 7 | Conclusões e Próximas Etapas | 82 |
| | Referências Bibliográficas | 84 |

Capítulo 1

Introdução

Nas últimas décadas, houve um aumento significativo nas ocorrências de desastres naturais e dos prejuízos decorrentes no contexto mundial. Esse crescimento é particularmente notável a partir da década de 70 (TOMINAGA *et al.*, 2015). O Brasil, segue essa tendência mundial, com um aumento significativo nas ocorrências de desastres naturais a partir de 1960, sendo os principais fenômenos, inundações e enchentes, movimento de massa e tempestades (TOMINAGA *et al.*, 2015). De acordo com DOS SANTOS ALVALÁ *et al.* (2019), no ano de 2010 aproximadamente 8.266.566 pessoas estiveram expostas a ameaças de movimento de massa, enchentes e inundações repentinas. Isso indica que, para cada 100 habitantes, 9 residiam em zonas de risco no Brasil. De maneira geral, as comunidades que residem em zonas de risco enfrentam não apenas a exposição aos perigos de paisagens vulneráveis, mas também desafios sociais significativos. Quando ocorrem desastres naturais, como deslizamentos de terra, essas populações sofrem com danos parciais ou totais em suas moradias, além da desestruturação de suas redes sociais.

Os desastres naturais têm sido reconhecidos como um sério desafio pelas autoridades e pelo poder público. Entre eles, as inundações são os mais frequentes, enquanto os movimentos de massa possuem maior letalidade e capacidade de destruição. A implementação e manutenção de sistemas de alerta antecipado para eventos extremos são fundamentais para a prevenção de riscos associados a movimentos de massa. Esses sistemas têm como objetivo informar a população sobre o progresso de possíveis ameaças, prevenir vítimas mortais e fornecer opções seguras para minimizar danos ou prejuízos. Para isso, são indispensáveis pesquisas e estudos para aprofundar o entendimento dos elementos associados aos deslizamentos de terra, visando aprimorar a implementação de sistemas de alerta mais precisos, facilitando a gestão eficaz dos riscos.

Uma das abordagens mais difundidas é estabelecer limiares de precipitação, de forma empírica, com base na intensidade e duração de chuva, durante um determinado período, antes da ocorrência do deslizamento de terra. Existem numerosos

estudos analisando diversas variáveis pluviométricas além da intensidade e duração, no entanto, não há um consenso estabelecido na literatura sobre os parâmetros ou critérios técnicos a serem adotados em um sistema de alerta (CAINE, 1980; CHLE-BORAD *et al.*, 2008; GUZZETTI *et al.*, 2005, 2007, 2008; LUMB, 1975; MARTEL-LONI *et al.*, 2012; MENDES *et al.*, 2020; MIRUS *et al.*, 2018; REICHENBACH *et al.*, 1998; ROSSI *et al.*, 2012; SANTORO *et al.*, 2010). De maneira geral, limiares críticos baseados na precipitação tendem a resultar em uma alta incidência de alertas falsos, pois, quando utilizados de forma isolada, podem não representar as situações críticas, por não considerar a umidade nos perfis do solo nas encostas, nem os processos de infiltração, retenção e drenagem da água da chuva. É fato que os movimentos de massa, do tipo escorregamento são causados por precipitação pluviométrica (chuva), ocorrem quando o solo está saturado ou próximo da saturação. Nesse nível de umidade, a tensão de cisalhamento é rompida e se desencadeia o evento. Nesse aspecto, cabe ressaltar o importante serviço ambiental do solo no armazenamento e dinâmica da água. O solo funciona com um grande reservatório de água. Para exemplo, considere um volume de $1m^3$ de solo (um monólito com $1m^2$ de área e 1 metro de profundidade). Se esse mesmo $1m^3$ de solo tiver uma porosidade de 50%, metade desse $1m^3$ ($0.5m^3$) poderá conter água (500 litros de água se esse solo estiver saturado). Isso significa na prática que se o solo de uma encosta estiver totalmente seco (0% de umidade), seria necessária uma chuva de 500 milímetros (500 litros por m^2 para saturá-lo e causar o desencadeamento do escorregamento. A umidade do solo varia constantemente no tempo e no espaço e, como consequência, nem sempre uma chuva relativamente intensa (100 mm ou maior, por exemplo) causará o desencadeamento de um movimento de massa. Importa saber qual foi a umidade antecedente do solo para saber se uma determinada chuva colocará uma determinada área em situação de risco.

Sistemas que incorporam dados de precipitação e umidade do solo têm demonstrado ser cada vez mais promissores. Isso se deve ao fato de que um dos principais fatores desencadeadores de deslizamento de terra está relacionado à infiltração da água no solo, o que reduz sua resistência ao cisalhamento ao diminuir a coesão aparente devido à saturação do solo e ao aumento das pressões hidrostáticas.

Neste contexto, o objetivo deste estudo foi avaliar a eficácia de se introduzir informações de umidade do solo e precipitação no aperfeiçoamento de sistemas de alerta a movimentos de massa no município de São Bernardo do Campo - SP (SBC). Para atingir o objetivo, o estudo teve os seguintes objetivos específicos:

- a) desenvolver um algoritmo para a definição objetiva e reproduzível de eventos de infiltração e precipitação;
- b) Implementar modelos de predição e identificar quais variáveis melhor antecipam o perigo iminente de deslizamento de terra.

Capítulo 2

Justificativa

Dados de umidade do solo são muito escassos no Brasil, sendo o CEMADEN (Centro Nacional de Monitoramento e Alertas de Desastres Naturais) o órgão responsável pela maior rede de monitoramento de umidade do solo no Brasil. Com base nesses dados, realizamos uma análise exploratória, na qual o município de SBC foi selecionado como área de estudo por apresentar a maior concentração de sensores com dados consistentes. Além de um banco de dados extenso de deslizamento de terra.

O município de SBC, localizado ao sudeste da região Metropolitana do estado de São Paulo, é frequentemente impactado por movimentos de massa induzidos por fortes precipitações, resultando em danos socioeconômicos e ambientais. Em resposta, o poder público tem desenvolvido políticas públicas para melhorar a gestão de risco. A Defesa Civil desenvolveu o "Plano Municipal de Redução de Risco"(PMRR), visando à identificação de áreas de risco, e conduz ações como a "Operação Guarda-Chuva", voltada ao aprimoramento da resposta emergencial e à conscientização da população. O "Plano Preventivo de Defesa Civil"(PPDC), que utiliza limiar crítico e visitas técnicas para definir os níveis de alerta de toda a região Metropolitana do estado de São Paulo, com o objetivo é prevenir vítimas mortais, removendo preventivamente as populações em área de risco. Além de diversos estudos e levantamentos de mapas pedológicos e geológicos na região.

Inicialmente, os limiares críticos de precipitação para o município foram estabelecidos em 100 mm acumulados em 72 horas. No entanto, um estudo realizado por SANTORO *et al.* (2010) analisou dados de deslizamentos em encostas e precipitação acumulada, no período de 2000 a 2010, onde observou um aumento no número de deslizamentos abaixo do limiar. Como consequência, os índices foram rebaixados em todos os municípios do estado de São Paulo atendidos pelo PPDC, No município de SBC o novo limiar referência adotado é de 80 mm de precipitação acumulada em 72 horas. Os autores também compararam os acumulados de precipitação diárias e de três dias de dois eventos específicos de deslizamentos, ocorridos nos municípios

de São Roque e Votorantim, respectivamente nos dias 28 e 27 de janeiro de 2004. Embora ambos os municípios pertençam à mesma região operacional do PPDC, os padrões de distribuição das chuvas nos dias que antecederam os eventos foram significativamente distintos, evidenciando a necessidade de estudos específicos para cada município.

Modelos de predição é um tema pouco explorado na região, a identificação dos limiares críticos, incorporando dados de precipitação e umidade do solo poderão reduzir os alertas falsos e servir como base para enfrentar novos eventos. A discussão sobre os parâmetros ou critérios para estabelecer limiares críticos é de suma importância para o aprimoramento dos sistemas de alertas. Na literatura não tem estabelecido que a precipitação acumulada com tempo fixo de 72 horas utilizado no PPDC é o melhor parâmetro para deflagrar deslizamentos. Muitos artigos abordam a duração de um evento de precipitação de forma dinâmica.

Segundo MELILLO *et al.* (2015) os investigadores calculam os eventos manualmente e muitas vezes não especificam os critérios utilizados para selecionar o sensor que representa o deslizamento ou a forma em que foi calculada a precipitação acumulada e a duração do evento. Mesmo quando especificados, muitas vezes é difícil determinar se foram aplicados de forma consistente. A construção de um o algoritmo automatizado para reconstruir os eventos definindo: a precipitação acumulada, a duração, a saturação inicial e final, entre outras variáveis. Elimina possíveis vícios do pesquisador, garante consistência metodológica e a reproduzibilidade.

Modelos de avaliam a suscetibilidade à deslizamento de forma dinâmica no espaço e no tempo incorporando variáveis estáticas como topografia do terreno e variáveis dinâmicas como a precipitação também podem auxiliar os sistemas de alerta. Esta dissertação busca contribuir com o processo metodológico de calcular as variáveis deflagradoras de deslizamento de forma automatizada, além, de implementar modelos de predição visando aperfeiçoar os sistemas de alerta.

2.1 Organização da dissertação

No capítulo 3, é apresentada o Referencial Teórico sobre movimentos de massa e as abordagens utilizadas para definir limiares críticos, modelos estatísticos e físicos para sistema de alerta antecipado. No capítulo 4 são apresentados os conjuntos de dados que foram utilizados. No capítulo 5, os procedimentos que foram desenvolvidos, como a escolha da área de estudo, dos modelos de predição, das metodologias e o pre-processamento dos dados. No capítulo 6, são apresentados os resultados da aplicação da metodologia proposta. No capítulo 7, são apresentadas as conclusões deste trabalho e os possíveis trabalhos futuros. E, finalmente, no capítulo 8 as referências bibliográficas que foram utilizadas para a construção deste trabalho.

Capítulo 3

Referencial Teórico

3.1 Movimento de Massa

Movimentos de massa são fenômenos naturais que envolvem o deslocamento de materiais como solo, rocha ou vegetação, geralmente em áreas inclinadas ou montanhosas. A ocorrência desses movimentos é comumente associada a influência de diversos fatores, tais como: clima, geologia, topografia, hidrogeologia e uso e a cobertura do solo. Destacam-se também as ações antrópicas como desmatamento e ocupação inadequada das terras (sobretudo encostas), que acarretam em um desequilíbrio do ambiente fazendo com que aumente a suscetibilidade aos desastres (TOMINAGA *et al.*, 2009). Existem diversas classificações de autores para os movimentos de massa. Segundo CUNHA *et al.* (1991); GUIDICINI e NIEBLE (1976) e SELBY (1993) classificações mais detalhadas consideram a velocidade de deslocamento, a relação entre a cinemática com massa do material desprendido e estável, de acordo com o clima e a geomorfologia do ambiente. Segundo TOMINAGA *et al.* (2009) e AUGUSTO FILHO (1992) o movimento de massa pode ser classificado em quatro subtipos: rastejo, escorregamentos ou deslizamento, quedas de blocos e corridas de massa , conforme apresentado na Tabela 3.1 e ilustrado na Figura 3.1.

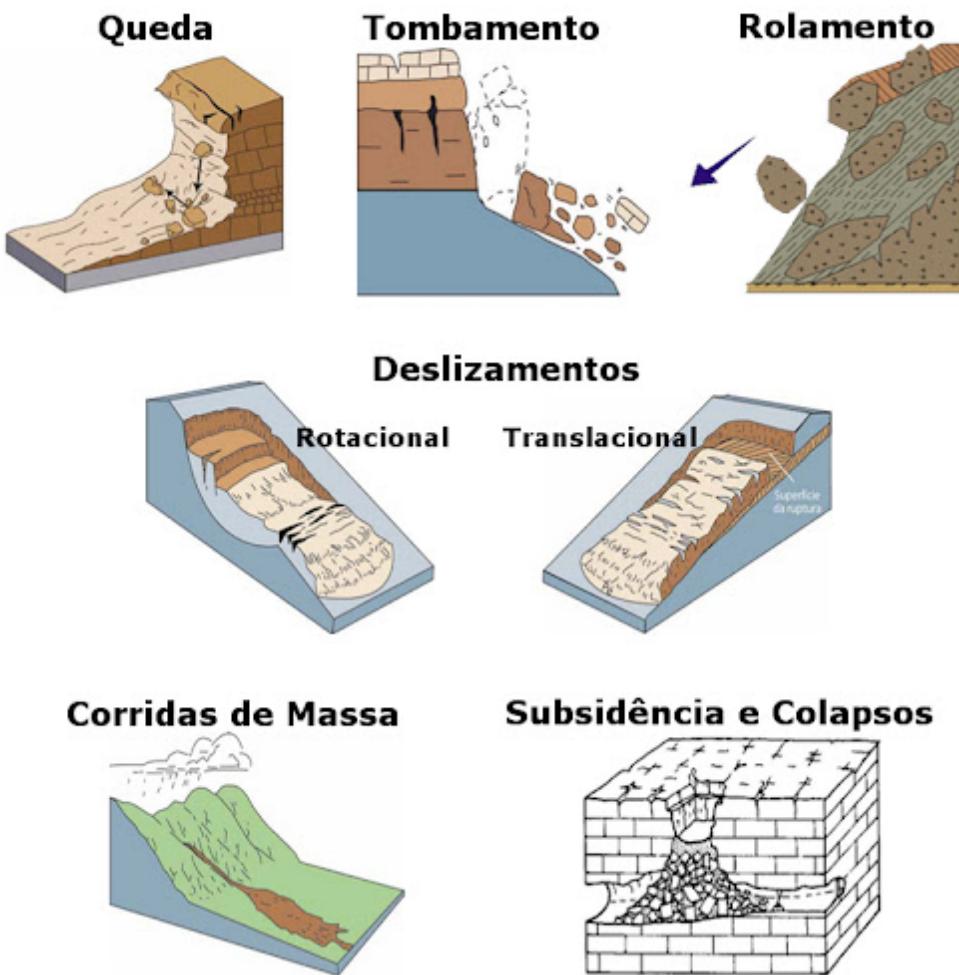


Figura 3.1: Principais tipos de movimentos de massa.

Fonte: CEMADEN

3.1.1 Rastejo

Os rastejos são caracterizados por serem lentos e contínuos, de poucos centímetros ao ano (cm/ano), não apresentam geometria e superfície de ruptura definidas e geralmente ocorrem em camadas superficiais de solo e de transição entre o solo e a rocha.

3.1.2 Queda de Blocos

As quedas de blocos rochosos são movimentos rápidos de blocos e/ou lascas de rocha, resultantes da ação gravitacional. São provocadas por fenômenos como a dilatação e contração térmica, sendo predominantes em encostas íngremes de paredões rochosos. Podem ser classificadas em quatro tipos; queda de blocos, que envolve o desprendimento de blocos do maciço em queda livre encosta abaixo; tombamento

Tabela 3.1: Principais tipos de movimentos de massa em encostas.

| Processos | - Dinâmica/Geometria/Material |
|--------------------------|---|
| Rastejo | <ul style="list-style-type: none"> - vários planos de deslocamento (internos) - velocidades muito baixas (cm/ano) e decrescentes com a profundidade - movimentos contínuos, sazonais ou intermitentes - solo, depósitos, rocha alterada/erodida - geometria indefinida |
| Escorregamentos | <ul style="list-style-type: none"> - poucos planos de deslocamento (externos) - velocidades médias (m/h) a altas (m/s) - pequenos a grandes volumes de material - geometria e materiais variáveis: Planares, Circulares, Em cunha |
| Quedas de Blocos | <ul style="list-style-type: none"> - sem planos de deslocamento - movimentos tipo queda livre ou em plano inclinado - velocidades muito altas (vários m/s) - material rochoso - pequenos a médios volumes - geometria variável: lascas, placas, blocos - Rolamento - Tombamento |
| Corridas de Massa | <ul style="list-style-type: none"> - muitas superfícies de deslocamento - movimento semelhante ao de um líquido viscoso - desenvolvimento ao longo das drenagens - velocidades médias a altas - mobilização de solo, rocha, detritos e água - grandes volumes de material - extenso raio de alcance, mesmo em áreas planas |

Fonte: TOMINAGA *et al.* (2009), AUGUSTO FILHO (1992).

de blocos, que consiste no movimento de rotação de blocos rochosos; rolamento de blocos, sendo movimento de blocos rochosos ao longo da encosta.

3.1.3 Escorregamento

Os escorregamentos, ou deslizamentos, são caracterizados por movimentos rápidos com um plano de ruptura bem definido, possuem curta duração e velocidade mais elevada proporcional à inclinação da superfície. Segundo AUGUSTO FILHO (1992), os escorregamentos são classificados com base na forma do plano de ruptura, podendo ser translacionais, rotacionais ou em forma de cunha.

Um escorregamento translacional é um tipo de deslizamento de terra que ocorre quando uma massa de solo, rocha e vegetação desce em encostas íngremes. É um fenômeno comum em montanhas e vales, onde os solos são pouco profundos.

Um escorregamento rotacional é um tipo de deslizamento de terra em que a massa de solo ou rocha se move em torno de um eixo, em um movimento aproximadamente rotatório.

3.1.4 Corrida de Massa

As corridas de massa são fenômenos caracterizados por movimentos de média a alta velocidade, nos quais grandes volumes de material se deslocam encosta abaixo, assemelhando-se a um fluido viscoso. Podem ser compostas por uma variedade de materiais, como solo pouco coesos, rocha, matéria orgânica, ar e água. São classificadas de acordo com as características do material em movimento e a velocidade do processo, sendo: corrida de lama, corrida de terra e corrida de detritos.

3.2 Fatores que Influenciam a Ocorrência de Movimento de Massa

Algumas causas da instabilização de encostas são variáveis, como a infiltração de água de chuva e a água subterrânea, enquanto outras são relativamente constantes, como a composição e estrutura do solo, além da inclinação de taludes naturais. A Tabela 3.2 apresenta uma subdivisão dos principais fatores que influenciam os movimentos de massa, conforme classificados por CRUDEN e VARNES (1996) em quatro grandes grupos.

Tabela 3.2: Fatores que Influenciam a Ocorrência de Movimento de Massa, adaptado de CRUDEN e VARNES (1996).

| | |
|----------------------|---|
| Agentes Geológicos | - Materiais cisalhados, fissurados ou fraturados - Contrastes na permeabilidade - Contrastes na espessura |
| Agentes Morfológicos | - Tectonismo ou vulcanismo - Qualquer tipo de erosão - Remoção da vegetação |
| Agentes Físicos | - Chuva intensa ou excepcionalmente prolongada - Intemperismo por congelamento e degelo - Degelo repentino, abalo sísmico, erupção vulcânica - Intemperismo de material expansivo |
| Agentes Antrópicos | - Escavação de encosta ou da sua base - Sobrecarga no topo ou na encosta - Diminuição do nível dos reservatórios - Desmatamento, irrigação, mineração, vazamentos - Vibração artificial |

Em relação ao município de São Bernardo do Campo (SBC), os principais agentes causadores de deslizamentos de terra registrados no banco de dados são fatores físicos, como a infiltração de água no solo devido à precipitação. Fatores antrópicos, associados ao processo de urbanização, também exercem influência, com registros

de deslizamentos causados por vazamentos nos sistemas de água ou esgoto, além de escavações para execução de obras.

3.3 Gestão de Risco de Deslizamentos em São Bernardo do Campo

A coordenadoria Municipal de Defesa Civil (COMDEC) é o órgão municipal responsável por fornecer apoio à população desenvolvendo o trabalho de atuação cotidiana de atendimentos e registro das ocorrências no que diz respeito a desastres naturais e pela execução, coordenação e mobilização de todas as ações de defesa civil no município.

No município de São Bernardo do Campo a prevenção de riscos de deslizamento segue as diretrizes do Plano Preventivo de Defesa Civil (PPDC), instituído em resposta a incidentes graves ocorridos durante os períodos chuvosos de 1987 e 1988, (decreto estadual nº 30.860 de 04/12/1989, redefinido pelo Decreto nº 42.565 de 01/12/1997), cujo principal objetivo é prevenir vítimas mortais, removendo preventivamente as populações que ocupavam áreas de risco antes que os deslizamentos atinjam suas moradias. O PPDC se baseia no monitoramento dos valores mínimos necessário de precipitação para desencadear um deslizamento (limiares críticos), obtidos a partir do estudo de correlação entre chuvas e deslizamentos em cada município, e nas vistorias de campo para verificação de evidências de movimentação dos terrenos (trincas, degraus de abatimento, etc.). Inicialmente TATIZANA *et al.* (1987a,b) definiu os limiares críticos para o município de Cubatão, Estado de São Paulo, tendo como referência os dados de precipitação acumulada (em mm) para um período de 72 horas (3 dias). Estes limiares foram utilizados na elaboração do PPDC, posteriormente foram extrapolados para outras regiões em São Paulo.

O PPDC estabelece quatro níveis operacionais, quais sejam: observação, atenção, alerta e alerta máximo. Os limiares usados como critério de entrada para cada nível são separados por região. Para o município de São Bernardo do Campo, os critérios são os mesmos adotados nos demais 37 municípios da Região Metropolitana de São Paulo, com exceção da Capital. A alteração dos níveis operacionais segue o seguinte procedimento: quando o valor da chuva acumulada de 72 horas ultrapassa os limiares críticos 80 mm, o nível operacional é alterado de “Observação” para “Atenção” e vistorias de campo são requeridas para avaliação dos terrenos. Caso seja constatado o registro de trincas, degraus ou qualquer outra feição de instabilidade, o nível operacional é alterado para “Alerta” e ações de remoção preventiva da população das áreas de risco iminente, indicadas pelas vistorias de campo. O nível operacional “Alerta Máximo” é adotado no caso de ocorrência de deslizamentos generalizados

no município e predição de chuvas com tendência de longa duração, como mostra a Tabela 3.3.

Em 2010 o município elaborou o Plano Municipal de Redução de Riscos (PMRR) (SÃO BERNARDO, 2010), o plano foi atualizado em 2021, e diagnosticou 34 áreas de risco contendo 126 setores de risco com 2010 moradias ou edificações com outros usos. O custo estimado para mitigação dos riscos nas áreas analisados no estudo é de R\$55.668.462,92 (COTTING *et al.*, 2024), as figuras 3.2, 3.3 e 3.4 apresentam alguns setores de risco delimitados pelo PMRR. Através do Decreto n° 17.303 o município estabeleceu a Operação Guarda-Chuva (OGC), visando adoção de procedimentos, para maior rapidez e eficácia às ações emergenciais de resposta, capacitação dos agentes públicos, incentivo à participação comunitária por meio de campanhas de conscientização e orientação e monitoramento em campo contínuo dos setores de risco mapeados pelo PMRR (BONGIOVANNI *et al.*, 2011).



Figura 3.2: Destaque para situação da encosta com solo exposto, cicatriz de deslizamento, árvores inclinadas.

Fonte: PMRR, 2021

Tabela 3.3: Principais ações e critérios dos níveis dos Planos Preventivos.

| NÍVEL | ENTRADA | SAÍDA | AÇÃO/MEDIDA PRINCIPAL |
|---------------|---|--|---|
| OBSERVAÇÃO | Início do período de vigência do plano – 1 de dezembro de cada ano. | Término do período de vigência – 31 de março. | Acompanhamento dos índices pluviométricos e da previsão meteorológica. |
| ATENÇÃO | Região Metropolitana de São Paulo: Acumulado de chuvas $\geq 80mm$ e previsão de chuvas com tendência de LONGA DURAÇÃO de QUALQUER intensidade ou Recebimento de informação de Risco de escorregamento remetida pelo CEMADEN. | previsão de não ocorrência de chuvas com tendência de LONGA DURAÇÃO de QUALQUER intensidade e Acumulado de chuvas $< 80mm$ em 72 h OU Recebimento da informação e Cessar Risco do CEMADEN, após mínimo de 24h da mudança de nível (aplicável apenas quando o ingresso se deu em razão de informação do CEMADEN). | Vistorias de campo nas áreas de risco anteriormente identificadas; Avaliação da necessidade de mudança de nível. |
| ALERTA | Registro de trincas, degraus ou qualquer outra feição de instabilidade em áreas habitadas que indique a possibilidade de escorregamentos, observadas em vistorias de campo, nas áreas de risco e nas proximidades. | previsão de não ocorrência de chuvas de longa duração (qualquer intensidade) e parcer favorável do – Instituto Geológico (IG); Restauração das vias e sistemas de drenagem e outras recomendações. | Remoção preventiva da população das áreas de risco eminentes, indicadas pelas vistorias de campo; implantar as ações recomendadas no relatório técnico emitido pelo IG. |
| ALERTA MÁXIMO | Registro de ocorrências generalizadas de escorregamentos nas áreas de risco ou proximidades e previsão de chuvas com tendência de longa duração. | previsão de não ocorrência de chuvas de longa duração (qualquer intensidade) e parcer favorável do IG; Restauração das vias e sistemas de drenagem e outras recomendações. | Proceder os mesmos itens definidos para o nível de alerta e; Remoção de toda a população das áreas de risco alto e muito alto. |

Fonte: PPDC 2017.



Figura 3.3: Deslizamento SBC.

Fonte: PMRR, 2021



Figura 3.4: Área de risco, SBC.

Fonte: PMRR, 2021

O Município de São Bernardo do Campo dispõe de volumoso acervo de estudos como: mapa geológico (PERROTTA *et al.*, 2006), cartas geotécnicas de aptidão à

urbanização (NOGUEIRA e CANIL, 2017), incidência de áreas suscetíveis a movimentos gravitacionais de massa e inundações em municípios abrangidos pelo Plano Nacional de Gestão de Risco e Respostas a Desastres Naturais, cartas de suscetibilidade a movimentos gravitacionais de massa e inundações (BITAR, 2014). Contudo a maioria dos estudos foca na criação de limiares críticos de precipitação e mapas de atributos geotécnicos e de suscetibilidade. No entanto, não há registros na literatura de modelos preditivos para alertas antecipados de deslizamentos que considerem simultaneamente as características espaciais e temporais da região.

3.4 Aviso Prévio de Deslizamento

O alerta antecipado é essencial para identificar indicadores de escorregamento de solo e mitigar danos, através a evacuação de residentes em áreas de risco potencial. A detecção precoce de um deslizamento pode ser realizada por meio do monitoramento em tempo real ou quase em tempo real de chuvas e das propriedades físicas do solo, com abrangência global, regional ou local (CHAE *et al.*, 2017). Em geral, os métodos de predição de deslizamentos de terra podem ser categorizados em três tipos: abordagens empíricas, estatísticas e físicas.

3.4.1 Limiares Empíricos

A maioria dos métodos de alerta antecipado que adota uma abordagem empírica são baseados na estimativa de limiares críticos de precipitação, utilizando dados históricos de precipitação em eventos com e sem deslizamentos de terra. O termo “limiar” define o nível mínimo ou máximo (crítico) de alguma quantidade necessária para que um processo ocorra (REICHENBACH *et al.*, 1998). Trabalhos pioneiros incluem estudos de (CAINE, 1980; LUMB, 1975), que estimaram uma curva de limiar inferior derivada de dois parâmetros: intensidade e duração de precipitação, associados a deslizamentos de terra. Esses limiares podem ser representados em coordenadas cartesianas, semi-logarítmicas ou logarítmicas. Uma revisão de literatura realizada por GUZZETTI *et al.* (2007) revelou que a maioria dos estudos publicados não especifica os critérios matemáticos ou estatísticos para a determinação dos limiares, que geralmente são definidos visualmente como o limite inferior das condições de precipitação associadas a deslizamentos. No entanto, esse procedimento, embora produza limiares visualmente aceitáveis, não são facilmente reproduzíveis.

Com base em um banco de dados contendo 2.626 eventos de precipitação que resultaram em deslizamentos de terra superficiais e fluxos de detritos em diferentes regiões do mundo, GUZZETTI *et al.* (2008) propuseram uma metodologia objetiva para a definição de limiares de precipitação, utilizando um modelo baseado na lei

de potência, equação 3.1.

$$I = \alpha D^{-\beta} \quad (3.1)$$

sendo I a intensidade média da precipitação (em mm/h), D é a duração da chuva (em h), e α e β são coeficientes positivos estimados por meio de inferência bayesiana. Os novos limites globais mínimos são significativamente inferiores aos outros limites globais propostos na literatura, os quais se baseavam em interpolação visual para ajustar as curvas. Segundo os autores, essa melhora resulta do uso de um banco de dados mais amplo e da aplicação de inferência bayesiana. BRUNETTI *et al.* (2010) propõem um método baseado em análise de frequência, primeiramente a intensidade e duração da precipitação são transformados em logaritmo, depois os dados que resultaram em deslizamentos de terra são ajustados pelo método dos mínimos quadrados com uma equação linear 3.2, equivalente à lei de potência da equação 3.1 em coordenadas lineares.

$$\log(I) = \log(\alpha) - \beta \log(D) \quad (3.2)$$

Em seguida, a diferença $\delta(d)$ entre $\log(I)$ e o valor predito é calculada. Então, a função de densidade de probabilidade da distribuição $\delta(d)$ é determinada através da técnica não paramétrica de estimativa da densidade do kernel (KDE) (SCOTT, 1992; SILVERMAN, 1986; VENABLES e RIPLEY, 2002) e modelada como uma distribuição gaussiana. Os limiares são definidos por diferentes probabilidades de excedência da distribuição ajustada. Por exemplo, para um limite de 5% (T5), é calculada a distância δ^* entre 5% e a média da distribuição, e esse valor é subtraído do intercepto, resultando em $\log(\alpha_5) = \log(\alpha_{50}) - \delta^*$. Isso gera uma curva paralela à curva de melhor ajuste, com o intercepto deslocado, figura 3.5. Assumindo que o catálogo de deslizamentos seja suficientemente completo e representativo da região, a probabilidade de ocorrer algum deslizamento desencadeado por precipitação abaixo desse limiar é inferior a 5%.

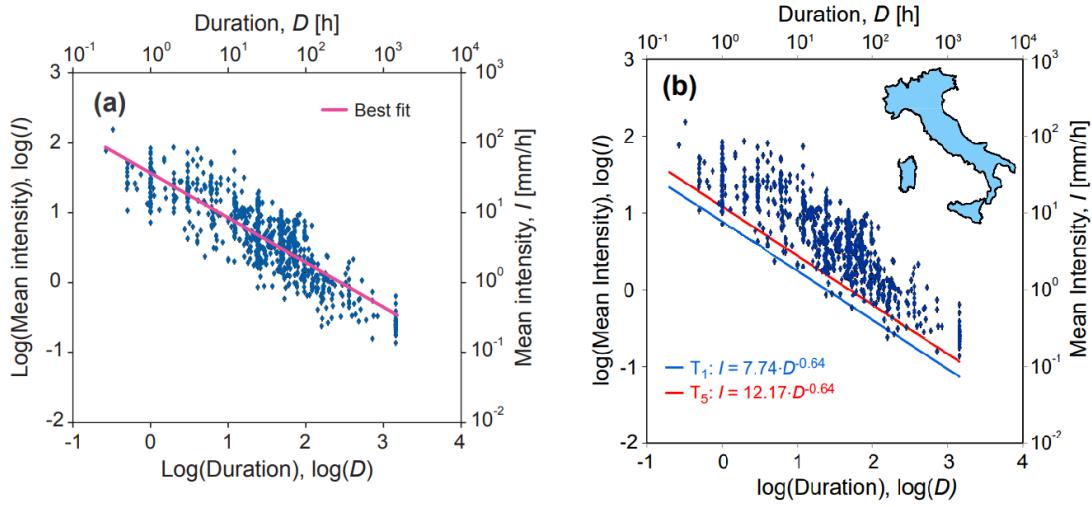


Figura 3.5: a) curva ajustada com 753 eventos resultantes em deslizamento. b) Curva de limiar com 1% e 5% de excedência.

Fonte: adaptado de, BRUNETTI *et al.* (2010).

A metodologia apresentada por BRUNETTI *et al.* (2010) foi adaptada por PERUCCACCI *et al.* (2012) para determinar limiares com base na precipitação acumulada e na duração, definida pela equação 3.3.

$$E = (\alpha \pm \Delta\alpha) \cdot D^{(\gamma \pm \Delta\gamma)} \quad (3.3)$$

onde, E é a precipitação acumulada do evento (mm), D é a duração do evento de precipitação (h), α um coeficiente positivo que define o intercepto da curva, γ o parâmetro de forma que controla a inclinação da curva, e $\Delta\alpha$ e $\Delta\gamma$ são as incertezas associadas a α e γ , respectivamente. O método utilizado para estimar os parâmetros α e γ é o mesmo utilizado anteriormente por BRUNETTI *et al.* (2010) com a adição do método não paramétrico bootstrap, que avaliar a distribuição média de amostras de uma população a partir de um conjunto de dados empíricos (EFRON, 1992; EFRON e TIBSHIRANI, 1994), no caso os valores médios de α e γ e as incertezas associadas a $\Delta\alpha$ e $\Delta\gamma$.

Na Itália, ROSSI *et al.* (2012) desenvolveram um sistema nacional de alerta precoce para deslizamentos de terra induzidos por chuvas chamado SANF. O sistema utiliza os seguintes parâmetros:

1. Limites de precipitação para possível ocorrência de deslizamentos de terra;
2. Medições de precipitação sub-horárias obtidas por uma rede nacional de pluviômetros de 1950;
3. Previsões quantitativas de precipitação.

Para definir os limites, eles utilizam a equação 3.1. São estabelecidos cinco níveis críticos para o sistema, denominados: muito abaixo do limite, abaixo do limite, no limite, acima do limite e muito acima do limite. Esses níveis são determinados calculando a probabilidade de ocorrência de deslizamentos, com os limiares definidos em 0,005%, 0,5%, 1,5% e 5%. Por exemplo, quando a intensidade média da precipitação (I) está entre os limites de 0,005% e 0,5%, a condição de precipitação é considerado abaixo do limite. O sistema compara os valores medidos e previstos de precipitação com os limites de intensidade e duração predefinidos duas vezes por dia, atribuindo a cada pluviômetro uma probabilidade de ocorrência de deslizamentos de terra.

Um sistema de alerta regional de deslizamento denominado SIGMA (Sistema integrado para gestão e monitoramento do alerta) foi desenvolvido para a região Emília Romagna, no norte da Itália. O sistema de alerta SIGMA é um algoritmo de decisão que combina diferentes limiares de precipitação e define um nível de alerta correspondente (MARTELLONI *et al.*, 2012). O estudo se baseou na hipótese de que valores anômalos ou extremos de chuvas são responsáveis pelo desencadeamento de deslizamentos. Para implementação do sistema de alerta SIGMA, a região foi dividida em 19 unidades territoriais (UT), principalmente com base em características fisiográficas e ambientais distintas. Cada UT é equipada com um pluviômetro automatizado que coleta leituras em intervalos de uma hora. Para cada pluviômetro são calculados os valores de precipitação acumulada, considerando acumulados de 1, 2 e 3 dias para deslizamentos rasos, e um acumulado de intervalo de tempo variável (até 240 dias) para deslizamentos em solos profundos ou de baixa permeabilidade. Os limiares são baseados no desvio padrão (σ) das séries temporais de precipitação acumulada, para cada UT são definidos valores de $1,5\sigma$, 2σ , $2,5\sigma$, 3σ , calibrados por meio de um algoritmo de otimização utilizando dados de 2004–2007, que contêm 888 ocorrências de deslizamentos. O período de 2008–2010, com 764 deslizamentos, foi utilizado para validar o modelo e 84% deles foram previstos corretamente, porém o modelo apresenta um elevado número de falso positivo.

No Brasil, TATIZANA *et al.* (1987a) conduziram um estudo de correlação entre precipitação acumulada e a intensidade que desencadeou deslizamentos para a município de Cubatão, SP. Utilizando dados registrados ao longo de 30 anos, os autores propuseram um ajuste exponencial expresso pela equação

$$I = k \cdot AC^{-b} \quad (3.4)$$

onde, I é a intensidade da precipitação (em mm/h), AC é o acumulado de chuva em 4 dias (em mm), e k e b são coeficientes positivos, variáveis de acordo com as propriedades geotécnicas e climáticas da região. Os coeficientes k e b foram calculados utilizando o método dos mínimos quadrados, ajustando a curva para

separar eventos com e sem registro de deslizamento. Essa metodologia é aplicada individualmente para cada sensor pluviométrico.

Foram conduzidos numerosos estudos analisando a correlação entre chuvas e deslizamentos até o momento, GUZZETTI *et al.* (2007) listaram 25 variáveis pluviométricas e climáticas descritas na literatura para estabelecer limiares de precipitação que indicam o início de deslizamentos. No entanto, não há um consenso estabelecido sobre os parâmetros ou critérios técnicos a serem adotados para medir as condições de precipitação que desencadeiam deslizamentos de terra, essa subjetividade dificulta a comparação de diferentes medições, restringe a definição e a comparação de limiares de precipitação e limita o uso desses limiares em sistemas de alerta antecipado. A estabilidade das encostas não é afetada apenas pela precipitação, portanto limiares críticos baseados na precipitação, quando utilizados de forma isolada, podem não representar as situações críticas, pois não consideram a umidade nos perfis do solo nas encostas, nem os processos de infiltração, retenção e drenagem da água da chuva. Segundo BERTI *et al.* (2012) nem sempre é possível obter uma fronteira distinta entre os eventos de chuva que provocam deslizamentos de terra e os eventos de chuva que não provocam deslizamentos de terra. Nesse caso, mesmo que o limite de precipitação seja excedido, nem sempre ocorrerão deslizamentos de terra. Portanto, a medição da umidade do solo em tempo real, em conjunto com os limiares de chuva, oferece a predição mais confiável e devem ser usados em sistemas de alerta antecipado (TOLL *et al.*, 2011).

Para propor limiares críticos baseados em precipitação e umidade do solo, MENDES *et al.* (2020) analisaram as ocorrências de deslizamentos em três municípios de São Paulo: Campos do Jordão, São José dos Campos e Ubatuba. O estudo identificou que os deslizamentos foram desencadeados pelo aumento da umidade do solo de 1,0 a 3,0 metros de profundidade, com valores críticos de umidade do solo entre 38,5% e 46%, além da intensidade da chuva horária e das chuvas acumuladas de 24 horas e 72 horas. Uma vantagem dessa abordagem em comparação com a do PPDC é a possibilidade de transitar entre os níveis operacionais, por exemplo de “OBSERVAÇÃO” para “ATENÇÃO”, sem a necessidade de vistoriar as áreas de risco in loco, quando os limiares críticos de chuva e umidade do solo são simultaneamente atingidos.

Em Washington, EUA, MIRUS *et al.* (2018) desenvolveram um sistema de alerta antecipado para deslizamentos de terras rasas induzidos pelas chuvas, para atuar ao longo do corredor ferroviário propenso a deslizamentos de terra entre Everett e Seattle. O sistema é baseado na equação proposta por (CHLEBORAD *et al.*, 2008),

$$P_3 = aP_{15} + b \quad (3.5)$$

onde P_3 e P_{15} representam a precipitação acumulada de 3 dias e 15 dias, respectivamente, e a e b são coeficientes positivos. O estudo propõe limiares híbrido de precipitação versus saturação do solo, substituindo o componente P_{15} por uma saturação média observada durante o mesmo período de tempo, S_{15} . Diferentemente da precipitação o teor de água no solo normalmente não atinge zero em condições de campo devido à água retida por forças capilares e adsorção, portanto o autor assume que o ponto de murchamento do solo ocorre na saturação média mínima registrada durante o período de monitoramento ($S_{15} = 0,49$). Gerando a equação 3.6. Os parâmetros da equação são calibrados por meio de um algoritmo de otimização utilizando dados ocorrência de deslizamento.

$$P_3 = a(S_{15} - 0,49) + b \quad (3.6)$$

Os limiares híbridos resultaram em uma maior precisão na predição de ameaças em comparação com versões que utilizam apenas precipitação, demonstrando uma taxa mais alta de verdadeiros positivos e taxas mais baixas de falsos positivos e falsos negativos para o mesmo conjunto de dados.

Com o propósito de aprimorar o sistema de alerta SIGMA (MARTELLONI *et al.*, 2012), SEGONI *et al.* (2018) investigaram duas metodologias para a integração de estimativas de umidade média do solo nesse sistema. Os valores de umidade são obtidos por um modelo de base física (CIARAPICA e TODINI, 2002) para cada unidade territorial (UT). O modelo mais simples adota um limiar de saturação do solo de 75%, derivado da média aritmética dos valores de cada unidade territorial (UT), como critério de decisão no algoritmo do SIGMA. A nova versão do algoritmo verifica o valor de saturação de uma UT, se estiver abaixo de 75%, nenhum deslizamento é esperado, acima desse limiar, o algoritmo do SIGMA é ativado para verificar as condições específicas de precipitação. A análise retrospectiva dos dados nas UT revela uma redução significativa nos alarmes falsos: uma diminuição de 28% nos alarmes falsos de primeiro nível (de 320 para 231), 17% nos alarmes falsos de segundo nível (de 169 para 141), e uma redução de 62% nos alarmes falsos de terceiro nível (de 13 para 5), com apenas um aumento de um alarme perdido, composto por três deslizamentos de terra. A segunda abordagem adotada por (SEGONI *et al.*, 2018) segue o mesmo procedimento utilizado para calcular os valores de desvio padrão das séries temporais de precipitação em (MARTELLONI *et al.*, 2012), aplicando-o às séries temporais de umidade do solo. Os limiares de desvio padrão de umidade do solo são incluídos no algoritmo de decisão do SIGMA, substituindo os valores de precipitação acumulado com longos períodos de até 243 dias por umidade do solo. A nova versão do algoritmo, denominada SIGMA-U incorpora limiares de precipitação acumulado de 1, 2 e 3 dias, juntamente com umidade

do solo acumulado para períodos de 1 a 15 dias. A análise retroativa mostra que o SIGMA-U apresenta uma redução de 22% nos alarmes perdidos em comparação com o SIGMA. Quanto aos falsos alarmes, observa-se uma redução significativa à medida que o nível de alarme aumenta: 8% (de 780 para 721) no nível 1, 31% (de 287 para 197) no nível 2 e 48% (de 22 para 11) no nível 3.

3.4.2 Abordagens Estatísticas

Como as condições de estabilidade de um talude geralmente resultam de uma combinação complexa de diversos fatores, os limiares de precipitação obtidos com base nos métodos empíricos, frequentemente, representam apenas os valores mínimos que podem desencadear deslizamentos de terra, com base nos dados disponíveis. Portanto, um certo grau de incerteza inerente à modelagem dos limiares é inevitável. Além disso, a baixa qualidade dos dados pode ampliar essa incerteza no modelo. Para considerar a incerteza nos limiares de precipitação, FRATTINI *et al.* (2009) propuseram o método de regressão logística, o qual permite definir a probabilidade de uma determinada intensidade de chuva, em uma dada duração, possa desencadear um deslizamento de terra. Esse método envolve a separação dos eventos de chuva em dois grupos principais: aqueles associados a eventos de deslizamento e aqueles que não o estão. Para cada evento associado a deslizamento, são coletadas medições de precipitação do pluviômetro mais próximo do local do deslizamento.

BERTI *et al.* (2012) propuseram um método Bayesiano bidimensional, que calcula a probabilidade condicional de ocorrência de deslizamentos (A) dada a ocorrência conjunta de duas variáveis de controle, no caso intensidade média diária (B) e duração (C) da precipitação, equação 3.7.

$$P(A | B, C) = \frac{P(B, C | A) \cdot P(A)}{P(B, C)} \quad (3.7)$$

sendo:

$P(B, C | A)$ = probabilidade condicional de B e C dado A (também chamada de verossimilhança), ou seja, a probabilidade de observar um evento de chuva de intensidade média B e duração C quando ocorre um deslizamento de terra A,

$P(A)$ = probabilidade anterior de A, ou seja, a probabilidade de ocorrer um deslizamento de terra independentemente de ocorrer ou não um evento de precipitação,

$P(B, C)$ = probabilidade marginal de B e C , ou seja, a probabilidade de observar uma precipitação com intensidade média B e duração C independentemente de ocorrer ou não um deslizamento de terra,

$P(A | B, C)$ = probabilidade condicional de A dado B e C (também chamada de probabilidade posterior), que é a probabilidade de observar um deslizamento de terra quando ocorre um evento de chuva de intensidade média B e duração C.

Na metodologia proposta por BERTI *et al.* (2012), todos os eventos de precipitação registrados são plotados no gráfico de intensidade e duração, Figura 3.6. Pontos pretos indicam eventos que resultaram em deslizamentos de terra, enquanto os pontos cinzas representam aqueles que não resultaram em deslizamento. O gráfico é dividido em uma grade de 9×13 células, e a probabilidade de ocorrência de deslizamentos é calculada para cada célula utilizando a equação 3.7.

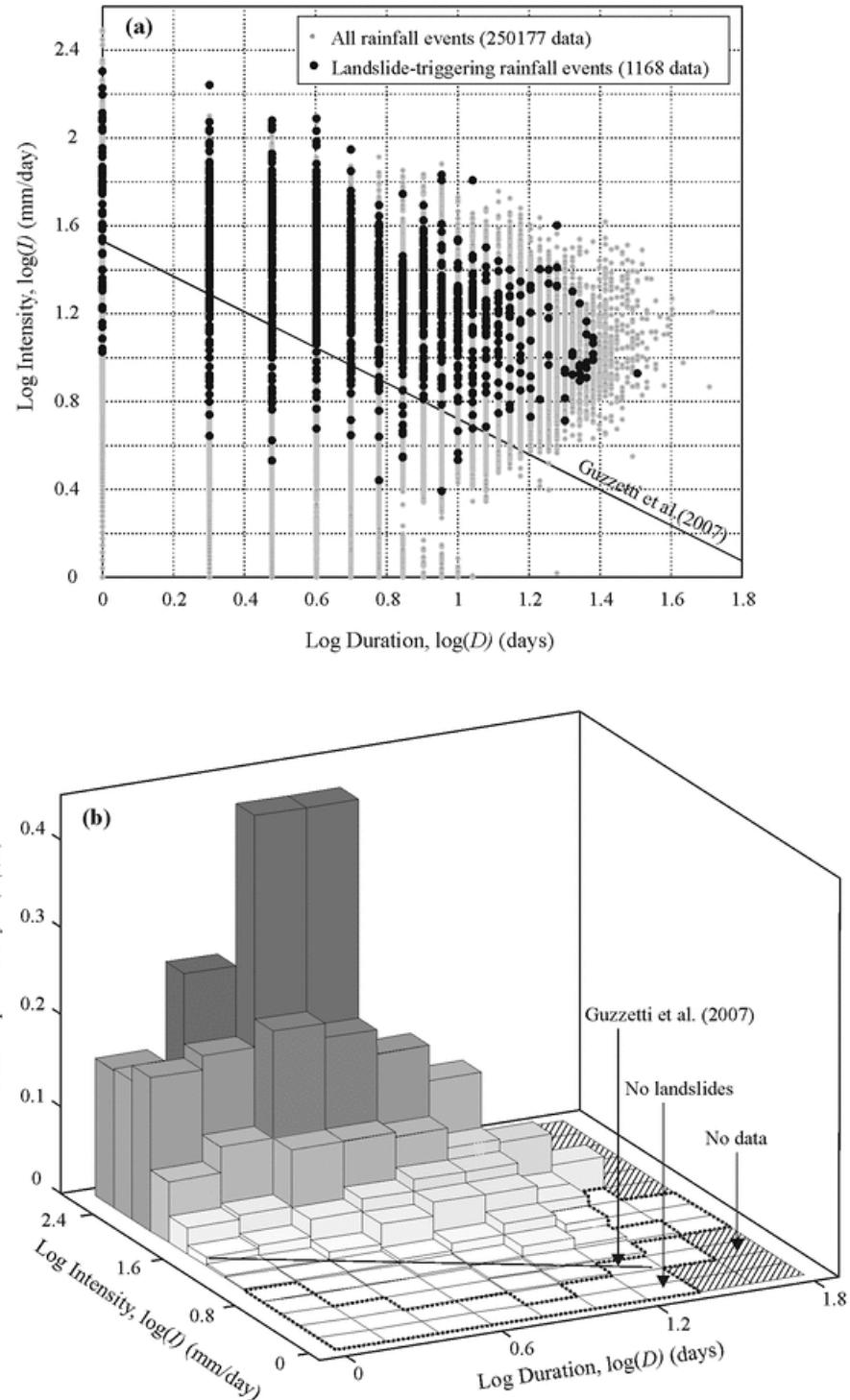


Figura 3.6: Análise bayesiana bidimensional. (a) Gráfico de intensidade - duração. (b) Histograma da probabilidade de deslizamento de terra.

Fonte: BERTI *et al.* (2012)

Segundo BERTI *et al.* (2012), uma das vantagens da abordagem bayesiana é a capacidade de expressar todas as incertezas em termos de probabilidades. A regra de Bayes leva em consideração que, em condições geológicas complexas, o mesmo evento de precipitação pode ou não gerar um deslizamento de terra, dependendo de diversos

fatores, como comportamento hidrológico, redução da resistência ao cisalhamento, taxa de deformação em longo prazo, falha progressiva ou intervenções antrópicas. Além disso, as previsões bayesianas podem ser atualizadas dinamicamente com novos dados, ajustando a probabilidade posterior $P(A|B, C)$ para $P(B, C|A)$.

Com o objetivo de avaliar o potencial da umidade do solo como alerta precoce de deslizamentos de terra regionais, WICKI *et al.* (2020) compararam dados de umidade do solo de diversos bancos de dados com um banco de dados nacional de deslizamentos de terra na Suíça, baseado em descrições de eventos em artigos de jornais e revistas. Um modelo de regressão logística foi aplicado para modelar a probabilidade de ocorrer um deslizamento com base nas características da umidade do solo. O método consiste em classificar os eventos em dois grupos: associados ou não a deslizamentos. Os parâmetros do modelo foram ajustados usando o método de máxima verossimilhança, visando otimizar a capacidade de discriminação entre as duas classes em comparação com as observações. Contudo, se o modelo não for capaz de separar perfeitamente as duas classes em comparação com as observações, a aplicação de um limite específico será sempre um compromisso entre maximizar o número de alarmes verdadeiros e minimizar o número de alarmes falsos. Diversos modelos com propriedades distintas foram analisados e avaliados quanto à acurácia, sendo que a melhor performance foi alcançada combinando a saturação antecedente com a mudança de saturação, resultando em uma acurácia de 75% na deflagração de deslizamento. Além disso, o modelo foi testado para diferentes distâncias de predição, variando de 5 a 40 km, revelando uma redução da qualidade da predição com o aumento da distância entre o local de medição de umidade do solo e o deslizamento de terra, sendo ideal uma distância de predição entre 10 e 15 km.

Os modelos anteriormente discutidos consideram apenas aspectos temporais de precipitação e umidade do solo que desencadeiam deslizamentos, ignorando outros fatores espaciais condicionantes. Uma abordagem amplamente difundida na literatura para incorporar os aspectos espaciais é a combinação de modelos estatísticos com Sistemas de Informações Geográficas (SIG) para avaliar a suscetibilidade a deslizamentos de terra. Essa suscetibilidade representa a probabilidade de ocorrência de deslizamentos em uma área, com base nas condições locais do terreno (BRABB, 1984), oferecendo estimativas quantitativas de "onde"esses eventos são esperados. Existem inúmeros estudos abordando modelos estatísticos de suscetibilidade a deslizamentos de terra. REICHENBACH *et al.* (2018) analisaram 565 artigos publicados em periódicos internacionais revisados por pares no período de janeiro de 1983 a junho de 2016, os métodos mais comuns nos artigos analisados são: regressão logística, rede neural, sobreposição de dados, análises baseadas em índice e peso de evidência, com uma preferência crescente por métodos de aprendizado de máquina nos últimos anos. Os autores relatam uma ampla gama de variáveis utilizadas nos modelos e

as classificam em cinco categorias principais: morfológica, geológica, cobertura da terra, hidrológica e outras. Entre elas, as variáveis morfológicas são as mais utilizadas, sendo as mais comuns: elevação, relevo, declive, aspecto e curvatura, que podem ser derivadas de Modelos Digitais de Elevação (MDE).

Em um estudo de caso realizado em Fanhões-Trancão, Portugal, FABBRI *et al.* (2003) identificaram o declive e o aspecto do terreno como as variáveis mais significativas para o modelo. Os autores concluíram que a relevância das variáveis varia conforme o tipo de deslizamento e o gatilho associado, destacando a importância do conhecimento e da compreensão dos processos de deslizamento na seleção de variáveis explicativas. Embora os modelos de suscetibilidade indiquem onde os eventos de deslizamentos são esperados, eles não permitem prever o momento ou a frequência desses eventos (GUZZETTI *et al.*, 2005). Para superar essa limitação, HUANG *et al.* (2022) combinaram mapas de suscetibilidade com limiares críticos de precipitação. Foram testados três modelos de aprendizado de máquina para gerar os mapas: regressão logística (LR), máquina de vetores de suporte (SVM) e floresta aleatória (RF), sendo RF o mais acurado. O sistema de alerta emite avisos quando os níveis críticos de precipitação são atingidos em áreas com alta suscetibilidade.

Tratar os dados de precipitação separadamente das características espaciais da paisagem negligencia a interação natural entre as variáveis. AHMED *et al.* (2023) propuseram uma versão Bayesiana do modelo aditivo generalizado (GAM) binomial que incorpora simultaneamente efeitos estáticos e dinâmicos. As covariáveis estáticas incluem variáveis padrão utilizadas na literatura de suscetibilidade, enquanto as dinâmicas abrangem a precipitação antecedente, variando de 1 a 14 dias, a janela de 8 dias foi selecionada como a combinação ideal, além de índices de vegetação com resolução temporal de 16 dias.

3.4.3 Abordagens Físicas

As Abordagens físicas ou determinísticas são baseados em modelos numéricos que consideram a relação entre precipitação, pressão da água nos poros, conteúdo volumétrico de água, tensão de sucção e estabilidade de encostas (CHAE *et al.*, 2017).

Os modelos determinísticos são de interesse para sistemas de alerta de deslizamentos de terra, pois podem prever a quantidade de precipitação necessária para desencadear falhas nas encostas, bem como a localização e o horário dos deslizamentos esperados (CHUNG *et al.*, 2017). Para entender como os processos hidrológicos influenciam a suscetibilidade a deslizamentos, IVERSON (2000) combinou um modelo de infiltração hidrológica com um modelo de estabilidade de talude infinito. O modelo de infiltração, baseado em uma solução aproximada da equação de Richards

(RICHARDS, 1931) no regime de difusão linear, permite modelar a hidrologia de solos quase saturados. A solução obtida para a carga de pressão induzida pela precipitação é então aplicada ao modelo de estabilidade de talude infinito para calcular os fatores de segurança da encosta.

O programa TRIGRS (do inglês, Transient Rainfall Infiltration and Grid-based Regional Slope-Stability), descrito por (BAUM *et al.*, 2002), é uma extensão do modelo de IVERSON (2000) que considera o escoamento superficial para dispersar o excesso de água e a intensidade variável da chuva para encostas dependentes do tempo e da profundidade. A ferramenta integra cálculos em uma estrutura GIS baseada em malha para prever a iniciação de deslizamentos de terra rasos induzidos por precipitação, célula a célula.

Segundo TSAI *et al.* (2008) os modelos anteriores baseado em física frequentemente simplificavam demais o processo de infiltração, usando uma equação de Richards do tipo difusão linear ou negligenciando a influência de forças internas do solo na resistência ao cisalhamento do solo não saturado. Semelhante a outros autores, TSAI *et al.* (2008) acoplaram modelos de análise de estabilidade de taludes infinitos com modelagem hidrológica, porém consideram a equação completa de Richards e o critério de falha de Mohr-Coulomb estendido FREDLUND *et al.* (1978) para descrever a resistência ao cisalhamento insaturado.

3.5 Algoritmo de Reconstrução de Evento de Precipitação e Infiltração

Como foi citado na seção 3.4.1, não há um consenso estabelecido sobre os parâmetros ou critérios técnicos a serem adotados para medir as condições de precipitação que desencadeiam deslizamentos de terra (GUZZETTI *et al.*, 2007). Uma inspeção na literatura revela que muitos artigos calculam a duração dos eventos de precipitação e a precipitação acumulada manualmente de forma heurística por um pesquisador ou um grupo de pesquisadores. Segundo MELILLO *et al.* (2015) os investigadores não especificam como os eventos de precipitação são determinados. Mesmo quando especificados os critérios utilizados para selecionar o sensor que representa o deslizamento, a forma em que foi calculada a precipitação acumulada e a duração do evento, muitas vezes é difícil determinar se foram aplicados de forma consistente.

De acordo com BERTI *et al.* (2012), identificar eventos de precipitação que desencadeiam deslizamentos é relativamente simples quando o mesmo ocorre após uma forte precipitação precedida por um período prolongado de seca. No entanto, essa identificação geralmente é desafiadora, pois os deslizamentos podem resultar de sequências complexas de precipitação com múltiplos episódios de intensidade e

duração variáveis, dificultando a determinação de um evento desencadeador. Um certo período de tempo sem precipitação pode ser um parâmetro usado como critério para separar a sequência de precipitação (BRUNETTI *et al.*, 2010).

Três dos autores de (BERTI *et al.*, 2012) definiram, de forma independente, os eventos de precipitação desencadeadores por análise visual e implementaram um algoritmo de detecção automatizado para avaliar possíveis discrepâncias de interpretação. Os eventos foram categorizados em três classes: bem definido (tipo 1), incerto (tipo 2) ou indefinido (tipo 3), correspondendo a 66 %, 10 % e 24 % dos casos, respectivamente. A figura 3.7 apresenta as diferenças entre o método manual aplicado pelos especialistas e o algoritmo das três classes estabelecidas. A identificação do início da precipitação desencadeadora foi identificada como uma das maiores fontes de incertezas na caracterização dos eventos, enquanto o término é definido pelo momento de ocorrência do deslizamento.

Para uma reconstrução objetiva dos eventos de precipitação que causaram deslizamentos de terra, MELILLO *et al.* (2015) propuseram um algoritmo que identifica as condições de chuva associadas aos deslizamentos, medindo a duração e a precipitação acumulada desses eventos. Para o escopo do algoritmo , um evento de precipitação em um registro de precipitação é um período de precipitação contínua ou um conjunto cronológico de períodos de precipitação contínua, separados de eventos de precipitação precedentes e sucessivos por períodos sem precipitação.

O algoritmo é dividido em dois blocos lógicos principais. O primeiro bloco realiza a reconstrução automática dos eventos de precipitação. O segundo bloco seleciona os eventos de precipitação que resultaram em deslizamentos de terra e determina a duração da precipitação, e a precipitação acumulada do evento, responsável pelos deslizamentos de terra. Os parâmetros que o algoritmo utiliza para realizar a reconstrução são: Sensibilidade do pluviômetro (0,2 mm), valor mínimo de detecção do sensor, valor mínimo de precipitação considerada relevante (0,2 mm), período seco para separar eventos de precipitação, considerado o inverso da evapotranspiração média anual.

O algoritmo provou ser capaz de reconstruir com precisão a maioria (87,7%) dos eventos de precipitação com menos de 10 % de diferença em relação ao método manual realizado tradicionalmente. A vantagem do algoritmo consiste no fato de que o mesmo realiza uma reconstrução objetiva e reproduzível dos eventos de precipitação, reduzindo a subjetividade na análise e o tempo necessário para determinar os eventos de precipitação.

O algoritmo (MELILLO *et al.*, 2015) foi aprimorado em 2018 pelo mesmo autor (MELILLO *et al.*, 2018) para selecionar automaticamente o pluviômetro mais adequado para cada deslizamento, considerando a distância entre o pluviômetro e o deslizamento. O cálculo da precipitação acumulada foi alterado para considerar

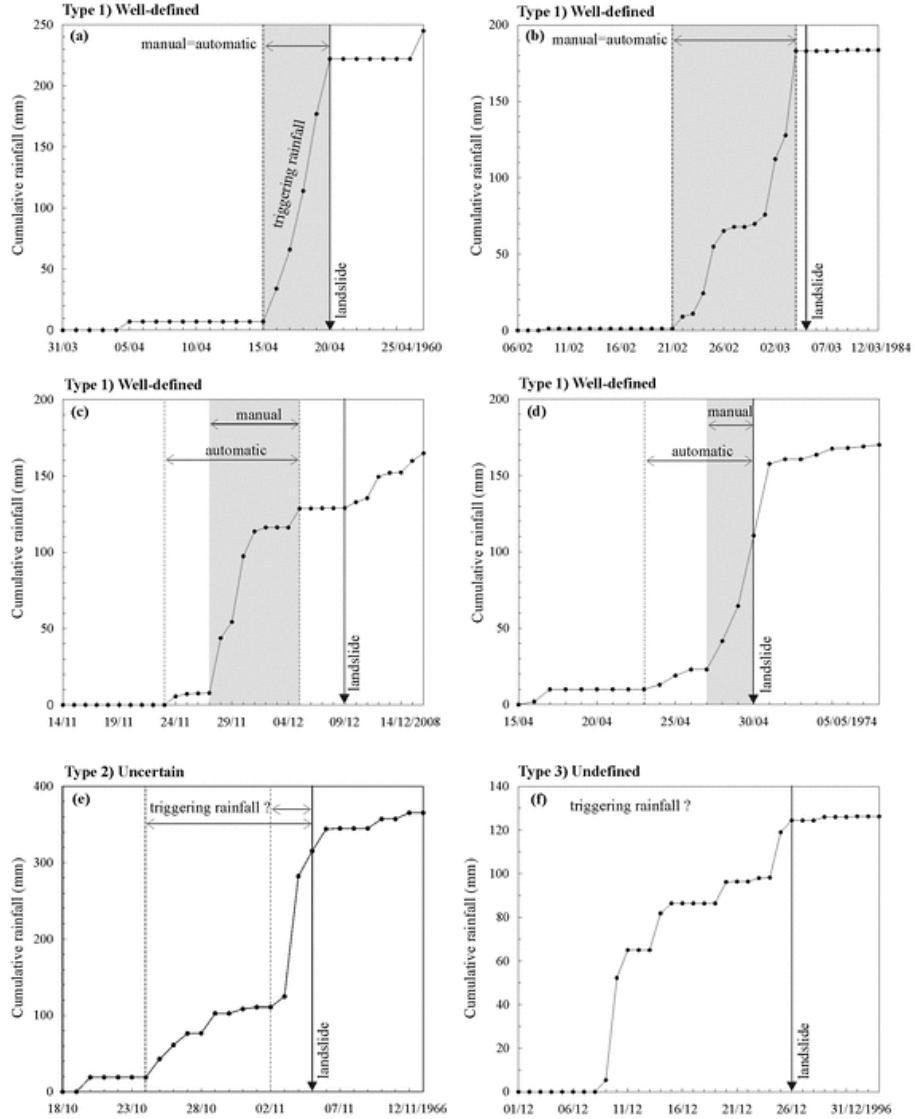


Figura 3.7: Classes indicando a facilidade de identificação dos eventos de precipitação. Exemplos de identificação manual (áreas cinzas) e automática (áreas tracejadas).

Fonte: BERTI *et al.* (2012)

o tempo decorrido desde o início da chuva. Um fator de decaimento constante (o artigo sugere, $k = 0.84$) é aplicado à chuva acumulada para representar a perda de água por infiltração e escoamento ao longo do tempo, conforme proposto por (CROZIER, 1999).

$$E_t = E_0 + kE_1 + k^2E_2 + \dots + k^nE_n = \sum_{i=0}^n k^i [E_i] \quad (3.8)$$

onde E_0 é a precipitação acumulada nas 24 h antes do tempo de ocorrência do deslizamento, E_i é a precipitação acumulada nas 24 h do i -ésimo dia antes do deslizamento, e n é a duração do evento de precipitação em dias. O modelo também incorpora procedimentos para calcular limiares de chuva, seguindo as metodologias de (BRUNETTI *et al.*, 2010; PERUCCACCI *et al.*, 2012).

Da mesma forma que a precipitação, os parâmetros para medir as condições da umidade do solo são de suma importância para os modelos. WICKI *et al.* (2020) utilizaram dados históricos de umidade do solo, de diversos banco de dados na Suíça, para reconstruir, de forma automática, eventos de infiltração que desencadearam deslizamento. As séries temporais de umidade do solo foram normalizadas pelo valor mínimo e máximo para representar a porcentagem de saturação do solo, e em seguida foram delimitados eventos de infiltração com o aumento de saturação de pelo menos 2% (ou 4% em conjuntos de dados ruidosos) e um intervalo mínimo de 3 horas entre eles, e então classificada como desencadeadora ou não desencadeadora de deslizamentos, com base na ocorrência de deslizamentos dentro de uma distância específica predefinida. Cada evento de infiltração é caracterizado pelas seguintes propriedades: saturação antecedente (saturação no início do evento), mudança de saturação do início ao fim do evento, duração do evento, taxa de infiltração (mudança de saturação dividida pela duração do evento), taxa máxima de infiltração em 3 horas, média e máxima saturação nas 2 semanas anteriores, desvio padrão antecedente e variação do desvio padrão. Os parâmetros do algoritmo , como o aumento mínimo de saturação e o intervalo mínimo entre eventos, foram determinados por meio de inspeção visual das séries temporais e ajustados para capturar o máximo de eventos individuais possível.

Após revisão bibliográfica, observa-se que existem estudos avaliando diversas variáveis pluviométricas e climáticas na literatura, porém não há um consenso estabelecido sobre os parâmetros ou critérios técnicos a serem adotados em sistemas de alerta antecipado. A incorporação da umidade do solo nos modelos tem demonstrado ser promissora, reduzindo o número de falsos alertas. Este estudo propõe um algoritmo desenvolvido para reconstruir eventos de infiltração e precipitação que desencadearam deslizamentos de terra. Os eventos reconstruídos são utilizados como entrada para o treinamento de dois modelos: um baseado em limiar crítico e uma regressão logística.

Capítulo 4

Materiais

Neste capítulo é apresentada a metodologia empregada neste trabalho para investigar as previsões de umidade do solo e a correlação entre dados pluviométricos, umidade do solo e deslizamento de terra, fornecendo subsídios para o aperfeiçoamento dos sistemas de alerta antecipado de deslizamento de terra. O organograma na Figura 4.1 ilustra os procedimentos adotados, desde a aquisição dos dados até o tratamento, modelagem e implementação dos modelos.

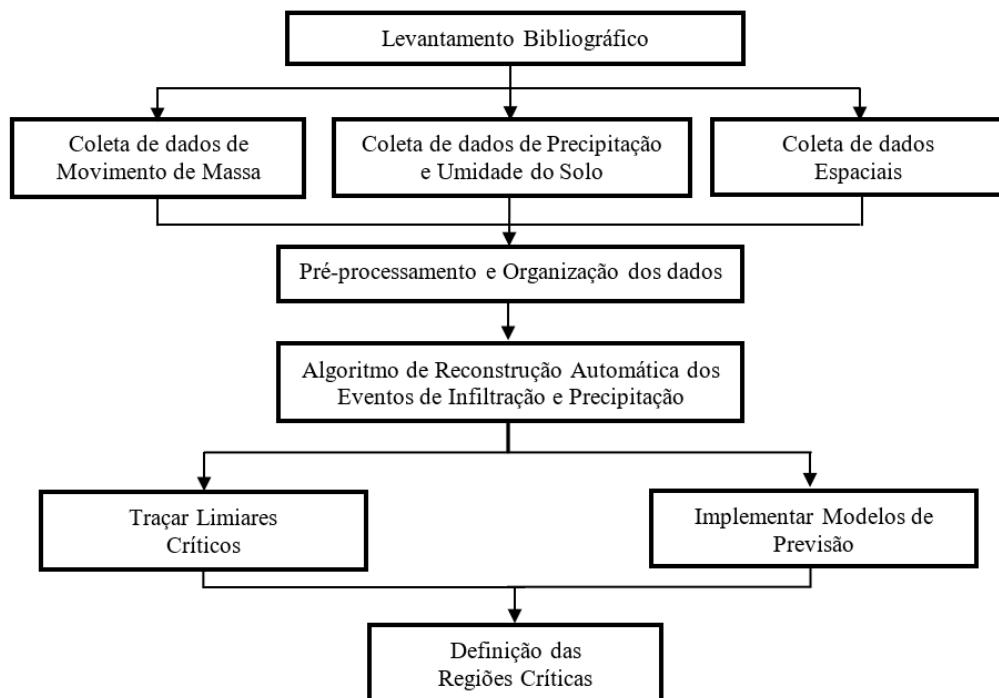


Figura 4.1: Organograma das etapas do trabalho

Fonte: Paróprio Autor

4.1 A Área de estudo

A área de estudo se concentra na região urbana do município de São Bernardo do Campo, localizado ao sudeste da região Metropolitana do estado de São Paulo, correspondente a $409,53 Km^2$ e com aproximadamente 810 mil residentes (IBGE, 2010). A seleção foi motivada pela maior quantidade de mapas disponíveis e estações geotécnicas com dados consistentes, abrangendo diversas áreas de risco e possuindo um longo histórico de ocorrências de deslizamentos de terra.

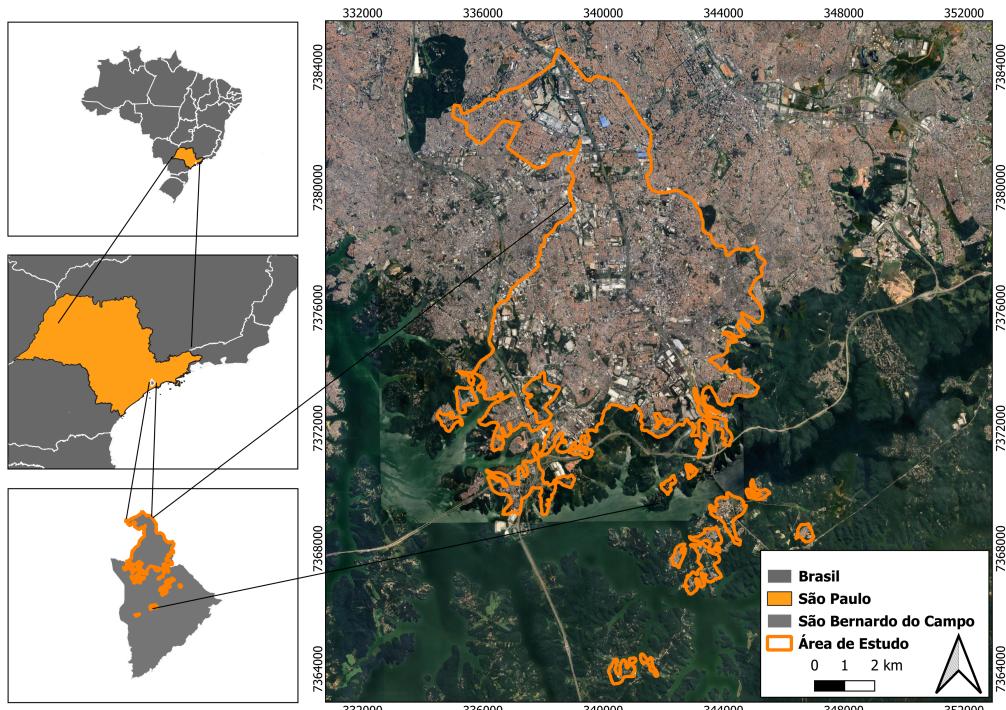


Figura 4.2: Área de estudo.

Fonte: Próprio Autor

4.2 Banco de Dados

Neste trabalho, foram utilizados variáveis estáticas e dinâmicas que influenciam a suscetibilidade a deslizamentos. As variáveis estáticas incluem fatores pedológicos, como tipo de solo; e hidrológicos, como a distância aos rios. Além disso, foram considerados dados de uso e cobertura do solo, limites municipais, rodovias, perigo e suscetibilidade de deslizamento disponibilizados pelo Serviço Geológico do Brasil (CPRM) (INSTITUTO GEOLÓGICO, 2015, 2020), bem como atributos topográficos do TOPODATA (VALERIANO, 2008), como declive, altimetria, curvatura, orientação da curvatura. As variáveis dinâmicas incluem dados de umidade do solo e precipitação, disponibilizados pela instituição de pesquisa Cemaden (Centro Nacional de Monitoramento e Alertas de Desastres Naturais) que podem ser acessa-

dos através do site (CEMADEN, 2022) e registros de ocorrências de deslizamentos do Instituto Geológico (IG) (RAFAELLI, C. B. S. AND FERREIRA, C. J., 2022) e da Defesa Civil Municipal de SBC.

4.2.1 Variáveis Estáticas

Os dados da CPRM são disponibilizados no formato shapefile e organizados em Sistema de Informação Geográfica (SIG), permitindo a manipulação e execução de análises espaciais. A seleção das variáveis pedológica e topológica foi motivada por sua ampla utilização na literatura para análise de suscetibilidade, enquanto as variáveis perigo e suscetibilidade a movimento de massa foram escolhidas para analisar as características dos deslizamentos de terra no município de SBC e avaliar os modelos.

Os diferentes tipos de solo podem apresentar suscetibilidades distintas à erosão, mesmo sob condições semelhantes de declividade, cobertura vegetal e manejo. Isso ocorre porque suas propriedades influenciam de maneira variada os movimentos de massa, oferecendo maior ou menor resistência à infiltração de água e, consequentemente, favorecendo ou inibindo a ocorrência desses eventos (BERTONI e LOMBARDI NETO, 1999), figura 4.3.

O uso e a cobertura do solo podem afetar a instabilidade das encostas devido a ações antrópicas, como cortes de taludes para construção de estradas e edificações em áreas urbanas. A variável foi considerada como estática, uma vez que o curto período analisado neste estudo não resulta em grandes mudanças nesses fatores, figura 4.4.

A distância aos rios é calculada a partir da distância euclidiana do pixel da imagem que representa o deslizamento até o curso d'água mais próximo, onde a proximidade pode aumentar a saturação e a instabilidade das encostas durante chuvas fortes. A distância da rodovia é calculada a partir da distância euclidiana do pixel da imagem que representa o deslizamento até rodovia mais próximo, onde a proximidade pode aumentar a instabilidade devido a cortes de taludes para construção das rodovias.

A CPRM define perigo como o potencial do movimento de massa em causar perda de vidas, ferimentos ou outros impactos na saúde, danos às propriedades, perda de bens e serviços, distúrbios sociais e econômicos e danos ao meio ambiente, figura 4.5. Quanto o mapa de suscetibilidade é definido como a possibilidade de ocorrência de movimento de massa em um determinado local, figura 4.6.

Os dados topográficos utilizados neste estudo foram obtidos a partir do Modelo Digital de Elevação (MDE) do TOPODATA, com resolução espacial de 90 metros. A altimetria representa a altura relação ao nível do mar e pode influenciar as condições

microclimáticas, a cobertura vegetal e os processos de erosão que afetam a estabilidade das encostas, figura 4.7. O declive, por sua vez, está diretamente relacionado à propensão a deslizamentos, uma vez que inclinações mais acentuadas aumentam a influência das forças gravitacionais sobre o solo e as rochas, figura 4.8.

A orientação das vertentes (aspecto) é definida como o ângulo azimutal correspondente à maior inclinação do terreno, no sentido descendente. É expressa em graus, de 0 a 360 graus, dividido em 8 classes. Esta variável está associada ao balanço de radiação (VALERIANO, 2008). As vertentes orientadas para o Equador recebem maior incidência de radiação solar, o que pode influenciar os regimes térmico e hídrico, afetando a evapotranspiração, a umidade do solo e, consequentemente, a estabilidade das encostas, figura 4.9. A Curvatura vertical dividida em 3 classes, reflete as mudanças na inclinação que influenciam a velocidade da água e o potencial de erosão ao longo da inclinação, figura 4.10. A Curvatura horizontal dividida em 3 classes, indicando o potencial de convergência ou divergência de água em encostas, impactando as taxas de erosão, figura 4.11.

Devido as variáveis de estudo serem de diferentes naturezas, quantitativas e qualitativas, foi necessária a codificação das mesmas, tabela 4.1.

Tabela 4.1: Codificação das variáveis estáticas condicionantes a escorregamento de terra.

| Pedologia (P) | Modalidades |
|-----------------------------------|--------------------------------------|
| P1 | Área Urbana |
| P2 | Cambissolos |
| P3 | Gleissolos |
| P4 | Neossolos |
| P5 | Latossolos |
| Uso do Solo (U) | Modalidades |
| U1 | Superfícies Antrópicas Não Agrícolas |
| U2 | Superfície Natural ou Seminatural |
| U3 | Áreas Descobertas |
| U4 | Superfícies Antrópicas Agrícolas |
| Altitude (A) | Modalidades (m) |
| A1 | 65 - 293 |
| A2 | 293 - 521 |
| A3 | 521 - 750 |
| A4 | 750 - 980 |
| Declive (I) | Modalidades (%) |
| I1 | 0 - 3 |
| I2 | 3 - 8 |
| I3 | 8 - 20 |
| I4 | > 20 |
| Orientação da Vertente (O) | Modalidade |
| O1 | Norte |
| O2 | Nordeste |
| O3 | Leste |
| O4 | Sudeste |
| O5 | Sul |
| O6 | Sudoeste |
| O7 | Oeste |
| O8 | Noroeste |
| Curvatura Vertical (CV) | Modalidade |
| CV1 | Côncavo |
| CV2 | Retilíneo |
| CV3 | Convexo |
| Curvatura Horizontal (CH) | Modalidade |
| CH1 | Divergente |
| CH2 | Planar |
| CH3 | Convergente |

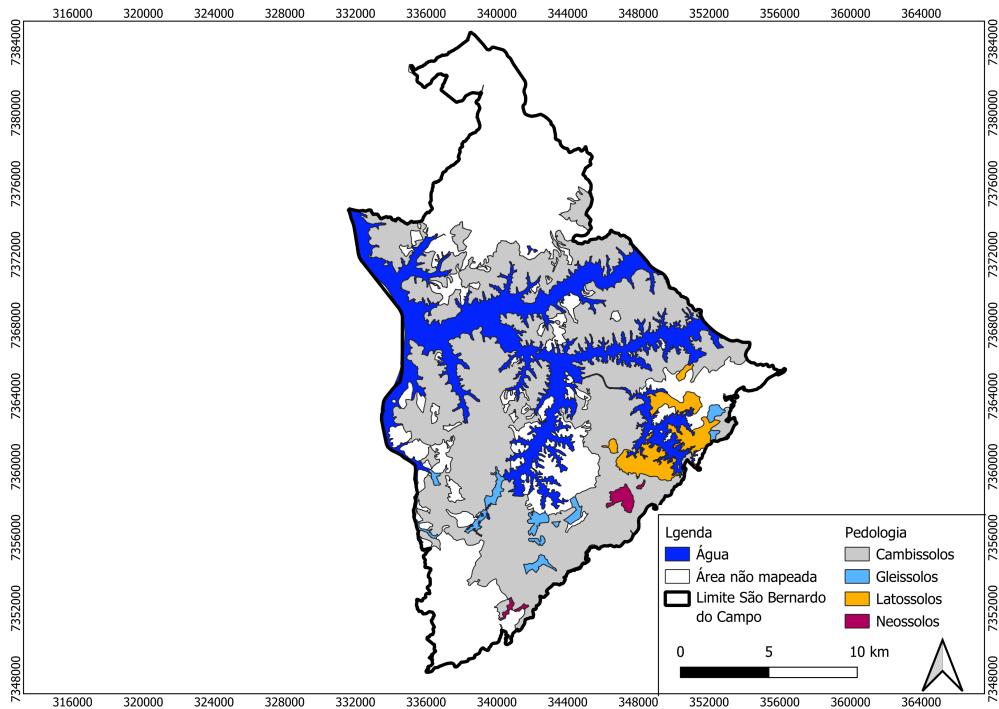


Figura 4.3: Mapa de classes de solo de SBC.

Fonte: adaptado de CPRM

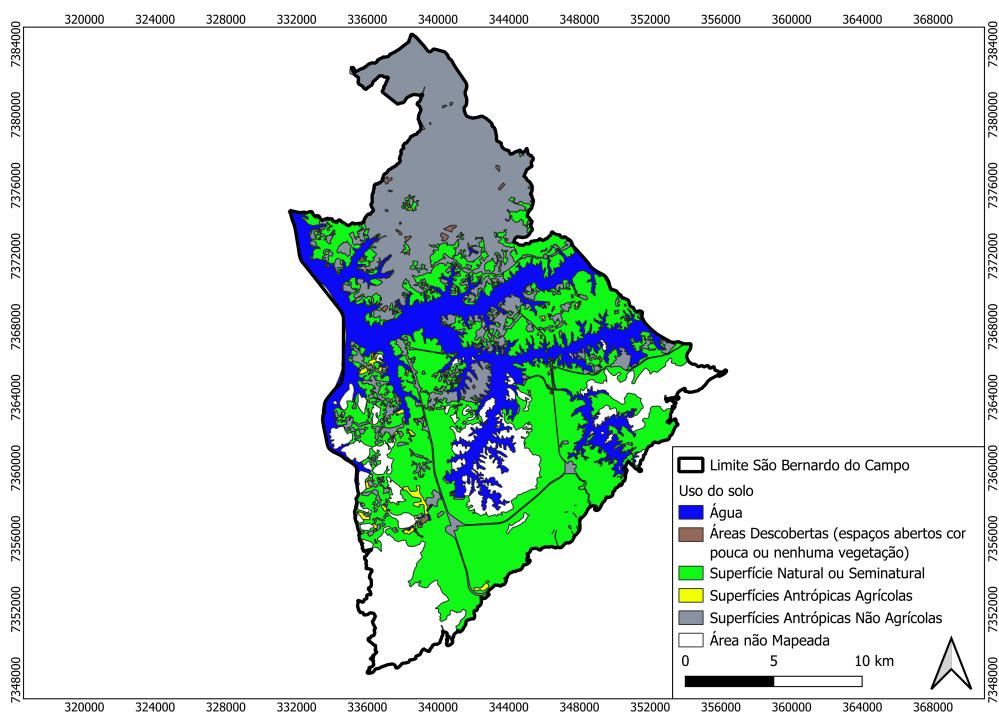


Figura 4.4: Mapa de Uso de Solo de SBC.

Fonte: adaptado de CPRM

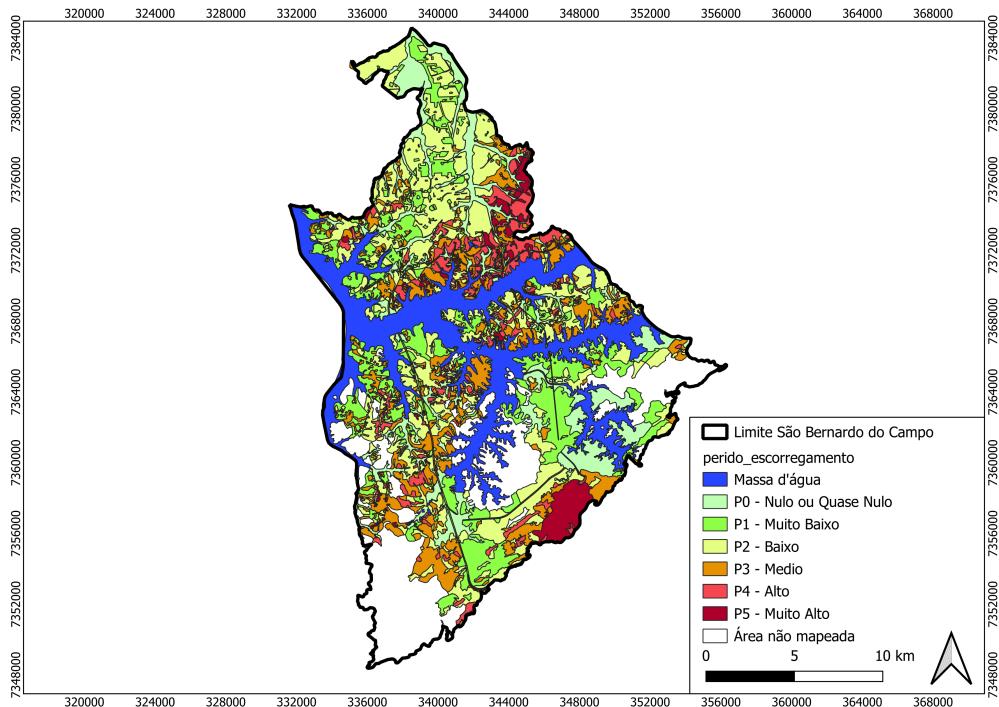


Figura 4.5: Mapa de perigo de deslizamento de SBC.

Fonte: adaptado de CPRM

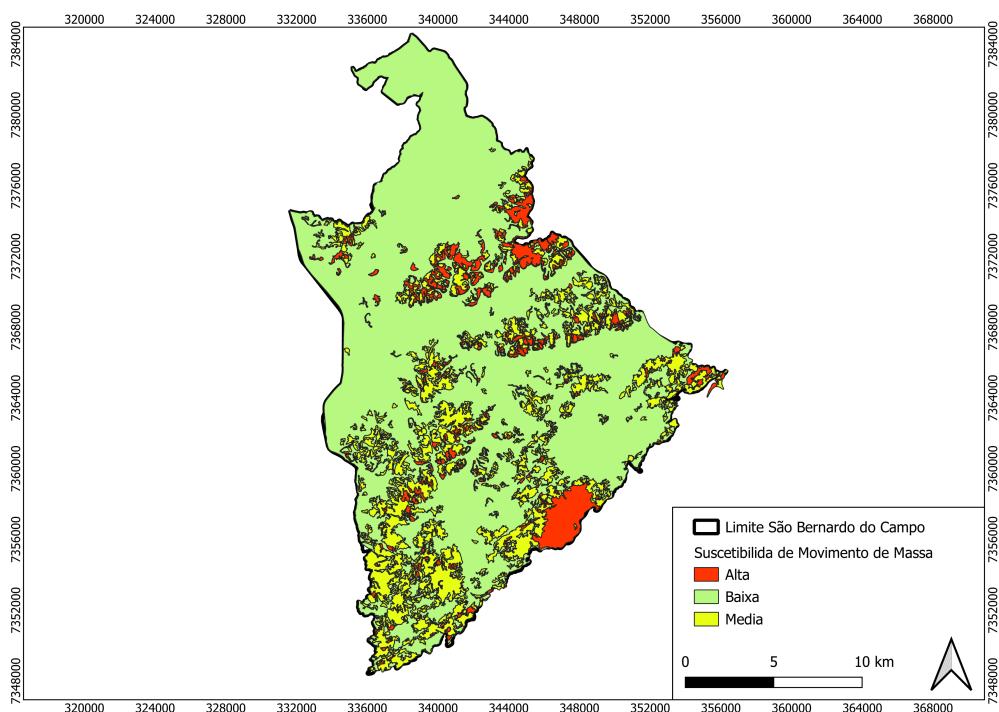


Figura 4.6: Mapa de suscetibilidade de deslizamento de SBC.

Fonte: adaptado de CPRM

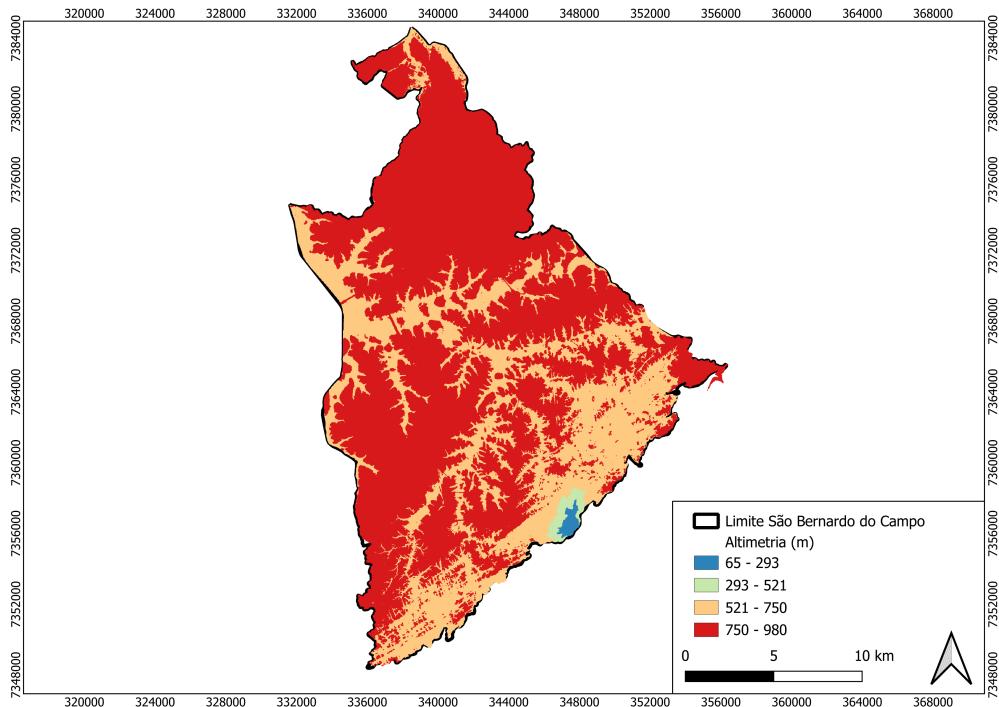


Figura 4.7: Mapa de Altimetria de SBC.

Fonte: adaptado de TOPODATA

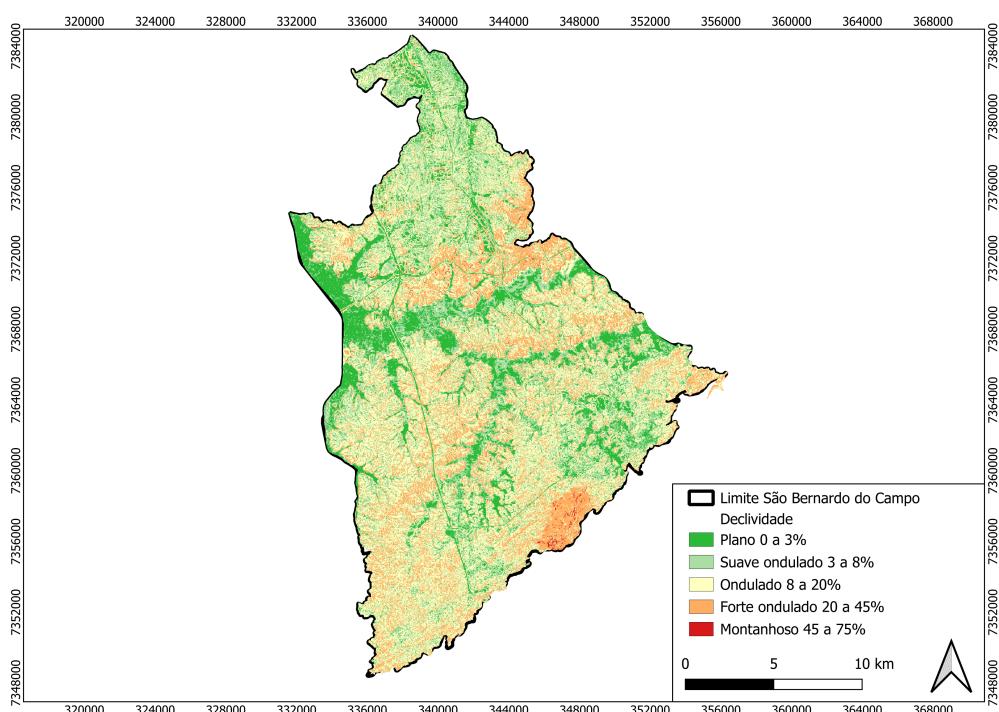


Figura 4.8: Mapa de Declividade de SBC.

Fonte: adaptado de TOPODATA

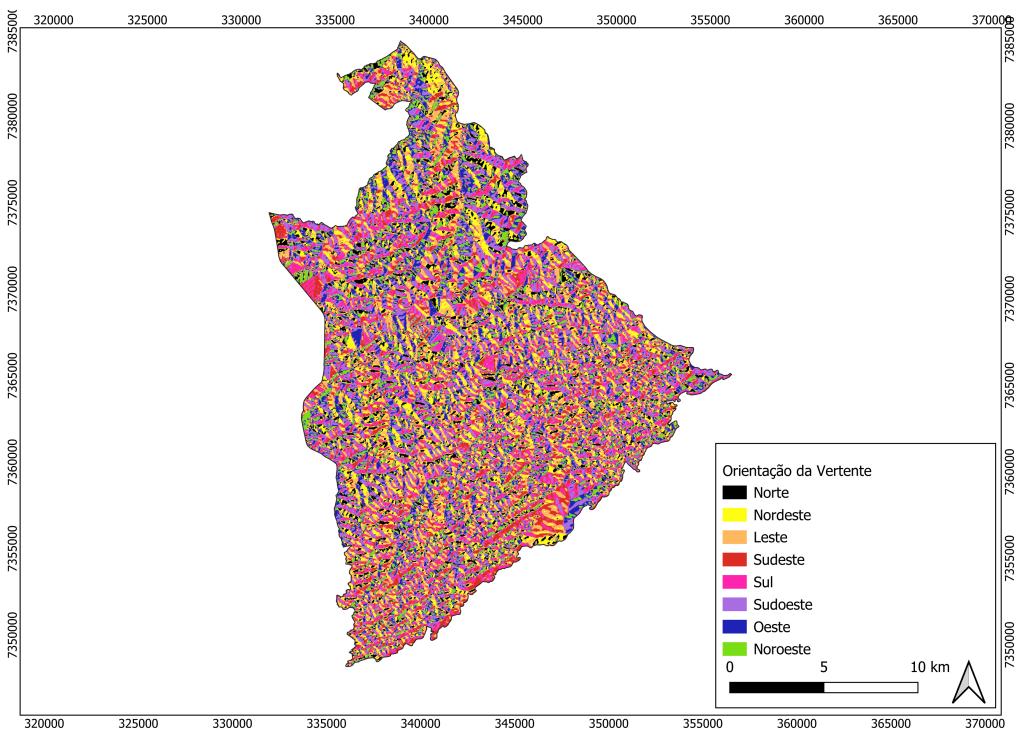


Figura 4.9: Mapa de orientação da vertente de SBC.

Fonte: adaptado de TOPODATA

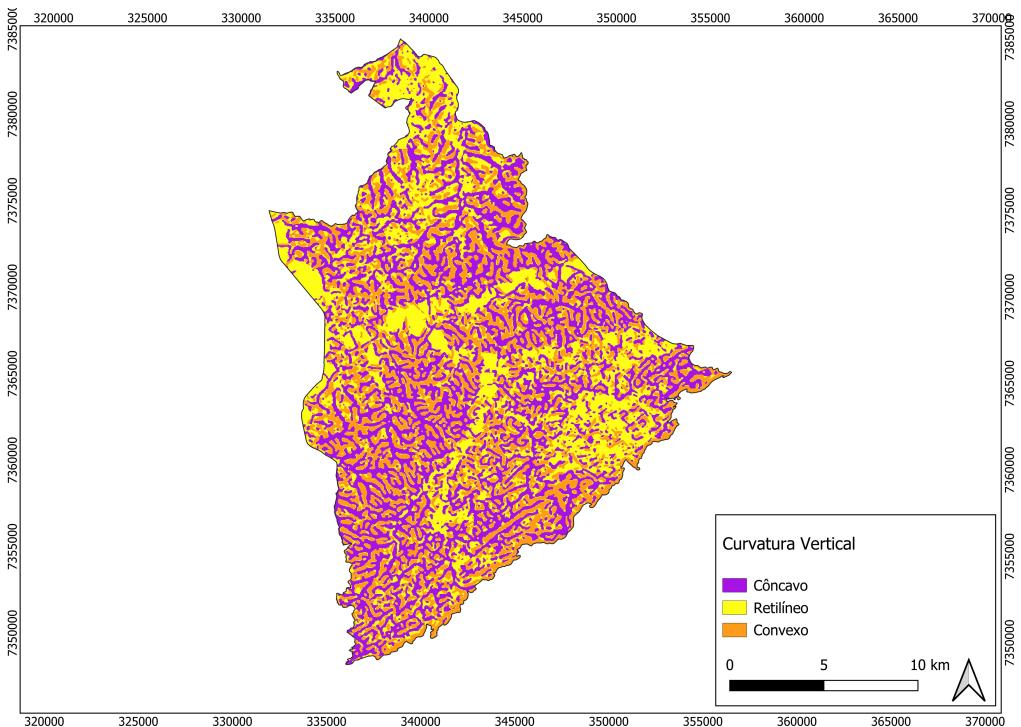


Figura 4.10: Mapa de curvatura vertical de SBC.

Fonte: adaptado de TOPODATA

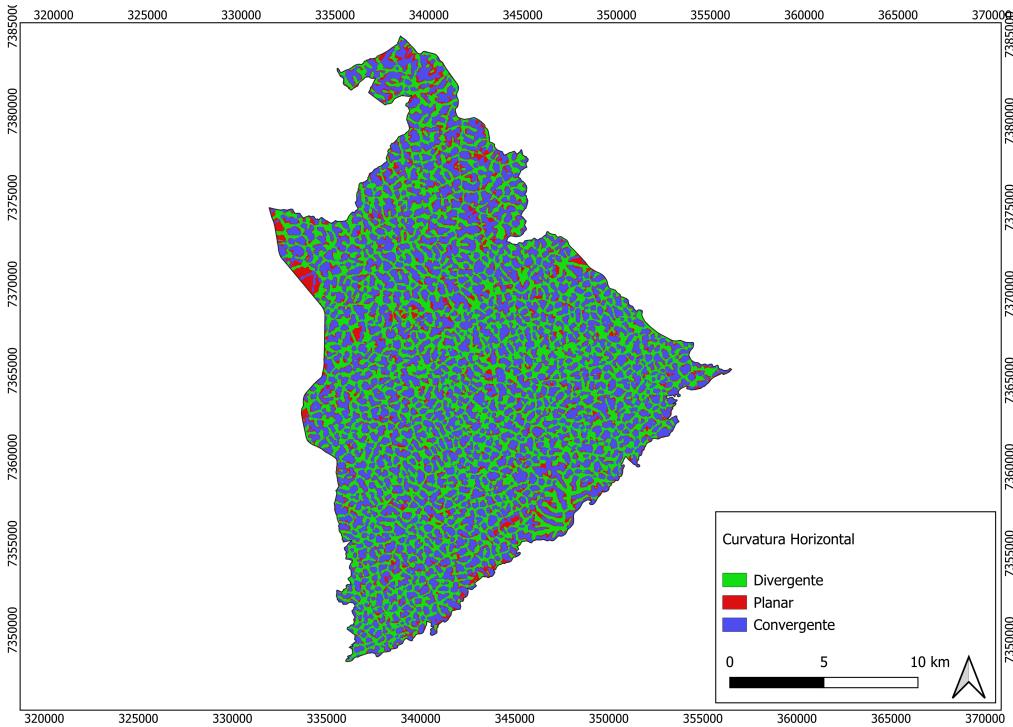


Figura 4.11: Mapa de curvatura horizontal de SBC.

Fonte: adaptado de TOPODATA

4.2.2 Variáveis Dinâmicas

As ocorrências de deslizamentos são filtradas do banco de dados do Instituto Geológico que é constituído de eventos e de desastres do estado de São Paulo, datados de 1967 a 2023. As principais fontes para o levantamento de ocorrências de acidentes e desastres são: Defesa Civil Municipal, Sistema Integrado de Defesa Civil Estadual (SIDEC), Subsecretaria Estadual de Proteção e Defesa Civil (SUPDEC), banco de dados de Desastres Naturais do Instituto de Pesquisas Meteorológicas UNESP (IPMET-UNESP), Sistema Integrado de Informações sobre Desastres S2ID, Secretaria Nacional de Proteção e Defesa Civil (Sedec), notícias de veiculação em mídia eletrônica, relatórios de vistorias e pareceres realizado pelo IG e registros municipais.

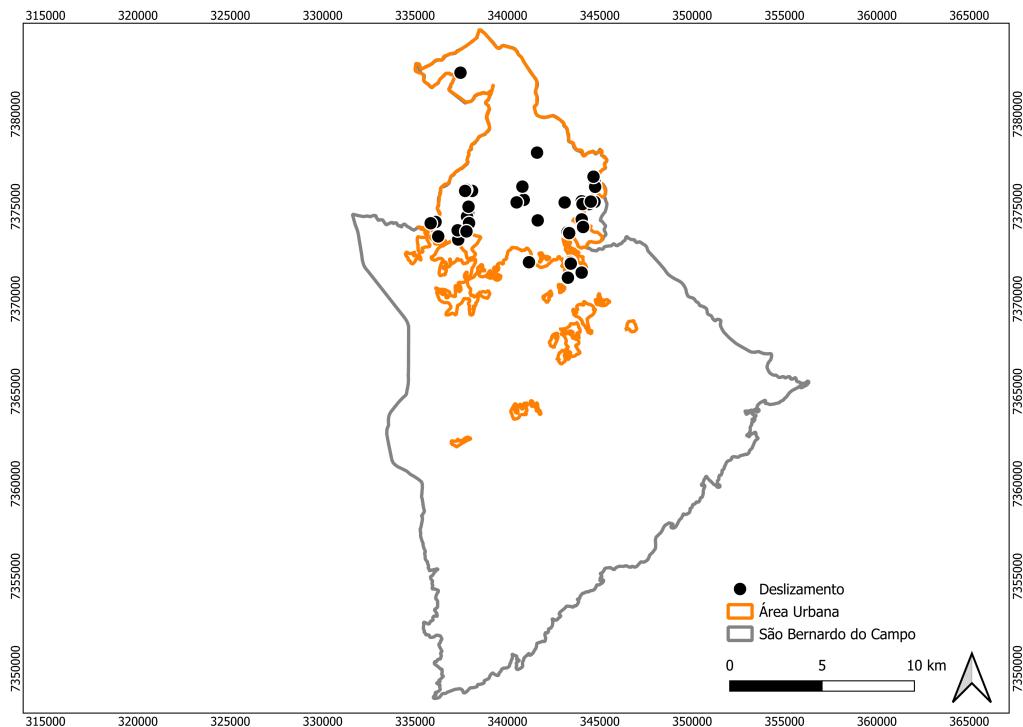


Figura 4.12: Ocorrências relacionadas a deslizamento de terra.

Fonte: Próprio Autor

Os dados de umidade do solo e precipitação são provenientes das estações geotécnicas, situada no estado de São Paulo. As estações utilizam o sensor EnviroSCAN Figura 4.13 para efetuar medições de umidade do solo em seis profundidades distintas, variando de meio metro a três metros, com intervalos de meio metro entre os sensores. O sensor EnviroSCAN emprega a técnica de medição baseada na capacidade do solo, ou seja, a leitura de umidade do solo é realizada de maneira indireta, e os valores obtidos são posteriormente convertidos em umidade volumétrica por meio da curva de calibração. A coleta de dados de precipitação ocorre a cada hora, a menos que o sensor detecte chuva, caso em que os dados são registrados a cada dez minutos. O valor mínimo de precipitação detectado pelo sensor é de 0,2 mm. Quanto aos dados de umidade do solo, estes são coletados a cada dez minutos após a detecção de chuva pelo sensor de precipitação. No entanto, se não houver chuva por um período contínuo de doze horas, o sensor passa a coletar os dados de umidade do solo a cada hora.



Figura 4.13: Sensor de Umidade do Solo EnviroSCAN

Fonte:

<https://sentektechnologies.com/products/soil-data-probes/enviroscan/>

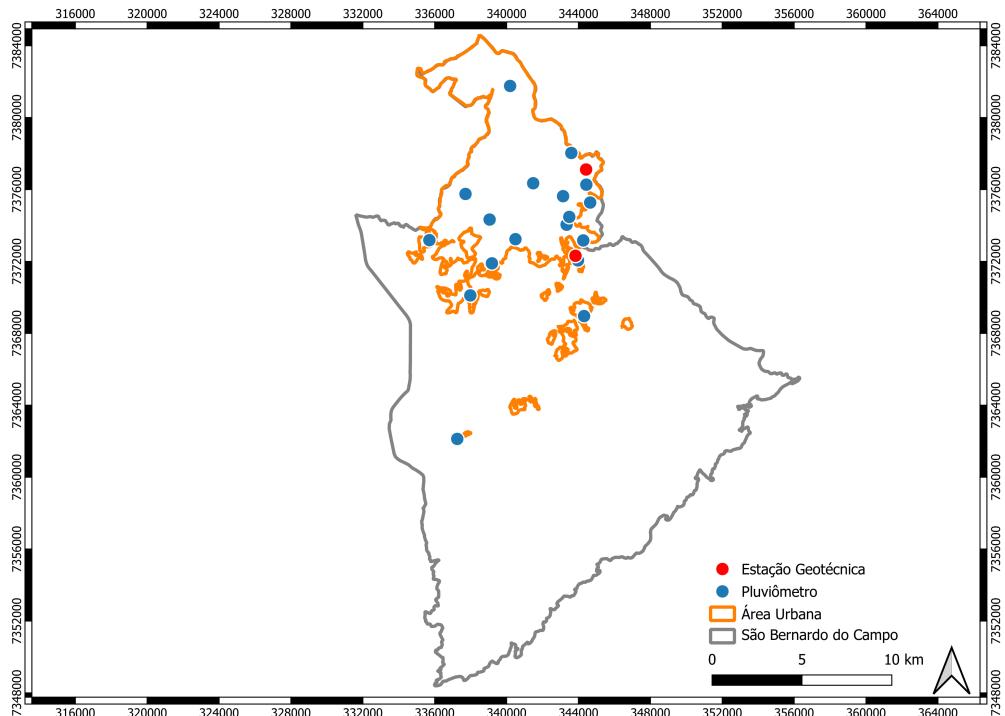


Figura 4.14: Distribuição das estações geotécnicas e pluviométricas.

Fonte: Próprio Autor

Capítulo 5

Método Adotado

Este estudo propõe um algoritmo desenvolvido para reconstruir eventos de infiltração e precipitação que desencadearam deslizamentos de terra. Os eventos reconstruídos são utilizados como entrada para o treinamento de dois modelos: um baseado em limiar crítico e uma regressão logística.

Todo o código desenvolvido neste trabalho foi implementado na linguagem de programação Python 3.11. A escolha dessa linguagem se deve à sua ampla utilização e à extensa disponibilidade de bibliotecas para análise e tratamento de dados. As principais bibliotecas empregadas foram: pandas (versão 2.2.2) e numpy (versão 1.26.4) para manipulação de dados; Scikit-learn (versão 1.6.1) para implementação dos modelos; e matplotlib (versão 3.10.0), plotly (versão 5.24.1) e seaborn (versão 0.13.2) para visualização dos dados. Os mapas foram produzidos no software QGIS 3.38.3.

5.1 O Algoritmo

O algoritmo proposto utiliza os registros de umidade do solo e precipitação para reconstruir eventos de infiltração e precipitação. A infiltração da água no solo consiste na entrada de água no solo (a água passa da superfície do solo para o interior do mesmo em função de um gradiente de potência total da água). Consequentemente, um evento de infiltração é definido como um período de aumento contínuo da umidade do solo, decorrente de processos de infiltração, os quais são separados por um período sem aumento na saturação do solo. Um evento de precipitação, por sua vez, corresponde a um período de chuva contínua ou uma sequência de períodos de precipitação que geram a infiltração de água no solo, separados de eventos anteriores e posteriores por intervalos secos (sem precipitação).

A estrutura lógica é separada em quatro etapas do algoritmo: (1) normalização dos dados de umidade do solo; (2) detecção de eventos de infiltração e precipitação que originou a infiltração; (3) associação das informações sobre deslizamento de terra

(ou ausência destes) aos eventos identificados; (4) calculo da saturação antecedente do solo, mudança de saturação do início ao fim, duração do evento, taxa de infiltração (mudança de saturação dividida pela duração), intensidade média de precipitação e precipitação acumulada para cada evento.

5.1.1 Etapa 1: Normalização dos Dados de Umidade do Solo

Os valores de umidade do solo de cada série temporal foram normalizados pelos valores mínimos e máximos, visando reduzir as incertezas da calibração do sensor. Para mitigar o efeito de possíveis dados anômalos (outliers), o valor para representar o máximo da saturação do solo utilizado corresponde a $\theta_{99,9\%}$, enquanto θ_{Min} corresponde ao solo seco medido em campo,

$$\theta_{norm} = \frac{\theta - \theta_{Min}}{\theta_{99,9\%} - \theta_{Min}} \quad (5.1)$$

onde θ é o medido, θ_{Min} o mínimo, e $\theta_{99,9\%}$ o percentil 99,9 % da umidade do solo da série temporal e θ_{norm} corresponde a saturação S(%). Em seguida, é calculada a média dos três sensores correspondentes às profundidades de 0–50 cm, 50–100 cm e 100–150 cm, representando a saturação média do solo na faixa de 0 a 150 cm de profundidade.

5.1.2 Etapa 2: Detecção de Eventos

O algoritmo começa detectando automaticamente eventos de infiltração isolados com aumento de umidade ao longo de 1 hora. São considerados os pontos com pelo menos 2% de aumento de umidade para omitir ruído de medição dos sensores. Em seguida, eventos com intervalo inferior a 6 horas são agrupados, formando um único evento. As datas do início e fim do evento são registradas. A data de início do evento de precipitação é definida como a primeira data anterior ao início da infiltração quando não há registros de precipitação por 24 horas. A figura 5.1 representa um período da série temporal de saturação da estação Palmira Grassioto, destacando em azul os eventos de infiltração delimitado pelo algoritmo.

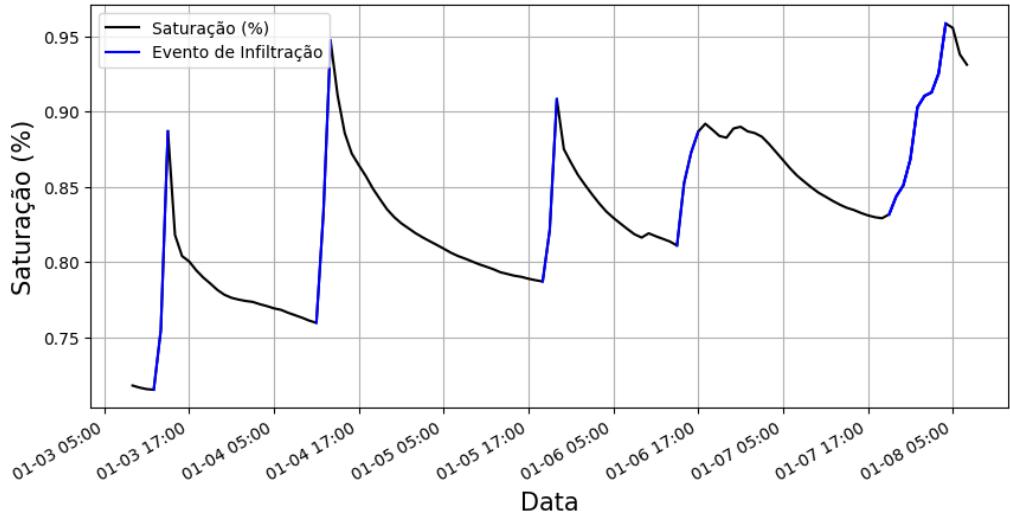


Figura 5.1: Exemplo do algoritmo detectando os eventos.

Fonte: Próprio Autor

5.1.3 Etapa 3: Associação dos Eventos com Deslizamentos de Terra

Cada registro de deslizamento de terra do banco de dados contém sua localização espacial e temporal. O algoritmo utiliza esses registros para associar a ocorrência de deslizamentos a eventos de infiltração e precipitação captados por um único sensor de umidade do solo e pluviométrico. O processo é ilustrado pela figura 5.2. Dado um registro de deslizamento (pin vermelho), o algoritmo calcula uma área circular centrada em sua localização geográfica, com um raio parametrizado (R), que depende de particularidades do local como morfologia e densidade dos sensores. Dentro desse raio estão todos os possíveis sensores (pins azuis) que podem representar o fenômeno. O algoritmo seleciona o sensor com maior proximidade, sem dados faltantes no período da ocorrência do deslizamento.

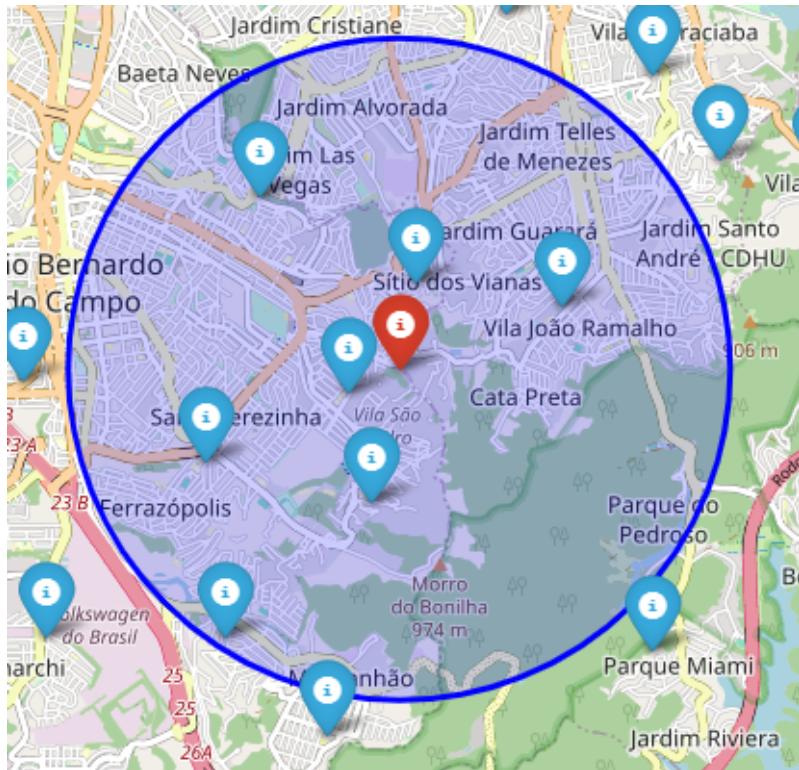


Figura 5.2: Associação dos eventos dentro de um raio de busca pré-estabelecido.

Fonte: Próprio Autor

Um deslizamento pode ocorrer durante um evento de infiltração, desta forma o aumento da umidade do solo após a ocorrência do deslizamento não pode ser considerado, pois não influencia no processo de falha da encosta. Neste caso, a data registrando o final do evento de infiltração e precipitação é alterada para a data de ocorrência do deslizamento. O deslizamento também pode ocorrer após o evento de infiltração, seja devido à movimentação de uma célula de tempestade (WICKI *et al.*, 2020), ou por causa da maneira em que são coletados os dados, conforme é descrito na Seção 6.1. Portanto, a janela temporal para classificação de deslizamento foi estendida para um parâmetro de T horas após o fim de um evento de infiltração.

A Figura 5.3 exibe a série temporal de porcentagem da taxa de saturação do solo, com eventos de infiltração destacados pela linha azul e ocorrências de deslizamentos indicadas pela linha vermelha tracejada. Podemos notar tanto eventos sem deslizamentos quanto eventos associados a múltiplos deslizamentos.

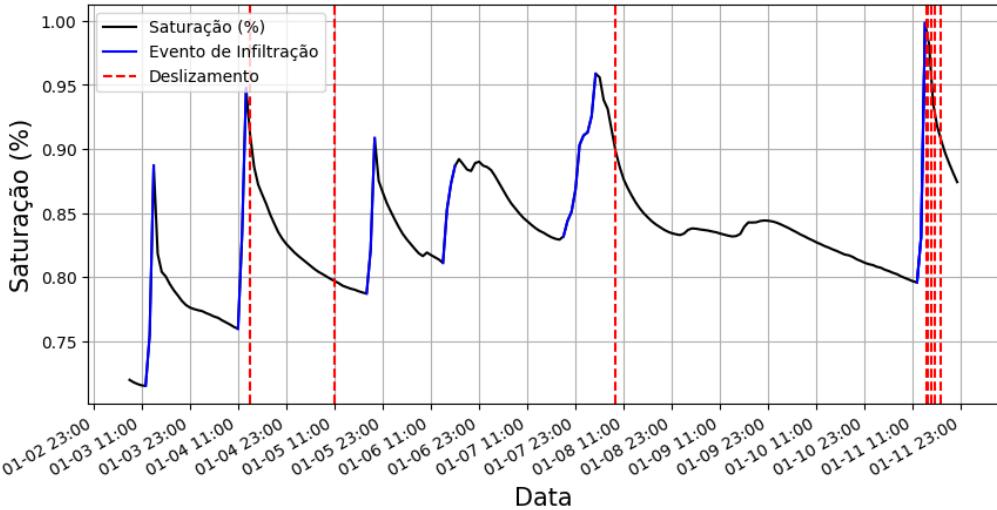


Figura 5.3: Associação dos eventos com os deslizamentos.

Fonte: Próprio Autor

5.1.4 Etapa 4: Cálculo das Métricas

O algoritmo calcula as métricas que caracterizam a instabilidade da encosta para cada evento, independentemente de estarem associados a um deslizamento. As métricas incluem a saturação antecedente do solo (saturação do solo no início de um evento), a mudança da saturação (alteração da saturação do início ao fim do evento), a duração do evento, a taxa de infiltração (mudança da saturação dividida pela duração), a intensidade média de precipitação e a precipitação acumulada. Para representar a perda de água por infiltração e escoamento ao longo do tempo, calculamos a precipitação acumulada, conforme proposto por CROZIER (1999), equação 3.8.

5.2 Modelo 1: Limiar

Embora a abordagem baseada na intensidade e duração da precipitação seja a mais utilizada na literatura, os dados de ocorrências de deslizamento utilizados no presente trabalho são registrados com a data e hora em que a Defesa Civil é acionada, gerando uma lacuna temporal entre o momento real do evento e o registrado, resultando em diversos deslizamentos sem precipitação associada no horário da ocorrência. Essa limitação inviabiliza o uso da abordagem, levando à adoção de limiares baseados na precipitação acumulada e na duração do evento, equação 5.2.

$$E = \alpha D^\beta \quad (5.2)$$

onde, E é a precipitação acumulada do evento (mm), D é a duração do evento de

precipitação (h), α e β são coeficiente positivo que define o intercepto e a inclinação da curva e são otimizados seguindo o método do (BRUNETTI *et al.*, 2010) e em seguida são definidos os níveis de excedência. A saturação final é registrada no momento final do evento reconstruído pelo algoritmo e o nível de excedência é definido como base no percentil dos dados de saturação. Os parâmetros dos limiares são otimizados utilizando dados históricos de ocorrência de deslizamentos correspondentes ao período de julho de 2021 a dezembro de 2022, enquanto o intervalo de janeiro a dezembro de 2023 foi empregado para a avaliação do modelo.

O modelo proposto combina método clássico de limiares de precipitação com limiares de saturação do solo, através de um algoritmo de decisão usando os níveis de excedência definidos. Metodologia desenvolvida com base na hipótese de que adição de informações sobre a umidade do solo reduz o número de falsos alertas, uma vez que o aumento da umidade diminui a resistência ao cisalhamento, devido à redução da coesão aparente dos poros. Para avaliar o desempenho em reduzir os falsos alertas o modelo é comparado com o limiar de precipitação por meio da matriz de confusão.

5.3 Modelo 2: Regressão Logística

A Regressão Logística é uma das técnicas da estatística multivariada mais utilizadas na avaliação de suscetibilidade a deslizamentos de terra (REICHENBACH *et al.*, 2018). Esse método avalia a relação entre uma variável dependente, expressa em duas ou mais classes, e um conjunto de variáveis independentes, que podem ser categóricas ou numéricas. A variável dependente pode assumir natureza dicotômica ou binária, nominal ou ordinal. No caso binário, a variável possui duas classes, geralmente codificadas como 0 ou 1 (por exemplo, ausência ou presença de deslizamentos de terra). Quando há três ou mais classes, a variável dependente pode ser de natureza ordinal, quando existe uma relação de ordem natural entre as classes (por exemplo, níveis de satisfação: "não satisfeito", "satisfeito" e "muito satisfeito") ou natureza nominal quando esta ordem não existe (por exemplo, tipo sanguíneo: A, B, AB ou O).

O modelo de regressão logística pode ser especificado por três componentes: (i) a componente sistemática, que especifica uma função linear entre as variáveis independentes; (ii) a componente aleatória, que identifica a distribuição de probabilidade da variável dependente; e (iii) a função de ligação, que descreve a relação matemática entre a componente sistemática e o valor esperado da componente aleatória (FIGUEIRA, 2006).

Neste trabalho é abordado apenas o caso da regressão logística múltipla com a variável dependente binária. Antes de apresentá-la, é interessante apresentar o

desenvolvimento da regressão logística binária.

O caso da regressão logística binária temos apenas uma variável independente. Digamos então, que a componente sistemática é dada por

$$g(x) = \beta_0 + \beta_1 x \quad (5.3)$$

onde x é a única variável independente e β_0, β_1 são coeficientes.

A componente aleatória Y segue a distribuição de Bernoulli.

$$Y = \begin{cases} 1, & \text{se ocorrer deslizamento} \\ 0, & \text{se não ocorrer deslizamento} \end{cases} \quad (5.4)$$

Tendo uma probabilidade desconhecida p expressa como uma função de x :

$$p(x) = P(Y = 1|x) \quad (5.5)$$

A probabilidade de sucesso é $0 \leq p(x) \leq 1$ e a probabilidade de fracasso é $q = 1 - p(x)$. A função de ligação é obtida igualando a equação *logit* à função $g(x)$, estabelecendo a relação entre a componente sistemática e a componente aleatória, e garantindo que os valores previstos estejam dentro do intervalo $(0,1)$. Assim, a equação da regressão logística pode ser expressa como:

$$\ln\left(\frac{p(x)}{1 - p(x)}\right) = \beta_0 + \beta_1 x \quad (5.6)$$

O objetivo da regressão logística é estimar a probabilidade desconhecida p , em seguida: aplicando o antilogaritmo na equação 5.6 para isolar p , temos:

$$\left(\frac{p(x)}{1 - p(x)}\right) = e^{\beta_0 + \beta_1 x} \quad (5.7)$$

equivalentemente,

$$p(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (5.8)$$

Os parâmetros β_0 e β_1 do modelo são estimados pelo método da máxima verossimilhança. Esse método busca determinar os estimadores dos parâmetros do modelo de regressão logística que maximizam o logaritmo da função máxima verossimilhança a partir do conjunto de dados observados, ou seja, o objetivo é encontrar os valores dos parâmetros que tem maior probabilidade de replicar o padrão de observação presente na amostra.

Seja $\beta = (\beta_0, \beta_1)$ o vetor de coeficientes, e os pares (x_i, y_i) de observação a probabilidade condicional $P(y_i = 1|x_i) = p(x_i)$ e $P(y_i = 0|x_i) = 1 - p(x_i)$. A função de verossimilhança é:

$$L(\beta) = \prod_{i=1}^n p(x_i)^{y_i} [1 - p(x_i)]^{1-y_i} \quad (5.9)$$

Aplicando logaritmo natural em ambos os lados da equação, obtemos a função log-verossimilhança:

$$\ln[L(\beta)] = \sum_{i=1}^n [y_i \ln p(x_i) + (1 - y_i) \ln(1 - p(x_i))]. \quad (5.10)$$

Dada as características não lineares nos parâmetros da equação, a resolução do mesmo requer procedimentos iterativos como o método de Newton-Raphson, o método gera uma sequência de soluções que convergem para o ponto de máximo da função de verossimilhança, para mais informações do funcionamento do método (HASTIE *et al.*, 2009).

Quando temos mais de uma variável independente o modelo é denominado regressão logística múltipla, as equações, os parâmetros e a função de verossimilhança são muito parecidos com o que foi mostrado anteriormente por se tratar de uma generalização da regressão logística binária.

Neste contexto, a componente sistemática é dada por

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_t x_t \quad (5.11)$$

em que $X = (x_1, x_2, \dots, x_t)$ é o conjunto das variáveis independentes e $\beta = (\beta_0, \beta_1, \dots, \beta_t)$ são os coeficientes.

Igualando a equação *logit* à função $g(x)$, obtemos a equação da regressão logística múltipla expressa por

$$\left(\frac{p(x)}{1 - p(x)} \right) = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_t x_t} \quad (5.12)$$

Isolando p , temos:

$$p(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_t x_t)}} \quad (5.13)$$

5.3.1 Método Stepwise

O método de seleção de variáveis stepwise (passo a passo) consiste em um algoritmo iterativo que combinando aspectos da seleção à frente (*forward*) e da seleção

para trás (*backward*), a partir de um critério de seleção estabelecido. Esse procedimento busca identificar o subconjunto de variáveis que maximiza a qualidade do modelo sem incluir preditores irrelevantes.

Na abordagem *forward* (seleção à frente), inicialmente são ajustados modelos univariados para cada variável independente, comparando-os com um modelo contendo apenas o intercepto. A variável independente que apresenta menor valor p é selecionada para o modelo. Em seguida, modelo busca selecionar a segunda variável que produza o maior aumento na razão de verossimilhança quando adicionada ao modelo. O processo continua até que nenhuma variável acrescida no modelo cause aumento significativo na razão de verossimilhança.

Já a abordagem *backward* inicia o procedimento com todas as variáveis independentes no modelo. O teste da razão de verossimilhança é aplicado iterativamente para identificar a variável com menor contribuição estatística, se a variável for significativa, ela fica no modelo e o processo se encerra, caso contrário ela sai do modelo e o processo continua até a variável que menos contribui passar no teste de significância.

O método stepwise combina ambas as abordagens, permitindo não apenas a adição de variáveis relevantes, mas também a reavaliação contínua daquelas já incluídas, garantindo um modelo parcimonioso e estatisticamente robusto.

5.3.2 Amostragem

A aplicação do modelo regressão logística requer um pré-processamento das variáveis de entrada, que são divididas em estáticas e dinâmicas. As variáveis dinâmicas são os dados de saída do algoritmo de reconstrução de eventos, onde cada amostra contém os registros da localização geográfica dos deslizamentos. Nos casos de ausência de deslizamento são registradas as coordenadas geográficas dos sensores responsáveis pelos eventos de precipitação e saturação. As variáveis estáticas são as características espaciais do local. Ambos conjuntos de dados são colocados sobreposto com o mesmo sistema de georreferenciamento e, em seguida, amostradas para cada ponto de deslizamento. Da mesma forma, ocorrem com os pontos sem deslizamento.

Um problema inerente do fenômeno é o desbalanceamento extremos na proporção dos dados (razões de desbalanceamento de 100:1 ou maiores). O uso de todos os pixels sem ocorrência de deslizamento como amostra resultaria em um desequilíbrio severo entre as classes. Esse problema ocorre, principalmente, porque os algoritmos tradicionais assumem diferentes erros como igualmente importantes, supondo que as distribuições são relativamente equilibradas (HE e GARCIA, 2009). Modelos treinados com dados altamente desbalanceados podem apresentar uma acurácia elevada,

mas tendem a tende a prejudicar a identificação de exemplos pertencentes a grupos minoritário, que é o foco deste estudo.

Para lidar com esse problema, adotamos duas das principais abordagens mencionadas por (KRAWCZYK, 2016). A primeira modifica o algoritmo de otimização, enquanto a segunda está relacionada à coleta de dados para a construção da amostra. Utilizamos o parâmetro `class_weight='balanced'` da biblioteca Scikit-learn no modelo de regressão logística, para reduzir o viés em favor dos dados majoritários atribuindo pesos inversamente proporcionais à frequência das classes, para mais detalhes a implementação do método segue a abordagem proposta em (KING e ZENG, 2001). Além disso, testamos duas estratégias distintas para a amostragem dos dados sem ocorrência de deslizamento para avaliar quais variáveis dinâmicas tiveram influência estatisticamente significativa no horário e local onde os deslizamentos ocorreram.

A primeira abordagem consistiu em selecionar aleatoriamente um evento de precipitação e infiltração por dia e, posteriormente, sortear um ponto dentro de um raio de três quilômetros do sensor que registrou o evento para representar as características espaciais. O valor de três quilômetros foi adotado por ser o mesmo parâmetro utilizado na seleção do sensor que representa a ocorrência de deslizamentos. Na segunda abordagem, os pontos foram sorteados aleatoriamente dentro da área de estudo, independentemente da localização dos sensores. Para cada ponto, foi então sorteado a data de evento de precipitação e infiltração registrado pelo sensor mais próximo.

5.3.3 Dimensionamento dos Dados

A padronização dos dados evita viés em relação a qualquer recurso específico devido à sua magnitude. Por exemplo, a variável distância das rodovias é contínua e pode assumir valores superiores a 1000 metros, enquanto a curvatura vertical é categórica e possui apenas três categorias. Para uniformizar as escalas, utilizamos o método *StandardScaler* da biblioteca *scikit-learn*, que transforma os dados para terem média zero e desvio padrão um.

O processo de padronização ocorre em duas etapas. Primeiramente, a média e o desvio padrão de cada variável são calculados com base exclusivamente nos dados de treinamento. Em seguida, esses parâmetros são aplicados para transformar tanto o conjunto de treinamento quanto o de teste, assegurando a consistência na escala das variáveis.

5.4 Métricas de Ajuste

Para avaliar a capacidade de ajuste do modelo e definir um limite para distinguir eventos de infiltração que desencadearam deslizamentos daqueles que não o fazem, foi realizada a análise da Curva Característica Operacional do Receptor (ROC), conforme descrito em (FAWCETT, 2006).

Inicialmente, construiu-se uma matriz de confusão, comparando a classificação observada e a modelada dos eventos. Foram consideradas quatro categorias: verdadeiros positivos (TP, alarmes corretos), falsos positivos (FP, alarmes falsos), verdadeiros negativos (TN, eventos corretamente classificados como não desencadeadores) e falsos negativos (FN, alarmes perdidos).

Os gráficos ROC são gráficos bidimensionais nos quais a taxa de verdadeiros positivos é plotada no eixo Y e a taxa de falsos positivos é plotada no eixo X. Um gráfico ROC descreve compensações relativas entre verdadeiros positivos e falsos positivos (FAWCETT, 2006).

A acurácia mede a proporção de amostras classificadas corretamente na amostra total, com a fórmula:

$$Accuracy = \frac{TP + FN}{TP + FP + TN + FN} \quad (5.14)$$

A precisão é a proporção de amostras positivas corretamente previstas de todas as amostras positivas previstas, medindo a precisão das previsões do modelo para amostras positivas.

$$Precision = \frac{TP}{TP + FP} \quad (5.15)$$

O recall avalia a capacidade do modelo de identificar amostras positivas, representada pela proporção de amostras positivas previstas corretamente de todas as amostras positivas verdadeiras.

$$Recall = \frac{TP}{TP + FN} \quad (5.16)$$

A pontuação F1 é uma métrica composta que considera tanto recall quanto precisão, computada como a média harmônica de ambos. Ela é usada para avaliar de forma abrangente o desempenho do modelo de classificação.

$$F_1 Score = \frac{Precision \cdot Recall}{Precision + Recall} \quad (5.17)$$

Essas métricas permitem avaliar a capacidade do modelo em distinguir corretamente os eventos desencadeadores de deslizamentos, auxiliando na definição de um limiar adequado para sua classificação.

5.5 Pré Processamento e Organização dos Dados

Uma análise preliminar foi realizada para selecionar a área de estudo e as estações a serem utilizadas no presente estudo. No estado de São Paulo, há um total de 41 estações geotécnicas.

De acordo com o fluxograma apresentado na Figura 5.4, 11 estações foram excluídas devido a um grande número de dados ausentes, e outras 15 estações foram descartadas, pois não estão mais transmitindo dados, as demais estações estão aptas para a análise exploratória. A Tabela 5.1 apresenta os nomes das estações eliminadas juntamente com os critérios de eliminação.

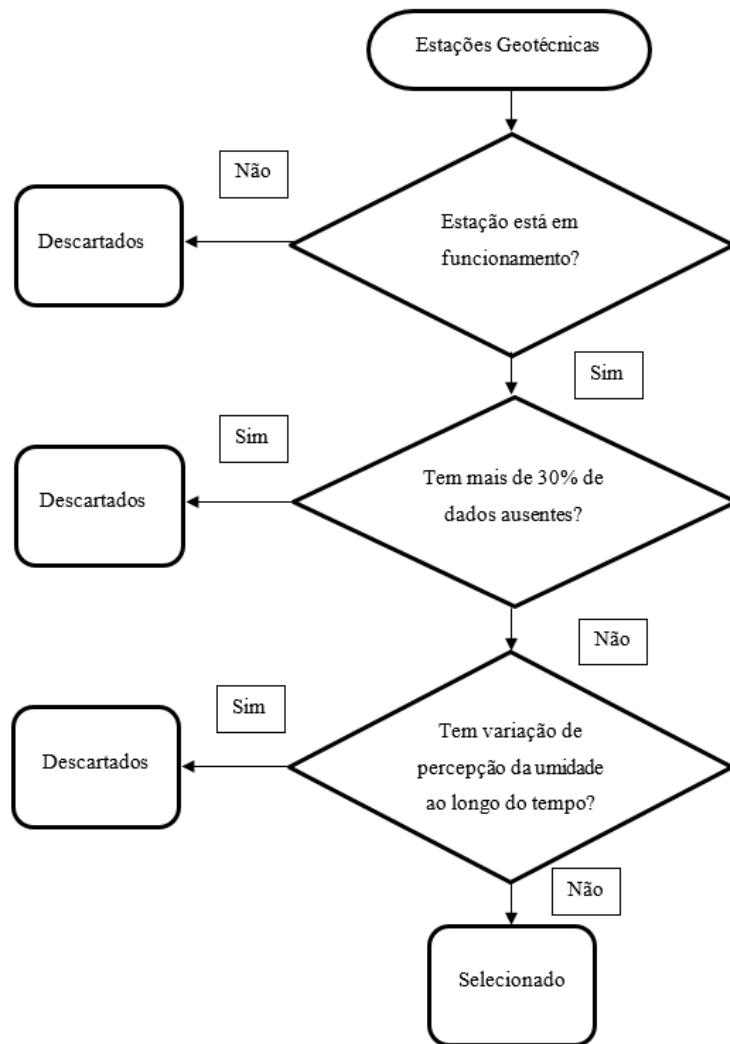


Figura 5.4: Fluxograma com os critérios para selecionar estações aptas ao estudo.

Fonte: Próprio Autor

Tabela 5.1: Estações Geotécnica Descartadas.

| Dados Faltantes | Não Operante |
|-------------------------------|-------------------------------------|
| EE Duarte Caetano - Cota 200 | Cachoeira Paulista (TESTES GEO) |
| Nicho VI - Monte Serrat | Vila Albertina |
| Seminário - Morro Nova Cintra | Pinhal do Miranda |
| HUVET - Morro Marapé | Jardim Alice |
| Sabesp - Morro São Bento | EE Emílio Justo - Vila Progresso |
| Caminho - Monte Serrat | Parque Tecnológico SJC (TESTES GEO) |
| Morro Barbosas | Sabesp - Morro José Menino |
| Morro Itararé | Morro Voturuá |
| Vila Albertina Rua F Centro | Vila Albertina Rua F Leste |
| Elias Costa | Vila Albertina Rua F Oeste |
| Feital | Sítio dos Vianas |
| | Jardim Zaira VIII |
| | Jardim Zaira VI |
| | Jardim Zaira - Deise |
| | Vila São João |

Das 15 estações remanescentes, uma análise individual foi realizada, abrangendo a identificação de outliers. Os valores de umidade do solo que aparecem isolados no gráfico de boxplot representado na Figura 5.5 foram considerados outliers, pois podem ter ocorrido devido a interferências no sensor. Esses valores não condizem com a realidade, uma vez que o processo de secagem do solo normalmente não apresenta mudanças abrupta em um intervalo de uma hora entre as leituras. Portanto, esses valores foram excluídos da análise.

Quanto à precipitação Figura 5.6, apesar de também conter outliers, os valores de 18,6 mm e 17,6 mm registrados em 31/03/2022 e 25/10/2021, respectivamente, são considerados condizentes com a realidade. Isso se deve ao fato de que houve um volume significativo de chuva registrado por outros sensores durante esses períodos. Portanto, esses valores foram mantidos na análise.

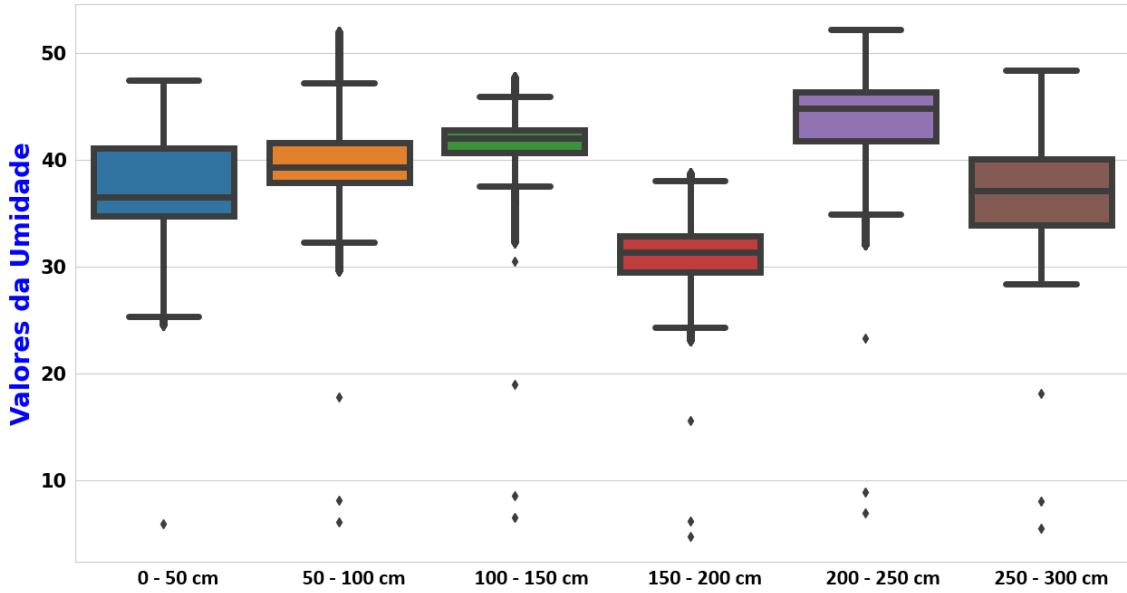


Figura 5.5: Boxplot da umidade do solo nas seis profundidades da estação geotécnica Vila Baiana, situada no município de Guarujá - SP.

Fonte: Próprio Autor

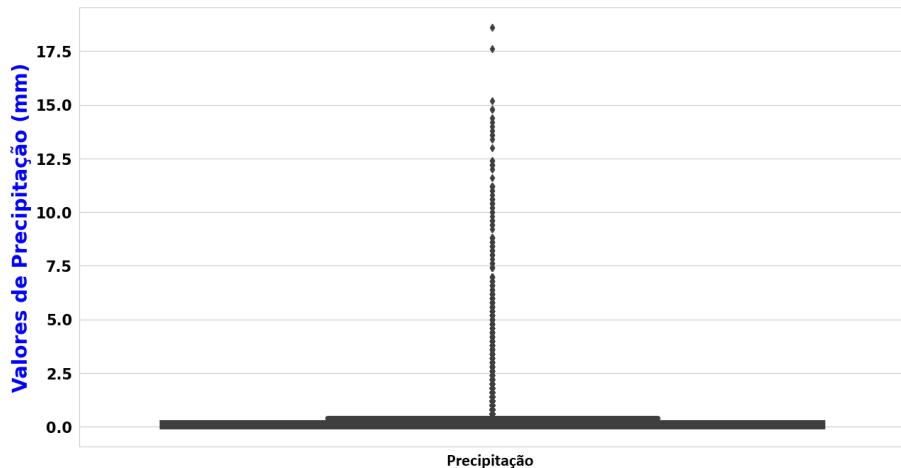


Figura 5.6: Boxplot da Precipitação da estação geotécnica Vila Baiana, situada no município de Guarujá - SP.

Fonte: Próprio Autor

Na ausência de chuva, a coleta de dados ocorreu a cada hora. Para atingir uma resolução temporal de dez minutos nos dados de umidade do solo, foi necessário preencher os intervalos usando o valor medido no instante anterior. Isso se deve ao fato de que, na ausência de chuva e de influências externas, a umidade do solo não apresenta uma variação significativa ao longo de uma hora.

Todas as estações analisadas exibiram uma correlação significativa entre os dados dos sensores de umidade do solo em várias profundidades, como ilustrado na Figura 5.7. Essa observação reflete uma dependência espacial, onde a proximidade entre os

sensores está diretamente associada a uma maior correlação entre eles. Nenhuma estação exibiu correlação entre a umidade do solo e a precipitação. Uma possível explicação para essa falta de correlação reside na natureza dos dados de precipitação, que são coletados em intervalos de dez minutos, resultando em uma predominância de valores zero, já que a maioria das medições ocorre durante períodos sem chuva.

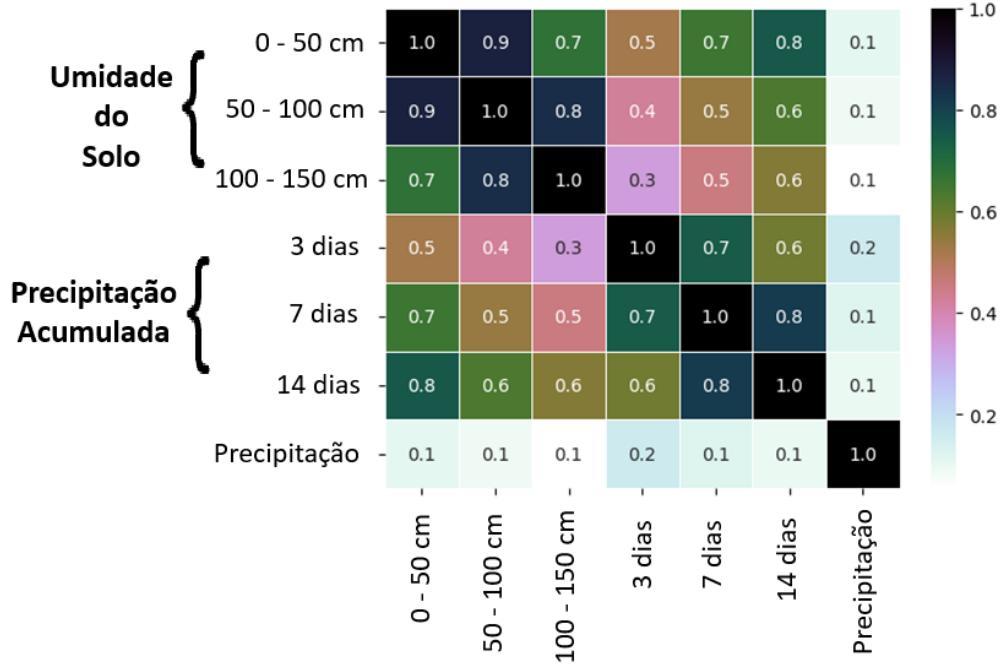


Figura 5.7: Correlação Pearson da estação Palmira Grassioto.

Fonte: Próprio Autor.

As estações Vila Baiana, Vila Albertina Torre Vanguarda, Riviera, Brazilia Tondi de Lima e Xixová Geotec apresentaram correlação fraca entre umidade do solo e a precipitação antecedente, similar ao apresentado na Figura 5.8. Os sensores de umidade do solo nessas estações exibem variações na percepção à umidade e uma tendência decrescente ao longo do tempo. Podemos observar na figura 5.9 que os dados de umidade durante o período de outubro de 2021 a abril de 2022 não reagem da mesma maneira à precipitação quando comparados ao período de outubro de 2022 a abril de 2023. O correto seria apresentar um aumento no valor da umidade.

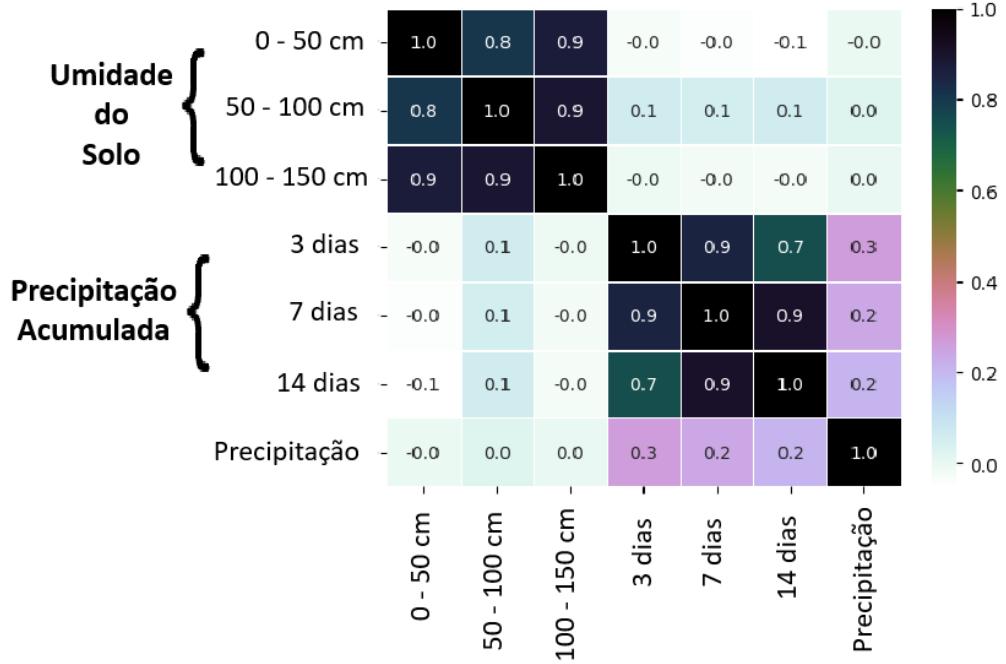


Figura 5.8: Correlação Pearson da estação Riviera.

Fonte: Próprio Autor.

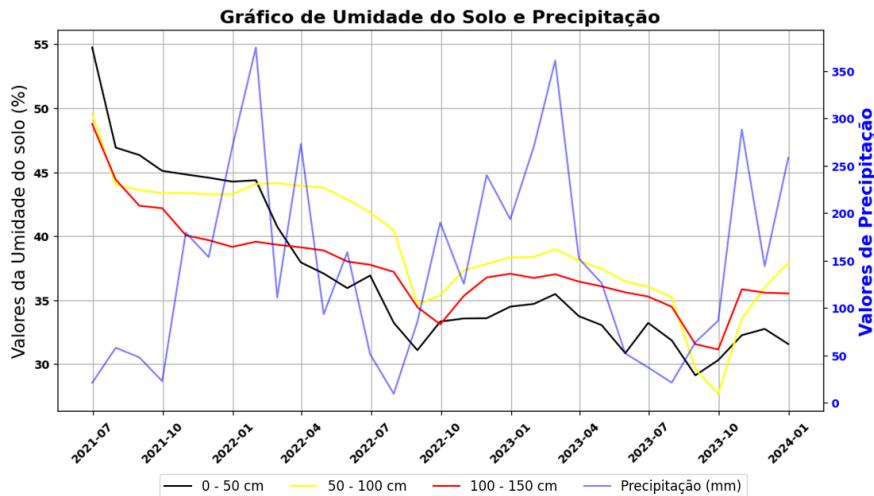


Figura 5.9: Gráfico da umidade do solo e precipitação da estação Riviera.

Fonte: Próprio Autor.

A correlação fraca entre umidade do solo e a precipitação antecedente, aliada à alteração da percepção da umidade do solo e o comportamento decrescente dessa variável ao longo do tempo, indica um mau funcionamento do sensor. Além de fornecer valores não confiável, seria prejudicial na modelagem do fenômeno. Portanto essas estações são descartadas. Das estações aptas listadas na Tabela 5.3, Palmira Grassioto e ETA Rio Grande foram selecionadas para o estudo. Os gráficos da precipitação e de umidade do solo nas profundidades 0 - 50 cm, 50 - 100 cm e 100 - 150

cm dessas estações são representados nas figuras 5.10 e 5.11, respectivamente.

Tabela 5.2: Estações Geotécnica Descartadas.

| Umidade Irreal |
|---------------------------------|
| João Dias |
| ALICE GEOTEC |
| Vila Baiana |
| Vila Albertina Torre Vanguarda |
| Riviera, Brazilia Tondi de Lima |
| Xixová Geotec |
| EGEO Ribeirão Pires |

Tabela 5.3: Estações Aptas.

| |
|------------------------|
| Jd. Frei Orestes |
| Vila Magini |
| Parque Boa Esperança |
| ETA Rio Grande |
| Palmira Grassioto |
| Sabesp - Vila Britânia |
| Parque Gerassi |

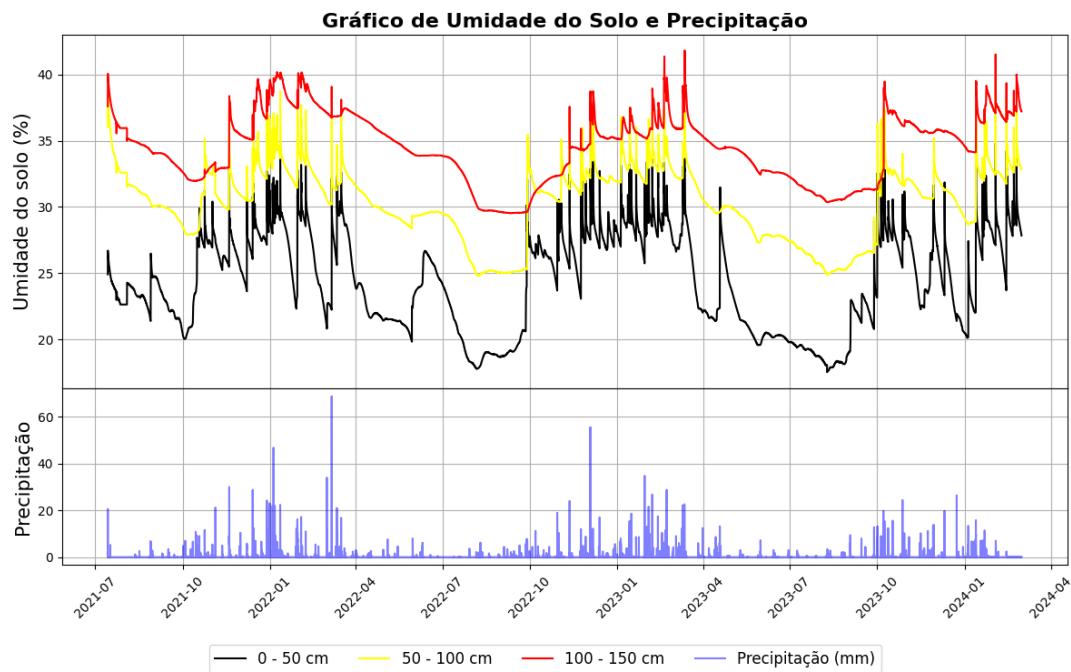


Figura 5.10: Gráfico da precipitação e da umidade do solo nas profundidades 0 - 50 cm, 50 - 100 cm e 100 - 150 cm da estação Palmira Grassioto.

Fonte: Próprio Autor.

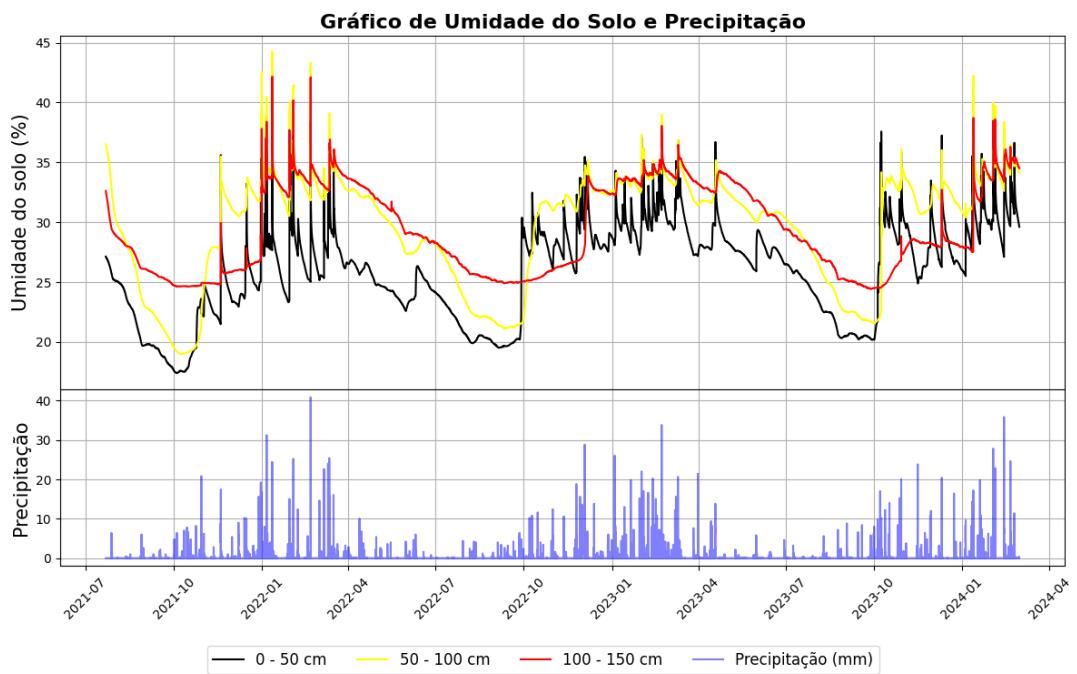


Figura 5.11: Gráfico da precipitação e da umidade do solo nas profundidades 0 - 50 cm, 50 - 100 cm e 100 - 150 cm da estação ETA Rio Grande.

Fonte: Próprio Autor.

O banco de dados de desastres naturais (RAFAELLI, C. B. S. AND FERREIRA, C. J., 2022) contém um total de 7165 registros de ocorrências relacionadas a deslizamento de terra. Destes, 2723 ocorreram no período de 1993 a 2023 em São Bernardo do Campo. Foi selecionado o período que contém dados de umidade do solo disponível, julho de 2021 a dezembro de 2023. Na primeira etapa da análise, foram removidos do banco os registros de vistorias técnicas realizadas pela defesa civil para avaliação de risco de deslizamento, assim como registros incompletos e outras ocorrências que não correspondem a deflagração de deslizamentos em si. Na segunda etapa, foram excluídas as ocorrências de deslizamentos induzidos por outros fatores, como vazamentos nos sistemas de água ou esgoto e execução de obras. Resultando em 50 ocorrências ao longo do período amostral de 2021 a 2023. Após o pré-processamento e a seleção das ocorrências de deslizamentos, os dados estão prontos para serem utilizados como entrada no algoritmo de reconstrução de eventos de precipitação e infiltração.

Capítulo 6

Resultados e Discussão

6.1 Análise de Incerteza Associada aos Dados

A modelagem desenvolvida integra diversos conjuntos de dados de fontes distintas, os quais podem apresentar disparidades em termos de formato, escala e unidade. Essas disparidades podem introduzir incertezas que têm o potencial de influenciar os resultados do modelo. Essa seção apresenta uma análise qualitativa das possíveis fontes de incerteza associadas aos dados.

Os registros de ocorrências da Defesa Civil são marcados com a data e hora em que o órgão é acionado. No entanto, existe uma lacuna temporal entre o momento exato do evento e o acionamento da Defesa Civil, que pode variar de alguns minutos a horas. Essa defasagem pode resultar em registros com valores de precipitação inferiores aos ocorridos durante o evento real ou em dias em que não houve chuva registrada. Além disso, os dados provenientes de notícias de mídia eletrônica geralmente não incluem o horário do evento, o que também pode introduzir incertezas. O número de movimentos de massa que ocorreram no período da pesquisa, provavelmente é maior do que os cadastrados. Isto se deve ao fato de que os agentes da defesa civil são chamados pela população, para atender as áreas habitadas, eventos que não têm impacto social significativo podem não ser registrados, já que a Defesa Civil pode não ser acionada e as ocorrências podem não ser divulgadas na mídia.

Alguns registros de ocorrências foram excluídos devido à falta de informações necessárias para classificar o tipo de movimento de massa. A maioria dos registros de deslizamento de terra não inclui informações sobre a velocidade ou o plano de ruptura, o que dificulta a classificação da tipologia do movimento. Os eventos com dados insuficientes para serem classificados como deslizamento de terra foram desconsiderados.

Para os dados de umidade do solo, não foram fornecidos detalhes sobre o desenvolvimento da curva de calibração (curva que relaciona a umidade com a variável

lida pelo sensor - capacidade). Embora o fabricante desses sensores informe que o aparelho sensor tem acurácia, a verdade é que cada camada do solo é diferente em termos de textura, porosidade, mineralogia, densidade e etc. Isto demanda que, em cada ponto de monitoramento de umidade, seja necessária a construção de curvas de calibração para diferentes profundidades.

A confiabilidade das medições de umidade do solo não pode ser definitivamente estabelecida, pois não sabemos se foi construída uma curva específica derivada de diferentes tipos de solo, ou a exatidão da correspondência entre as leituras do sensor e medições diretas de umidade do solo no campo. No entanto, a ausência dessas informações não afeta substancialmente o desenvolvimento deste trabalho, uma vez que o objetivo principal é desenvolver um limiar crítico, e quaisquer imprecisões nas leituras dos sensores seriam supostamente sistemáticas.

6.2 Algoritmo

Os parâmetros do algoritmo dependem das características da região e da densidade de sensores. Os valores escolhidos são apresentados na tabela 6.1. EI e EP representam a variação mínima detectável pelos sensores de umidade do solo e dos pluviômetros, respectivamente. Valores inferiores ao mínimo detectável são tratados como interferências. II e TP representam o tempo sem alteração na umidade ou detecção de precipitação, utilizados para separar os eventos. RI e RP definem o raio de busca centrado no local de ocorrência do deslizamento para seleção do sensor que o representa, no caso o sensor mais próximo, sem dados faltantes. A constante $k = 0.84$ é uma constante que funciona como um fator de decaimento, por exemplo, a contribuição da precipitação medida quatro dias antes do deslizamento é reduzida para aproximadamente metade (49,8%). T representa o intervalo máximo de tempo após o término de um evento no qual um deslizamento pode ser associado a esse evento. Esse parâmetro depende principalmente da qualidade temporal em que as ocorrências são registradas, devido a possíveis atrasos relatados na seção anterior, foi adotado $T = 24$ horas.

Tabela 6.1: Parâmetros do algoritmo.

| Sensor | Siglas | Parâmetros | Valores |
|-----------------|--------|-------------------------|----------|
| Umidade do solo | EI | Eventos isolados | 2% |
| | II | Intervalo entre eventos | 6 horas |
| | RI | Raio de busca | 50 km |
| Pluviométrico | EP | Eventos isolados | 0.2 mm |
| | TP | Tempo sem precipitação | 48 horas |
| | RP | Raio de busca | 3 km |
| | K | Fator de decaimento | 0.84 |
| | T | Janela temporal | 24 horas |

O raio de busca é um dos parâmetros mais importante para o algoritmo, pois a escolha de um valor baixo resulta em uma área em torno do deslizamento sem nenhum sensor e um valor grande pode extrair dados de uma estação que não represente adequadamente o local do deslizamento devido a variação espacial da precipitação e umidade do solo. O ideal é obter o sensor o mais próximo possível do evento com dados consistentes, para a rede pluviométrica composta por 20 sensores foram testados valores de 1 a 4 km para RP, em contraste, baixa densidade de sensores de umidade do solo (apenas 2), não foi possível testar raio de buscas diferentes, resultando em um raio de busca maior que dos sensores pluviométricos, RI = 50 km.

O algoritmo identifica e relata os eventos sem sucesso na reconstrução, especificando as razões conforme detalhado na tabela 6.2. À medida que o raio de busca (RP) aumenta, podemos notar uma redução no número de ocorrências sem sensores na área designada e um aumento no número de eventos reconstruídos. O número de ocorrências sem detecção de precipitação ou alterações na umidade do solo se manteve praticamente constante, possivelmente esses casos ocorreram devido a fatores externos, como vazamentos em sistemas de água ou esgoto umedecendo o solo em locais fora da área de atuação do sensor, obras na região alterando a paisagem natural e a tornando mais suscetível, ou atrasos temporais entre a ocorrência e o registro do deslizamento. O valor escolhido foi RP = 3 km, pois valores maiores não resultaram em um aumento significativo no número de eventos reconstruídos, e vale notar que sensores muito distantes do ponto de deslizamento podem não refletir com precisão as condições pluviométricas locais. O algoritmo reconstruiu um total de 1.048 eventos de precipitação e 131 eventos de infiltração, das 50 ocorrências de deslizamento, reconstruiu 42, correspondendo a 84% dos casos. A discrepância entre o número de eventos reconstruídos é devido a densidade de sensores na região. Em um dia de chuva generalizada no município, o algoritmo tem 20 pluviômetros disponíveis para determinar os eventos de precipitação, mas apenas 2 sensores de

umidade do solo para os eventos de infiltração.

Tabela 6.2: Diferentes raio de entrada para o algoritmo.

| RP | 1 km | 2 km | 3 km | 4 km | 10 km |
|---|----------|----------|----------|----------|----------|
| Sem sensor dentro do raio de busca | 23 | 6 | 2 | 0 | 0 |
| Sem alteração na precipitação e umidade do solo | 4 | 5 | 5 | 5 | 5 |
| Sem alteração na precipitação | 5 | 6 | 1 | 2 | 0 |
| Eventos reconstruídos | 18 (36%) | 34 (68%) | 42 (84%) | 43 (86%) | 45 (90%) |

Para estudar a capacidade do algoritmo de reconstruir eventos de infiltração e precipitação, os eventos que geraram deslizamentos foram calculados de forma manual pelo especialista. Das 50 ocorrências de deslizamento o método manual obteve sucesso em reconstruir os mesmos casos do algoritmo, e o mesmo problema para dos demais, sendo 5 casos sem detecção de precipitação e alteração na umidade do solo e 2 casos com sensores muito distantes para representar o deslizamento. Em 6 casos (14,2%) o método do algoritmo definiu eventos de infiltração mais curto que o método manual, e em 15 (35%) casos o algoritmo definiu eventos de precipitação mais longo que o manual, essas diferenças são devidas as interpretações subjetivas do pesquisador em definir o momento que inicia o evento, a figura 6.1 ilustra um exemplo da diferença do método algoritmo para o manual em um evento de infiltração. Esse resultado é semelhante ao algoritmo para reconstruir eventos de precipitação obtidos por (BERTI *et al.*, 2012; MELILLO *et al.*, 2015), com 34% e 12,3% dos eventos de precipitação definidos manualmente sistematicamente mais curtos.

A maior incerteza na reconstrução dos eventos está na identificação do momento em que um evento se inicia. Essa incerteza ocorre, sobretudo, em eventos de longa duração com precipitações espaçadas em curtos intervalos de tempo, como ilustrado na figura 3.7. Uma diferencial do algoritmo em relação ao anteriores é a incorporação do método proposto por (CROZIER, 1999) reduzindo a diferença entre a precipitação acumulada reconstruída pelo algoritmo e a manualmente. O método adiciona um fator empírico de decaimento $k = 0.84$, representando a perda de água por evaporação, infiltração e escoamento ao longo do tempo, precipitações distantes do final do evento sofrem maior penalidade e exercem menor influencia na precipitação acumulada.

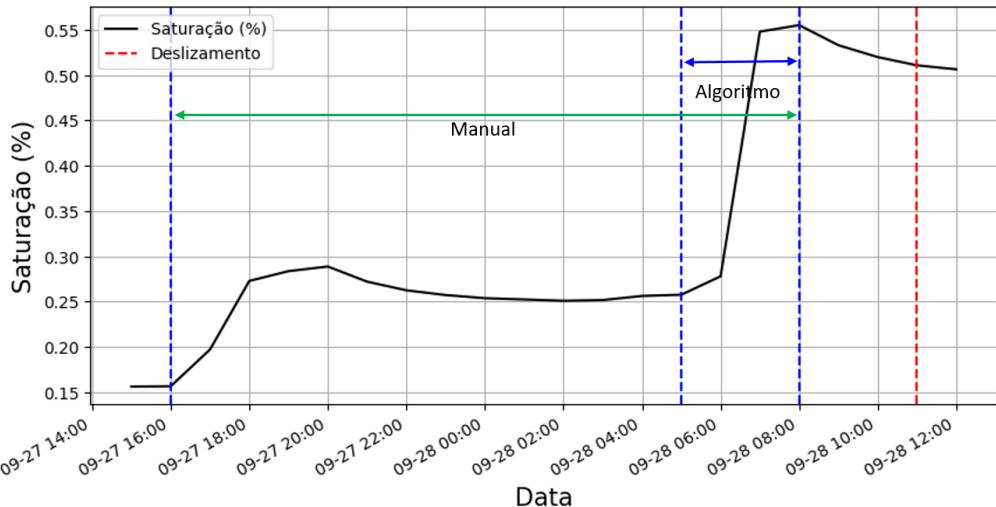


Figura 6.1: Comparação do evento de infiltração calculado pelo algoritmo e manual.

Fonte: Próprio Autor

O algoritmo reconstrói os eventos utilizando os parâmetros (II e TP) definidos pelo pesquisador que representam o tempo sem alteração na umidade ou detecção de precipitação para separação entre eventos sucessivos. Essas informações são as mesmas usadas por um pesquisador para separar eventos manualmente. O problema é que geralmente os critérios usados no método manual não são claros ou explícitos (ALEOTTI, 2004; BRUNETTI *et al.*, 2010; GUZZETTI *et al.*, 2007; REICHENBACH *et al.*, 1998; SANTORO *et al.*, 2010; TATIZANA *et al.*, 1987a), ou não são aplicados consistentemente. A falta de clareza nos parâmetros dificulta a reprodutibilidade da pesquisa e dificulta a comparação entre limiares diferentes. O algoritmo realiza uma reconstrução objetiva e reproduzível dos eventos garantindo a consistência metodológica. Outro aspecto do algoritmo é que os eventos são definidos de forma independente das informações de deslizamento, tais informações são associadas a um evento somente após o evento ter sido identificado. Uma vantagem significativa sobre os métodos tradicionais que consideram apenas eventos de precipitação que resultaram em deslizamento (BERTI *et al.*, 2012; BRUNETTI *et al.*, 2010; GUZZETTI *et al.*, 2007; LUMB, 1975; PERUCCACCI *et al.*, 2012; SANTORO *et al.*, 2010; TATIZANA *et al.*, 1987a) e ignoram os outros eventos, que representam maior parte dos dados.

Outra vantagem de utilizar o algoritmo é a redução do tempo para determinar os eventos e alterar os parâmetros de forma rápida, MELILLO *et al.* (2015) estimaram em cerca de 1 mês, para reconstruir os eventos de precipitação de 163 ocorrências em um período de 9 anos e verificar possíveis erros, indicando a dificuldade para um pesquisador para analisar conjuntos de dados extensos com vários sensores. Para o conjunto de dados usado nesse estudo com 50 ocorrências, o tempo de execução do código demora menos de 10 minutos.

Além dos limiares de precipitação os eventos calculados pelo algoritmo podem ser usados em diversas abordagens, como modelos estatísticos para avaliar suscetibilidade, probabilidade condicional e inferência bayesiana para analisar a possibilidade de ocorrência do deslizamento, ou analisar processos deferentes como tempestades de alta intensidade e inundações repentinas.

6.3 Análise Exploratória de dados

Após a aquisição e organização dos dados, foram realizadas análises para compreendê-los melhor. A Figura 6.2 apresenta o número de ocorrências de deslizamentos e a média da precipitação distribuídas mensalmente. Observa-se que, de maneira geral, a quantidade de ocorrências de deslizamento acompanha a precipitação média mensal. Das 42 ocorrências avaliadas, 85,7% ocorreram nos meses de dezembro a março, que correspondem a 56% da precipitação anual. O mês de janeiro registrou o maior número de ocorrências, com 19 registros, representando cerca de 45%.

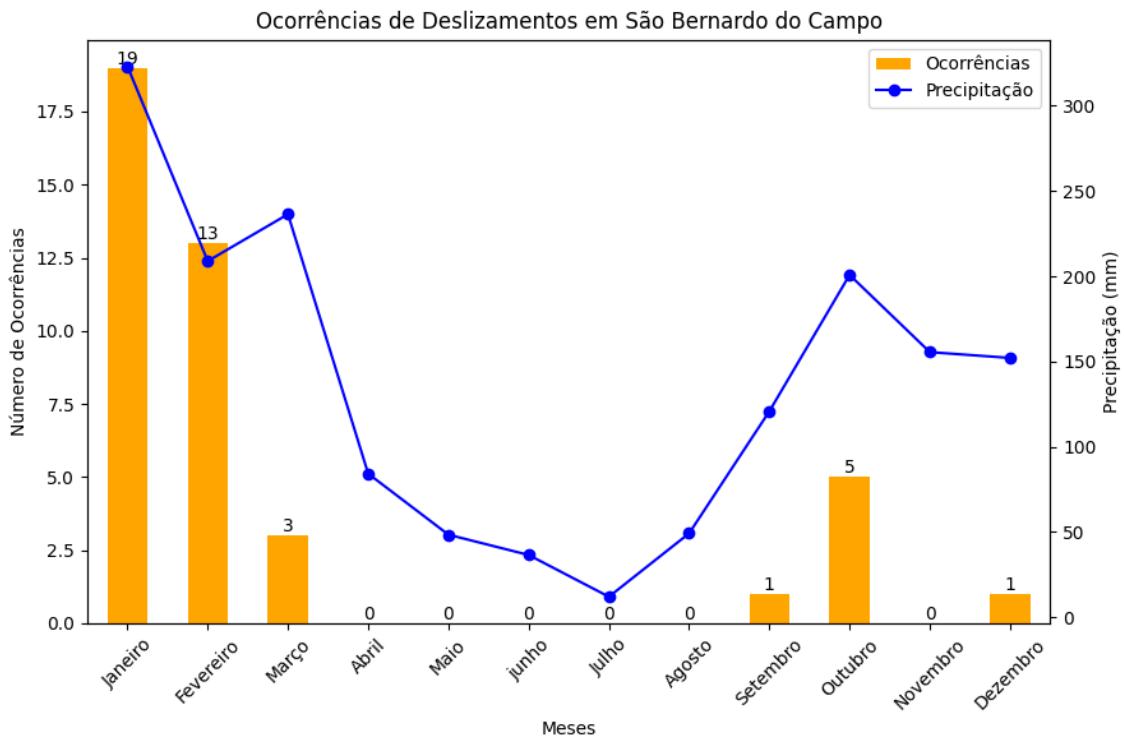


Figura 6.2: Distribuição de deslizamento em função das ocorrências em São Bernardo do Campo, SP

Fonte: Próprio Autor

A análise dos 42 eventos de precipitação que resultaram em deslizamentos reconstruído pelo algoritmo revela que o intervalo de duração varia de 1 a 16 dias, a precipitação acumulada calculada com o decaimento entre 23 a 158 milímetros e precipitação acumulada total de 31 a 314 milímetros. Dentre esses eventos, 9

(21,4%) dos eventos tiveram duração entre 29 a 72 horas, 10 (24%) apresentaram precipitação acumulado total menor que 80 milímetros e 6 (14%) com duração e precipitação acumulado total menor que 72 horas e 80 mm. Esses valores estão abaixo dos limiares estabelecidos no Plano Preventivo da Defesa Civil de São Paulo (PPDC).

Por exemplo, o evento ocorrido no dia 4 de janeiro de 2022 durou 43 horas com a precipitação acumulada de aproximadamente 60 mm. Embora o PPDC utilize uma metodologia distinta, considerando um período fixo de 72 horas para o cálculo da precipitação acumulada, a comparação é viável, pois o menor evento identificado pelo algoritmo teve 29 horas de duração, e o critério adotado para separar os eventos considera um período de no mínimo 48 horas anterior ao início do evento sem precipitação, ou seja, o evento com menor duração 29 horas mais 48 horas sem precipitação é maior que o tempo fixo de 72 horas do PPDC. Dessa forma, todos os 6 eventos mencionados teriam o mesmo valor de precipitação acumulado considerando um período fixo de 72 horas de duração e não seriam detectados seguindo o PPDC.

Em relação as características espaciais, a tabela 6.3 mostra as frequências relativas dos deslizamentos de terra nas modalidades das variáveis estudadas, essa frequência foi importante no auxílio da escolha das variáveis para o modelo.

A altitude do município de SBC varia de 67 a 978 metros que foram divididos em 4 classes. Os deslizamentos ocorreram exclusivamente na classe A4, demonstrando baixa representatividade para o modelo. Não é possível descartar a possibilidade de ocorrências de deslizamentos nas demais classes, porém altitudes inferiores se concentram em uma pequena região do município, longe do local de interesse e afastada das áreas urbanas. A reclassificação da variável dentro do intervalo de 750 a 980 metros não seria apropriada, uma vez a amplitude de 230 metros, dividida em classes, não iria representar as condições que afetam a estabilidade das encostas, causadas pela altitude, como alterações nas condições microclimáticas, cobertura vegetal e processos erosivos. Portanto, optamos por retirar essa variável do modelo.

Inicialmente, a declividade demonstra ser uma variável representativa, condizendo com o conhecimento prévio de que deslizamentos ocorrem com maior frequência em áreas mais inclinadas. Podemos observar que 50% dos eventos registrados ocorreram na classe D4, correspondente às regiões com inclinações superiores a 20 graus. Por outro lado, a classe D1 com os menores valores de inclinação não teve nenhuma ocorrência.

A orientação da vertente apresenta uma distribuição aparentemente aleatória. Esse comportamento possivelmente está relacionado à divisão em oito classes, aliado ao baixo volume de dados, o que, dificulta a identificação de tendências associadas a uma orientação específica.

A curvatura vertical ocorre com maior frequência em vertentes côncavas (CV1),

refletindo o conhecimento prévio de que essas áreas côncavas são mais suscetíveis a deslizamentos (ALCÂNTARA *et al.*, 2025). A curvatura horizontal é mais frequente nas vertentes convergentes, seguido por vertentes planares e, por último, divergentes, o que pode indicar uma tendência das ocorrências. Portanto, essa variável é mantida no modelo.

No município, a cobertura pedológica predominante é composta por CAMBIS-SOLOS, que ocupam 63,5% do território não alagado, as demais classes apresentam menor representatividade: LATOSOLOS (3,5%), NEOSSOLOS (2,3%) e GLEIS-SOLOS (2,0%). As áreas urbanas compondo 28,6% do território não alagado não foi mapeada, sendo que maior parte dos deslizamentos (81%) ocorreu nessa região, enquanto 19% foram registrados em CAMBISSOLOS. Essa alta frequência de deslizamentos em áreas urbanas pode estar relacionada à forma de coleta dos dados, majoritariamente provenientes de registros da defesa civil acionada por cidadãos. Consequentemente, essa variável demonstra baixa representatividade para o modelo. De maneira similar à distribuição pedológica, as superfícies antrópicas não agrícolas correspondem a 97,6% das ocorrências de deslizamento, tornando o uso e cobertura do solo pouco representativa para o modelo.

Após análise da frequência das ocorrências, as variáveis estáticas que foram selecionadas para teste no modelo são: declividade, orientação da vertente, curvatura vertical, curvatura horizontal, distância dos rios e distância das rodovias.

Tabela 6.3: Frequência relativa das ocorrências de deslizamento de terra dentro das modalidades das variáveis condicionantes.

| Variável | Codificação | Ocorrências | FR (%) |
|-----------------------------|-------------|-------------|--------|
| Altitude | A1 | 0 | 0 |
| | A2 | 0 | 0 |
| | A3 | 0 | 0 |
| | A4 | 42 | 100 |
| Declividade | D1 | 0 | 0 |
| | D2 | 5 | 11,9 |
| | D3 | 16 | 38,1 |
| | D4 | 21 | 50 |
| Orientação | O1 | 2 | 4,7 |
| | O2 | 0 | 0 |
| | O3 | 6 | 14,3 |
| | O4 | 3 | 7,1 |
| | O5 | 9 | 21,4 |
| | O6 | 5 | 12 |
| | O7 | 5 | 12 |
| | O8 | 12 | 28,6 |
| Curvatura Vertical | CV1 | 27 | 64,3 |
| | CV2 | 7 | 16,6 |
| | CV3 | 8 | 19 |
| Curvatura Horizontal | CH1 | 11 | 26,2 |
| | CH2 | 2 | 4,7 |
| | CH3 | 29 | 69 |
| Uso do Solo | U1 | 41 | 97,6 |
| | U2 | 1 | 2,4 |
| | U3 | 0 | 0 |
| | U4 | 0 | 0 |
| pedologia | P1 | 34 | 81 |
| | P2 | 8 | 19 |
| | P3 | 0 | 0 |
| | P4 | 0 | 0 |
| | P5 | 0 | 0 |

6.4 Limiares

Utilizamos a equação 5.2 para determinar os limiares de precipitação do município de SBC. Os parâmetros foram otimizados com dados de julho de 2021 a dezembro de 2022, contendo 20 deslizamentos de terra com as condições de precipitação e infiltração reconstruídas pelo algoritmo. A figura 6.3 apresenta a distribuição dos 20 pontos referente as condições de precipitação acumulado e duração do evento que resultaram em deslizamentos de terra (pontos vermelhos) e a equação com os parâmetros otimizados por mínimos quadrados plotados em coordenada logarítmica. Em seguida a função de densidade de probabilidade da distribuição $\delta(d)$ foi estimada por meio do método não paramétrico de densidade do kernel (KDE), conforme ilustrado na Figura 6.4. Os limiares foram definidos por diferentes probabilidades de excedência, deslocando o intercepto $\log(\alpha_5) = \log(\alpha_{50}) - \delta^*$, onde δ^* representa a distância entre a linha vertical preta, correspondente ao melhor ajuste dos parâmetros da equação, e a vermelha, o nível de excedência desejado.

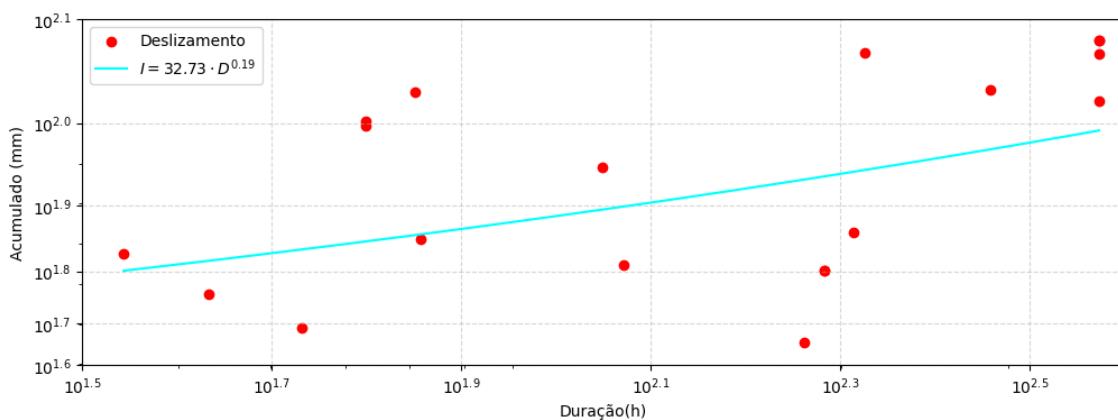


Figura 6.3: Eventos de precipitação que resultaram em deslizamentos de terra, no período de julho de 2021 a dezembro de 2022, com a curva das condições empíricas de Acumulado x Duração ajustada por mínimos quadrados.

Fonte: Próprio Autor

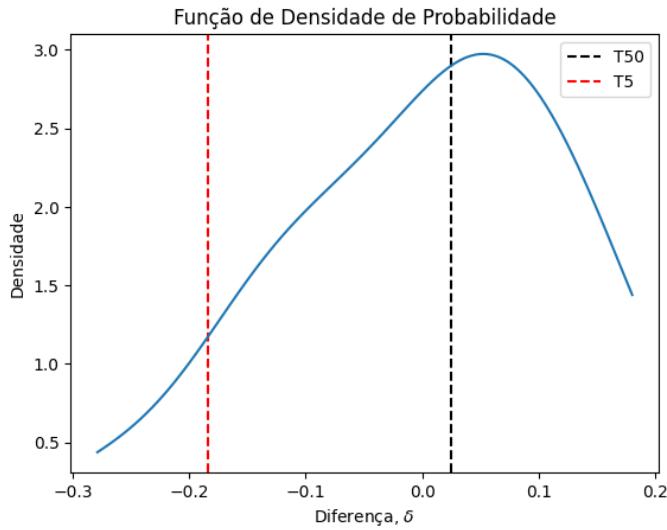


Figura 6.4: Estimativa da função de densidade de probabilidade da distribuição gaussiana.

Fonte: Próprio Autor

No presente trabalho os limiares de precipitação foram definidos em 1%, 10% e 30% de excedência e os limiares de saturação em 5% ($S_5 = 0,77$), 10% ($S_{10} = 0,85$) e 20% ($S_{20} = 0,95$). O conjunto de dados referente ao período de teste do modelo (intervalo de janeiro a dezembro de 2023) são apresentados na figura 6.5, contendo as três curvas de limiar de precipitação obtidas: 1% (T_1 , linha verde), 10% (T_{10} , linha azul) e 30% (T_{30} , linha vermelha), os pontos pretos representam eventos sem deslizamento e os vermelhos indicam ocorrências de deslizamento.

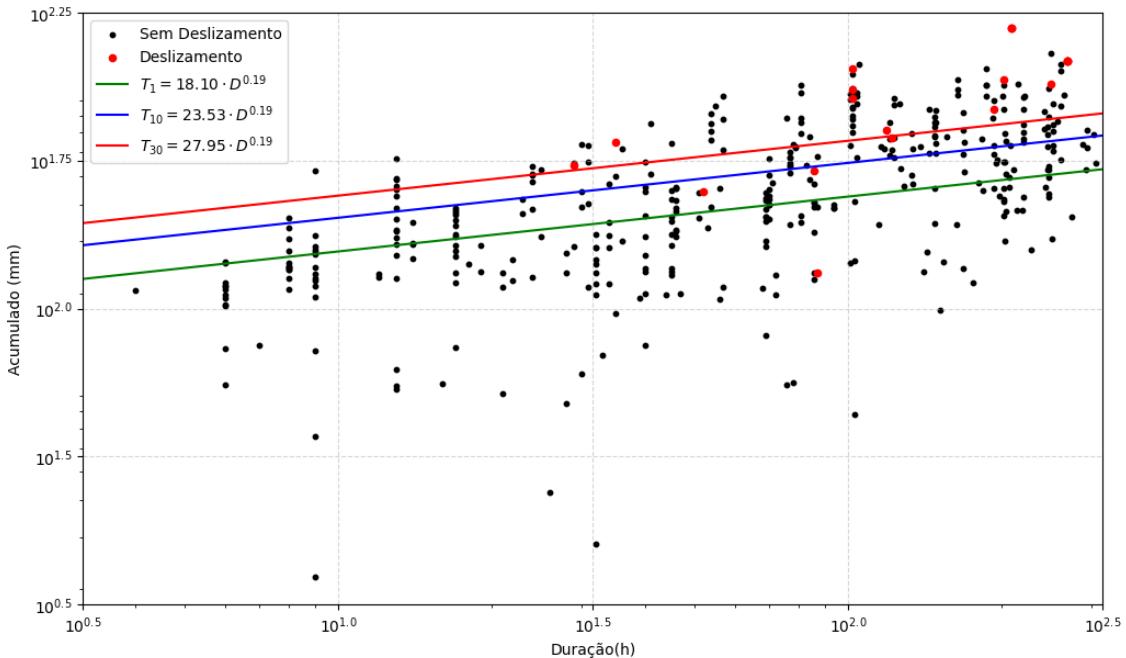


Figura 6.5: Curvas de limiar com os níveis de excedência.

Fonte: Próprio Autor

Um algoritmo de decisão foi desenvolvido para incorporar os limiares de precipitação com os de saturação do solo. Conforme ilustrado na figura 6.6 o algoritmo recebe como entrada os dados de precipitação acumulado, duração e saturação. Inicialmente, os valores da precipitação acumulado e duração são avaliados se estão acima de uma determinada curva de excedência de precipitação. Em seguida, é avaliado se a saturação ultrapassa algum limiar crítico, para então ser definido o nível de alerta adequado e as ações preventivas caso necessário. Os níveis de alerta e as ações utilizadas neste trabalho correspondem aos adotados no PPDC, tabela 3.3. Um modelo que utiliza as mesmas curvas de excedência de precipitação e os mesmos níveis operacionais de alerta, sem considerar a saturação do solo, figura 6.7, é adotado como referência o impacto da inclusão da variável saturação do solo na redução de alertas falsos.

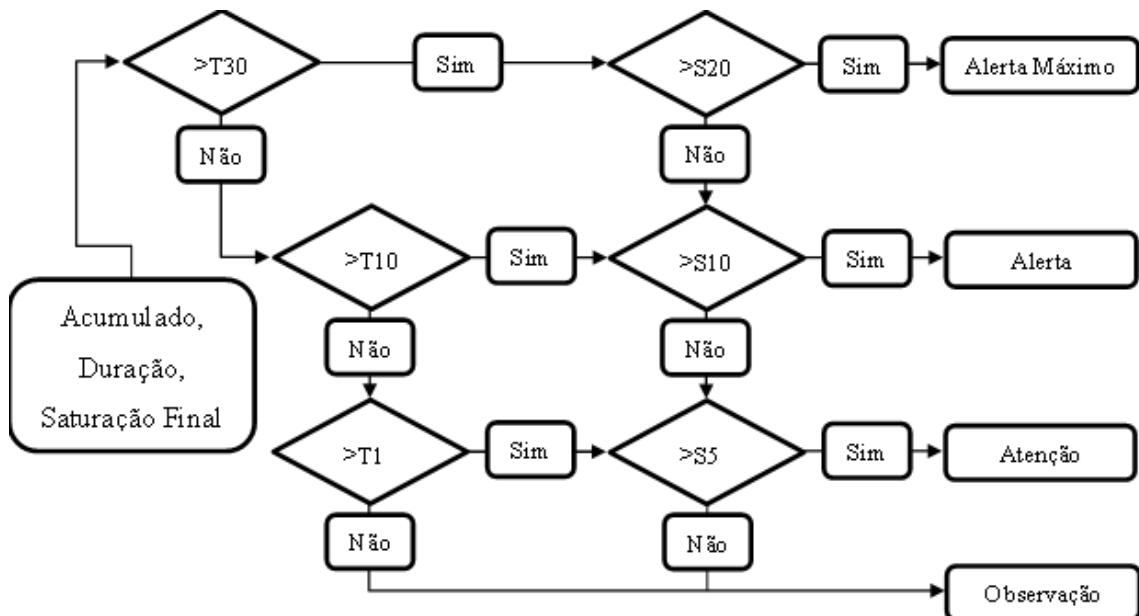


Figura 6.6: Algoritmo de decisão usando precipitação e saturação do solo.

Fonte: Próprio Autor

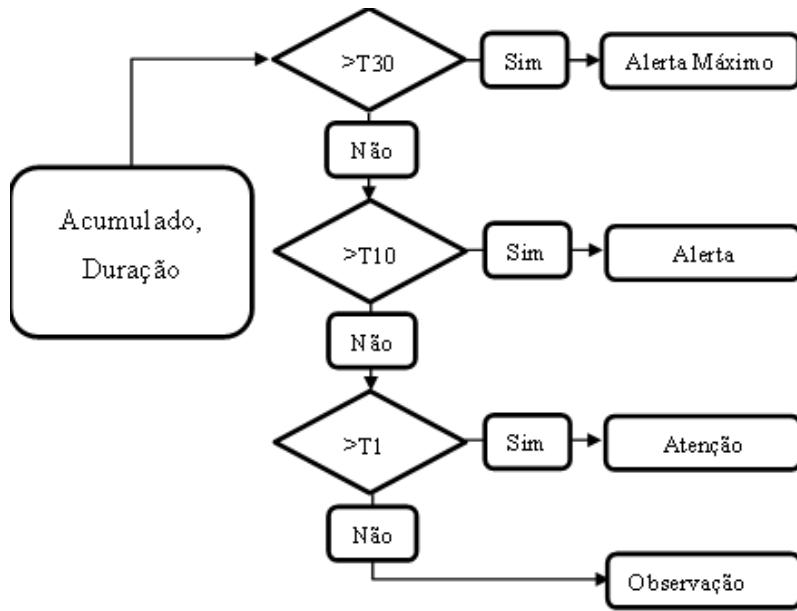


Figura 6.7: Algoritmo de decisão.

Fonte: Próprio Autor

O período de teste, de janeiro a dezembro de 2023, contém 22 eventos com ocorrência de deslizamento e 398 eventos sem ocorrência. A tabela 6.4 apresenta os valores de falsos positivos e verdadeiros positivos. Ambos os algoritmos de decisão previram corretamente 21 deslizamentos. A inclusão da saturação do solo resultou em um aumento de uma ocorrência no nível operacional "atenção", apesar da principal medida nesse nível ser a realização de vistorias em campo, enquanto no modelo anterior a ação recomendada para esse caso seria a remoção preventiva da população de áreas de risco, apenas uma ocorrência não oferece um impacto significativo.

Em compensação os falsos positivos nos níveis "Alerta" e "Alerta Máximo" foram reduzidos 43,9% e 66,3%, respectivamente, representando uma redução de 84 alertas, especialmente porque a principal medida nesses casos é a evacuação preventiva da população em área de risco. No nível "Atenção", houve um aumento de 15,6% nos falsos positivos, o que, apesar de indesejado, tem um baixo impacto em relação as melhorias dos níveis que oferecem mais perigo. Uma vantagem do sistema proposto em comparação com a do PPDC é a possibilidade de transitar entre os níveis operacionais, por exemplo, de "Observação" para "Atenção", sem a necessidade de vistoriar as áreas de risco in loco, além de poder direcionar as vistorias para regiões onde os sensores apresentam maior excedência.

Apesar da redução nos falsos positivos, o número de alertas falsos ainda é elevado. Em 2023, o modelo previu 62 alertas falsos nos níveis "Alerta" e "Alerta Máximo", ambos com recomendação de evacuação preventiva da população. O gráfico ilustrado na figura 6.8 auxilia a compreensão deste problema, onde os dados plotados apresentam saturação crítica superior a S20 (95%) e aqueles acima da curva T30 são

Tabela 6.4: Resultados da previsão dos algoritmos de decisão

| | | Precipitação | Precipitação e Saturação | Variação (%) |
|---------------------|---------------|--------------|--------------------------|--------------|
| Falso Positivo | Atenção | 81 | 96 | +15,6 |
| | Alerta | 57 | 32 | -43,9 |
| | Alerta Máximo | 89 | 30 | -66,3 |
| Verdadeiro Positivo | Atenção | 3 | 4 | +25 |
| | Alerta | 1 | 3 | +67 |
| | Alerta Máximo | 17 | 14 | -17,6 |
| | Total | 21 | 21 | 0 |

classificados como "Alerta Máximo". Nos dias 12, 19 e 21 de fevereiro de 2023, o município de SBC foi atingido por um grande volume de precipitação em grande parte do seu território, como cada sensor pluviométrico registra um evento ao detectar precipitação, o fenômeno resultou em múltiplos eventos acima do limiar sem deslizamento. O modelo assume como pressuposto a homogeneidade espacial da região, não considerando as características locais que podem influenciar a ocorrência de deslizamentos, tal limitação contribuiu para o elevado número de falsos alertas.

Outro fator que pode prejudicar o modelo é a baixa densidade de sensores de umidade do solo. Durante tempestades generalizadas, como as ocorridas nos dias 12, 19 e 21 de fevereiro, o município conta com 20 pluviômetros, mas apenas dois sensores de umidade do solo para representar toda a região. Essa limitação resulta em eventos registrados em locais distintos, com diferentes valores de precipitação acumulada, mas associados a um mesmo percentual de saturação do solo. A implementação de mais sensores de umidade do solo poderia melhorar a representação das condições de saturação do solo, reduzindo alertas falsos e melhorando a precisão do modelo.

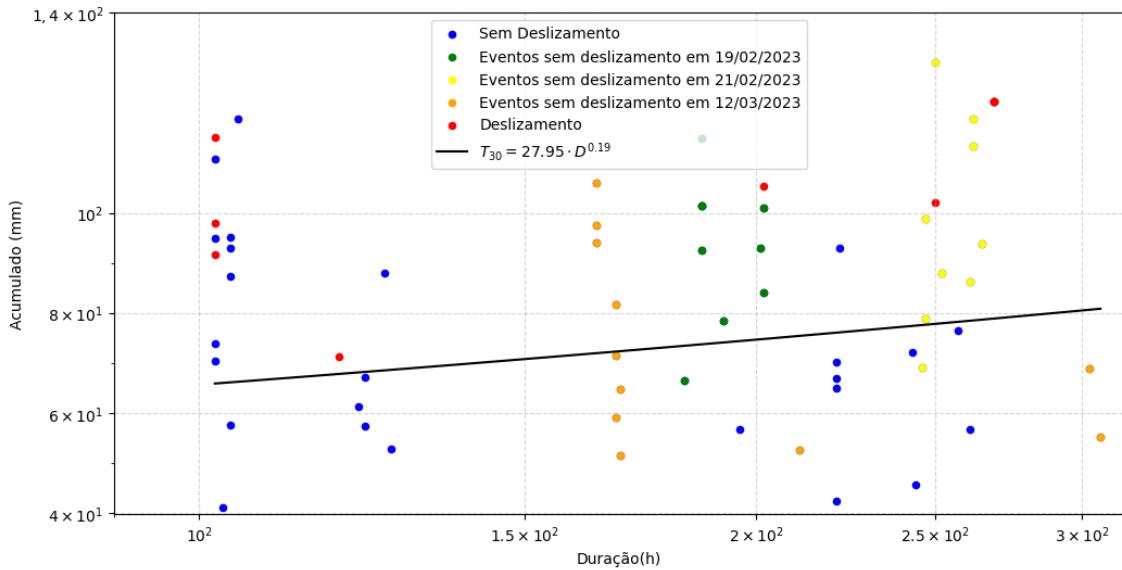


Figura 6.8: Limiar referente ao Alerta Máximo.

Fonte: Próprio Autor

Como o modelo não considera características espaciais, um único evento com o maior valor de precipitação acumulada no dia pode representar as condições mais críticas de estabilidade da encosta. Ao analisar os dados com os mesmos parâmetros para ambos os algoritmos de decisão, considerando apenas um evento por dia e desprezando aqueles sem deslizamento ocorridos em dia com deslizamento, resultaram em 37 dias com eventos de precipitação e infiltração sem ocorrência de deslizamento. Dos 22 deslizamentos registrados, 14 ocorreram em dias distintos ao longo de 2023.

A tabela 6.5 apresenta os resultados da previsão dos algoritmos de decisão. Ambos os modelos previram corretamente 13 dias com ocorrência de deslizamentos. Para os verdadeiros positivos, a inclusão da saturação do solo teve o mesmo efeito observado na análise anterior. Os falsos positivos nos níveis "Alerta" e "Alerta Máximo" somados foram reduzidos em 40% correspondendo a uma diminuição de oito dias. No período de 2023, o modelo tinha previsto 62 eventos com alertas falsos, em sua nova configuração representa 12 dias com alertas falsos com recomendação de evacuação preventiva da população em área de risco.

Tabela 6.5: Resultados da previsão dos algoritmos de decisão. Fonte: Próprio Autor.

| | | Precipitação | Precipitação e Saturação | Variação (%) |
|---------------------|---------------|--------------|--------------------------|--------------|
| Falso Positivo | Atenção | 11 | 8 | -27,3 |
| | Alerta | 2 | 8 | +75 |
| | Alerta Máximo | 18 | 4 | -77,8 |
| Verdadeiro Positivo | Atenção | 3 | 4 | +25 |
| | Alerta | 1 | 2 | +50 |
| | Alerta Máximo | 9 | 7 | -22,2 |
| | Total | 13 | 13 | 0 |

6.5 Regressão Logística

As variáveis restantes são precipitação acumulada, intensidade média, duração, saturação inicial, saturação final, mudança de saturação, declividade, orientação da vertente, curvatura vertical, curvatura horizontal, distância da rodovia, distância dos rios. Aplicamos o método de Stepwise utilizando o valor p com nível de significância de 5% como critério de entra e saída das variáveis em ambos procedimentos de amostragem para selecionar de forma automática as variáveis explicativas para o modelo de regressão logística.

Na primeira abordagem consistiu em selecionar aleatoriamente um evento de precipitação e infiltração por dia resultando em 102 amostras de não ocorrência de deslizamento, em seguida foi sorteado um ponto dentro de um raio de três quilômetros do sensor que registrou o evento e extraído as características espaciais do local. A tabela 6.6 mostra o processo de seleção das variáveis pelo método stepwise, na primeira iteração a variável selecionada para entrar no modelo que melhor explica o fenômeno segundo o valor p é a precipitação acumulado, na segunda iteração nenhuma variável foi significativa, todas apresentaram valor p superior a 5%, finalizando o método com a seleção apenas da variável precipitação acumulado.

Conforme a matriz de confusão, figura 6.9, o modelo previu corretamente 18 casos de deslizamento e atingiu uma acurácia de 71,8%. No entanto, uma métrica mais interessante para avaliação neste contexto é a precisão, figura 6.7. Para a classe de deslizamento, o modelo obteve precisão de na faixa de 56%, um valor considerado baixo, pois a ocorrência de falsos alertas pode comprometer a eficácia do sistema de monitoramento. Um número elevado de alarmes falsos pode levar à perda de credibilidade do sistema de alerta junto à população, reduzindo sua eficácia em

Tabela 6.6: Evolução do valor p no processo de seleção Stepwise

| Variável | Iteração 1 | Iteração 2 |
|------------------------|------------|------------|
| Intensidade Média | 0.1692 | 0.4749 |
| Declividade | 0.1361 | 0.8454 |
| Saturação Inicial | 0.0519 | 0.30989 |
| Saturação Final | 0.0011 | 0.26917 |
| Mudança da Saturação | 0.0004 | 0.81490 |
| Distância da Rodovia | 0.8083 | 0.41382 |
| Distância aos Rios | 0.0623 | 0.05844 |
| Curvatura Vertical | 0.5812 | 0.78070 |
| Precipitação Acumulado | 0.00001 | - |
| Orientação | 0.0730 | 0.64784 |
| Duração | 0.0002 | 0.79272 |
| Curvatura Horizontal | 0.0007 | 0.05401 |

situações de risco real e potencialmente resultando em perdas de vidas.

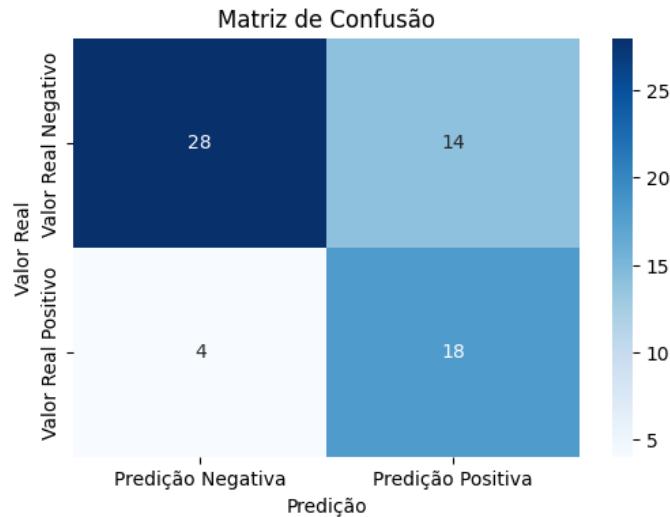


Figura 6.9: Matriz confusão para a primeira técnica de amostragem.

Fonte: Próprio Autor

Tabela 6.7: Relatório de Classificação do Modelo

| Classe | Precisão | Recall | F1-Score | Suporte |
|--------|----------|--------|----------|---------|
| 0 | 0.88 | 0.67 | 0.76 | 42 |
| 1 | 0.56 | 0.82 | 0.67 | 22 |

Nenhuma das variáveis espaciais atingiu nível de significância estatística suficiente para serem incluídas no modelo. Esse resultado contraria o conhecimento prévio sobre a influência desses fatores na ocorrência de deslizamentos de terra, uma vez que a literatura sugere que variáveis como declividade e curvatura do terreno desempenham um papel fundamental na estabilidade do solo (AHMED *et al.*, 2023;

ALCÂNTARA *et al.*, 2025; FABBRI *et al.*, 2003).

Na segunda abordagem, foram sorteados 60 pontos aleatoriamente dentro da área de estudo, independentemente da localização dos sensores. Para cada ponto, foi então sorteado a data de evento de precipitação e infiltração registrado pelo sensor mais próximo. A tabela 6.8 mostra o processo de seleção das variáveis pelo método stepwise, na primeira iteração a variável selecionada para entrar no modelo que melhor explica o fenômeno segundo o valor p é a precipitação acumulado, na segunda iteração é selecionada a declividade, em seguida realizado do backward e as variáveis precipitação acumulado e declividade se mantiveram no modelo. Na terceira iteração nenhuma variável foi significativa, todas apresentaram valor p superior a 5%, finalizando o método com a seleção das variáveis precipitação acumulado e declividade.

Tabela 6.8: Evolução do valor p no processo de seleção Stepwise

| Variável | Iteração 1 | Iteração 2 | Iteração 3 |
|------------------------|------------|------------|------------|
| Intensidade Média | 0.1492 | 0.1404 | 0.3983 |
| Declividade | 0.0012 | 0.0098 | - |
| Saturação Inicial | 0.0388 | 0.5575 | 0.2847 |
| Saturação Final | 0.0021 | 0.5751 | 0.2864 |
| Mudança de Saturação | 0.0008 | 0.9612 | 0.7934 |
| Distância da Rodovia | 0.1799 | 0.7829 | 0.5896 |
| Distância aos Rios | 0.0769 | 0.0228 | 0.0613 |
| Curvatura Vertical | 0.0677 | 0.2569 | 0.6812 |
| Precipitação Acumulado | 0.0001 | - | - |
| Orientação | 0.0075 | 0.3888 | 0.8401 |
| Duração | 0.001 | 0.1855 | 0.2566 |
| Curvatura Horizontal | 0.0374 | 0.7358 | 0.6635 |

Conforme a matriz de confusão, figura 6.10, o modelo previu corretamente 19 casos de deslizamento e atingiu uma precisão de 86% para a classe de deslizamento, figura 6.9. Embora essa abordagem tenha apresentado uma precisão superior à técnica de amostragem anterior, não é possível fazer uma comparação direta devido às diferenças no tamanho das amostras. No entanto, esse modelo é considerado mais adequado, pois incorpora a declividade além da precipitação acumulada, o que faz sentido físico onde terreno mais inclinado são mais propícios a ocorrência de deslizamentos.

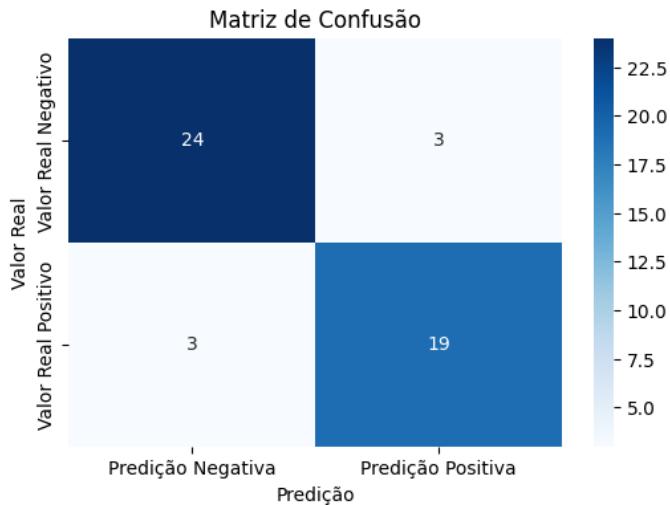


Figura 6.10: Matriz confusão para a segunda técnica de amostragem.

Fonte: Próprio Autor

Tabela 6.9: Relatório de Classificação

| Classe | Precisão | Recall | F1-Score | Suporte |
|--------|----------|--------|----------|---------|
| 0 | 0.89 | 0.89 | 0.89 | 27 |
| 1 | 0.86 | 0.86 | 0.86 | 22 |

Avaliamos o desempenho do modelo por meio da curva ROC, que representa a relação entre a taxa de verdadeiros positivos e a taxa de falsos positivos para diferentes limiares de decisão do modelo. Na figura 6.11, a linha tracejada azul indica um classificador aleatório, com probabilidade de acerto e erro igual. Isso significa que um modelo com desempenho próximo dessa linha não é melhor do que escolher aleatoriamente. Para quantificar a capacidade do modelo de separar corretamente as classes, utilizamos a métrica AUC (do inglês, *area under the curve*), representada em laranja. Um modelo ideal teria a curva mais próxima do canto superior esquerdo com valor de AUC igual a 1. O modelo indica AUC = 0.94 demonstrando um excelente poder de discriminação entre classes positivas e negativas.

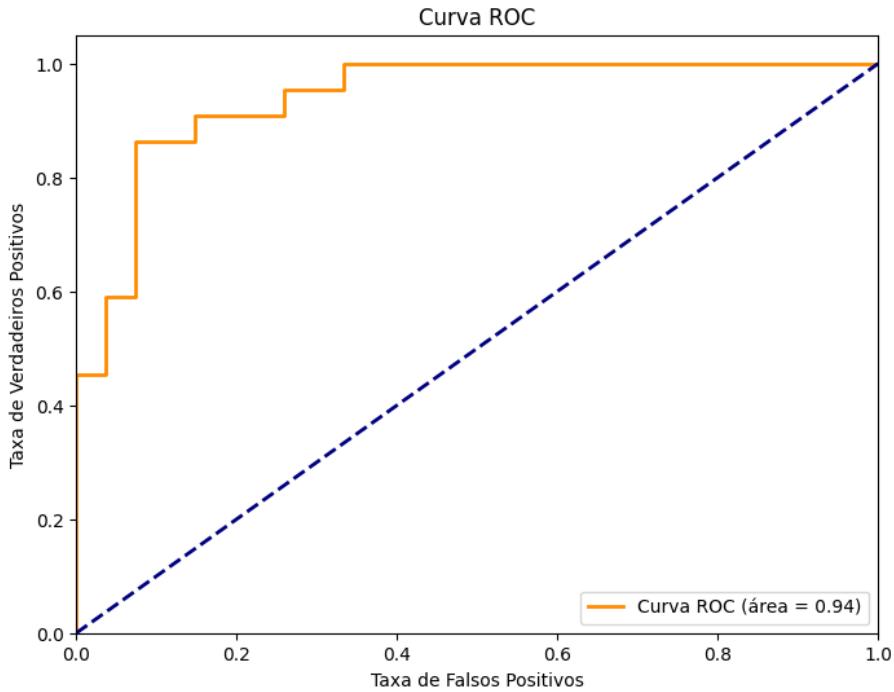


Figura 6.11: Curva ROC.

Fonte: Próprio Autor

A análise da contribuição de cada variável condicionante é essencial para compreender sua influência na ocorrência de deslizamentos de terra. Como fator desencadeante, conforme a figura 6.12 a precipitação acumulada apresentou a maior taxa de contribuição, correspondendo a 32%, seguida da declividade, com 4,5%. As demais variáveis, eliminadas pelo método stepwise, mostraram contribuições inferiores à da declividade, sendo que, em alguns casos, contribuíram negativamente para o modelo.

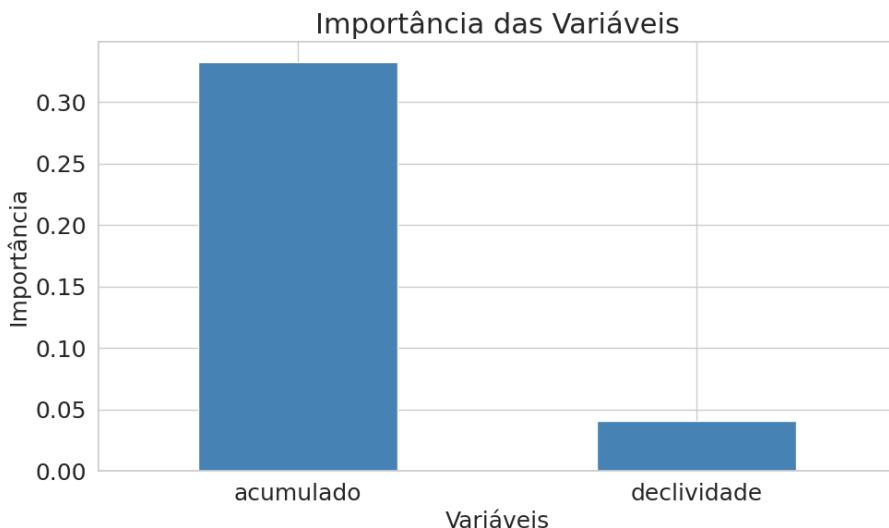


Figura 6.12: Importância das Variáveis do modelo.

Fonte: Próprio Autor

A probabilidade de ocorrência de deslizamento pode ser expressa pela seguinte equação:

$$p = \frac{1}{1 + e^{(-0.496 + 2.055 \cdot PA + 1.096 \cdot D)}} \quad (6.1)$$

em que $-0,496$ é o intercepto, PA é a precipitação acumulada e D é a declividade.

Os deslizamentos em São Bernardo do Campo são predominantemente induzidos por chuvas, o que evidencia a importância de um modelo dinâmico capaz de capturar a evolução da suscetibilidade em função das variações das variáveis de entrada, como a precipitação acumulada. Modelos tradicionais de suscetibilidade, em sua maioria, consideram apenas variáveis estáticas e, quando incorporam variáveis dinâmicas, geralmente utilizam resoluções temporais anuais, como a média da precipitação anual, o que impossibilita sua utilização em um sistema de alerta antecipado.

Nosso modelo possibilita predição horária das condições do terreno, permitindo a predição da evolução da suscetibilidade no espaço e no tempo. Para ilustrar essa dinâmica, realizamos simulações para os dias 18, 19, 20 e 23 de fevereiro de 2023, a partir da equação 6.1. A fim de compatibilizar as dimensões das variáveis, interpolamos os dados de precipitação acumulada, calculados pelo algoritmo, por meio do inverso do quadrado da distância, ajustando-os à resolução espacial da declividade. Podemos notar na figura 6.13 o aumento da suscetibilidade ao longo do tempo, à medida que a precipitação acumulada aumenta. Os pontos pretos ilustrados nas figuras indicam a distribuição espacial de deslizamentos nos respectivos dias de ocorrência. No dia 23, podemos notar que praticamente toda a região apresenta alta suscetibilidade a deslizamento, resultado da elevada contribuição da precipitação acumulada no modelo e devido a ocorrências de deslizamento em locais com baixa declividade.

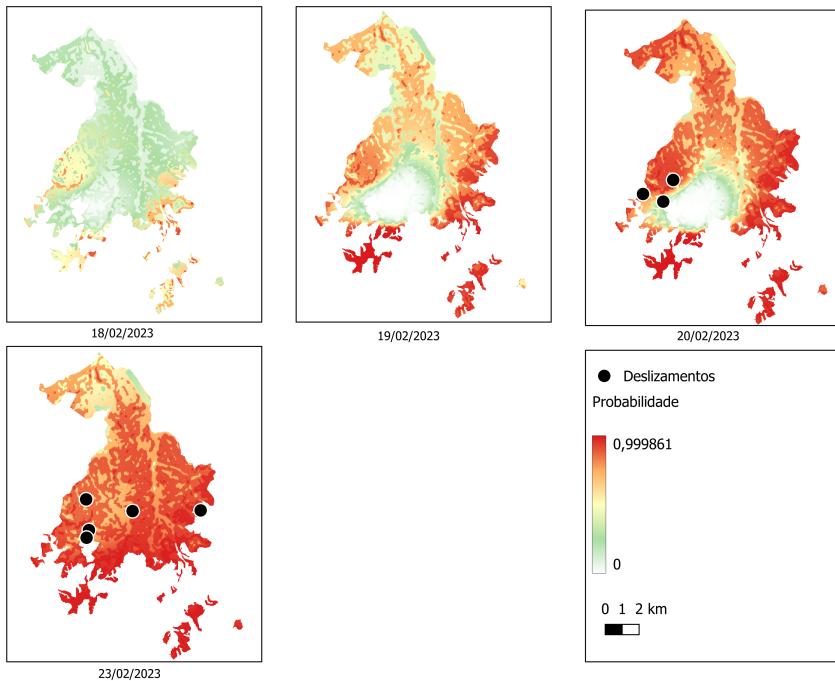


Figura 6.13: Suscetibilidade prevista a deslizamentos de terra para os dias 18, 19, 20 e 23 de fevereiro de 2023.

Fonte: Próprio Autor

A ausência ou a baixa significância estatística das variáveis topológica pode estar relacionada a três fatores, a característica da região, a forma de coleta dos dados e o curto período de análise.

A região urbana de SBC é relativamente pequena em comparação com outras regiões estudadas, que geralmente consideram escalas regionais (AHMED *et al.*, 2023; ALCÂNTARA *et al.*, 2025; HUANG *et al.*, 2022; REICHENBACH *et al.*, 2018). Inicialmente, a proposta do trabalho era analisar todo o município, entretanto, os registros de deslizamento no período analisado estão exclusivamente na área urbana. As ações antrópicas causadas pelo alto nível de urbanização na região podem ter influenciado a ocorrência de deslizamento, superando os efeitos das características físicas do terreno. Esse impacto pode ser observado no presente estudo, assim como nos mapas de suscetibilidade e perigo, nos quais diversos deslizamentos ocorreram em áreas classificadas como de baixa suscetibilidade e periculosidade, figura 6.14. Esses mapas, elaborado pelo (INSTITUTO GEOLÓGICO, 2015, 2020), foram desenvolvidos considerando a declividade, forma da encosta e amplitude altimétrica como as características principais condicionantes a deslizamento.

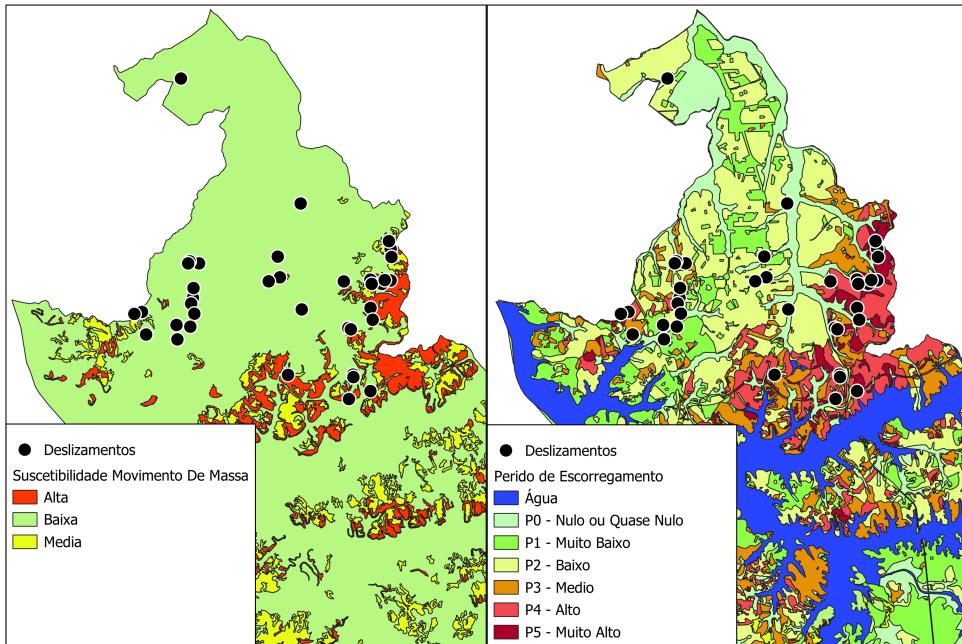


Figura 6.14: Mapas de Suscetibilidade e perigo.

Fonte: adaptado de, CPRM.

Conforme foi citado na seção 6.1 existe um erro inerente a forma em que os dados de deslizamentos foram coletados tanto no espaço como no tempo. Além disso, o banco de dados contem deslizamentos de pequeno porte que ocorreram em muros de arrimo, bem como deslizamentos de grande porte, a falta de detalhamento dos dados em relação a descrição das características e fatores externos não relacionados a precipitação que causaram o deslizamento dificulta a análise. Em uma revisão de literatura considerando 565 artigos REICHENBACH *et al.* (2018) relatou que quase 70% modelos estatísticos utilizam informações sobre deslizamentos obtidas de mapas de inventário preparados por meio da interpretação visual de fotografias aéreas estereoscópicas ou imagens de satélite. Essa abordagem pode não capturar pequenos deslizamentos, especialmente aqueles que ocorrem em áreas urbanizadas, que são rapidamente removidos e causam pouca alteração na paisagem.

O pequeno período de dados disponíveis para umidade do solo restringiu a análise. Os primeiros sensores no estado de São Paulo foram instalados em meados de 2021 resultando a apenas dois anos e meio, dos quais um ano e meio foi utilizado para treinamento e um ano para teste. Esse intervalo resultou em um total de 42 ocorrências, o que pode ser considerado um número pequeno de dados para treinar o modelo. Um conjunto de dados contemplando um período maior teria mais amostras de deslizamentos e poderia representar melhor a região.

Na análise apresentada na seção 5.5, identificamos diversas estações que não estão enviando dados ou apresentam valores irreais de umidade. Tanto o modelo de limiar

proposto neste trabalho quanto aqueles abordados na revisão da literatura indicam melhorias com a inclusão da umidade do solo. Dessa forma, a manutenção adequada dos sensores existentes, assim como a instalação de novos sensores, é essencial para aprimorar os modelos de predição.

Capítulo 7

Conclusões e Próximas Etapas

O algoritmo demonstrou boa capacidade de reconstrução dos eventos de infiltração e precipitação. Das 50 ocorrências de deslizamento, reconstruiu 42, correspondendo a 84% dos casos. As falhas na reconstrução ocorreram devido à ausência de sensores dentro da distância estabelecida como parâmetro ou à inexistência de variação na precipitação e na umidade do solo nos sensores próximos. Além de garantir uma reconstrução objetiva e reproduzível dos eventos, o algoritmo assegura a consistência metodológica reduzindo possíveis viés do pesquisador, reconstrói os eventos de forma independente das ocorrências de deslizamentos e reduz o tempo necessário para a análise dos eventos em comparação com o método de reconstrução manual.

Apesar curto período de tempo e poucos dados para o treinamento do modelo de limiar, a inclusão da umidade do solo, calculado pelo algoritmo, demonstrou potencial ao reduzir os falsos positivos nos níveis "Alerta" e "Alerta Máximo" em 43,9% e 66,3%, respectivamente, resultando em uma redução de 84 alertas falsos. No entanto, o número de falsos alertas ainda permanece elevado. Em um período de um ano, o modelo previu alertas falsos para 62 eventos distribuídos em 12 dias distintos. Isso possivelmente ocorre por uma limitação do modelo em não considerar outras variáveis condicionantes do fenômeno como características topográficas.

Para superar essa limitação, propomos um modelo espaço-temporal que incorpora condições de precipitação em tempo real, juntamente com a declividade, para a predição de deslizamentos de terra. O modelo apresenta resultados promissores em termos de todos os indicadores de desempenho, incluindo acurácia, recall, precisão, AUC. Além disso, a abordagem permite a predição horária, capturando de forma dinâmica o aumento da suscetibilidade de ocorrer deslizamentos à medida que a precipitação se intensifica.

A precipitação acumulada foi a variável com maior contribuição para o modelo, seguida da declividade. As demais variáveis, incluindo as derivadas da umidade do solo e as variáveis espaciais (exceto a declividade), não foram estatisticamente significativas para inclusão no modelo segundo o método de stepwise. Esse resultado

provavelmente está relacionado à pequena variação em algumas variáveis na região, à forma que os dados de deslizamento são coletados ou ao curto período de análise.

Uma das principais dificuldades enfrentadas neste estudo decorreram da ausência de detalhes na base de registros de movimentos de massa. Houve uma grande dificuldade em distinguir registros que de fato relatam a ocorrência de deslizamentos induzido por chuva daqueles que se referiam a outras situações, como visitas técnicas para avaliação de risco em residências, vazamentos nos sistemas de água ou esgoto e execução de obras. A padronização desses registros, incluindo informações detalhadas sobre fatores condicionantes, tipo de movimento de massa, localização exata, velocidade do deslocamento e características do plano de ruptura, contribuiria para o aprimoramento de futuros modelos de predição.

De maneira geral o modelo atendeu à proposta de prever a evolução espaço-temporal da suscetibilidade, porém, para valores elevados de precipitação acumulada, praticamente toda a região é classificada com alta suscetibilidade. Dessa forma, alguns aspectos ainda demandam investigação adicional antes da implementação em sistemas de alertas operacionais. Para as próximas etapas do estudo, podemos destacar a inclusão de variáveis espaciais mais adequadas ao contexto urbano, que possam auxiliar na discriminação das áreas mais suscetíveis, como densidade e estágio de ocupação do município, índice de saneamento básico e uso e ocupação do solo de forma mais detalhada, podem aprimorar a capacidade preditiva do modelo. Além disso, devemos testar a utilização de dados de previsão de precipitação no modelo, em vez de projeções passadas.

Em relação ao algoritmo de reconstrução de eventos de infiltração e precipitação as próximas etapas incluem a análise de complexidade, que visa determinar a ordem de tempo de execução do algoritmo e identificar possíveis otimizações para melhorar a eficiência computacional. Além disso, a paralelização do código será explorada para reduzir o tempo computacional, possibilitando sua aplicação em grandes volumes de dados. Outra melhoria interessante é o desenvolvimento de uma interface que facilite a utilização do algoritmo pelos usuários.

Referências Bibliográficas

- AHMED, M., TANYAS, H., HUSER, R., et al., 2023, “Dynamic rainfall-induced landslide susceptibility: a step towards a unified forecasting system”, *International Journal of Applied Earth Observation and Geoinformation*, v. 125, pp. 103593.
- ALCÂNTARA, E., BAIÃO, C. F., GUIMARÃES, Y. C., et al., 2025, “Machine Learning Reveals Lithology and Soil as Critical Parameters in Landslide Susceptibility for Petrópolis (Rio de Janeiro State, Brazil)”, *Natural Hazards Research*.
- ALEOTTI, P., 2004, “A warning system for rainfall-induced shallow failures”, *Engineering geology*, v. 73, n. 3-4, pp. 247–265.
- AUGUSTO FILHO, O., 1992, “Caracterização geológico-geotécnica voltada à estabilização de encostas: uma proposta metodológica”. In: *Conferência Brasileira sobre Estabilidade de Encostas-COBRAE. Anais*, pp. 721–733.
- BAUM, R. L., SAVAGE, W. Z., GODT, J. W., 2002, *TRIGRS: a Fortran program for transient rainfall infiltration and grid-based regional slope-stability analysis*. US Geological Survey Reston, VA, USA.
- BERTI, M., MARTINA, M., FRANCESCHINI, S., et al., 2012, “Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach”, *Journal of Geophysical Research: Earth Surface*, v. 117, n. F4.
- BERTONI, J., LOMBARDI NETO, F., 1999, “Erosão”, *BERTONI, J.; LOMBARDI NETO, F. Conservação do Solo. 4^a. ed. São Paulo: Ícone. Cap*, v. 7, pp. 68–93.
- BITAR, O. Y., 2014, *Cartas de suscetibilidade a movimentos gravitacionais de massa e inundações-1: 25.000: nota técnica explicativa*. Relatório técnico, IPT; CPRM.

- BONGIOVANNI, L. A., FUKUMOTO, M. M., ANTONIO, C. A., et al., 2011, “Operação guarda-chuva: prevenção e resposta a desastres naturais em São Bernardo do Campo-SP”. In: *13º Congresso Brasileiro de Geologia de Engenharia e Ambiental. São Paulo, Associação Brasileira de Geologia de Engenharia e Ambiental*.
- BRABB, E., 1984, “Innovative approaches to landslide hazard and risk mapping”. In: *International Landslide Symposium Proceedings, Toronto, Canada*, v. 1, pp. 17–22.
- BRUNETTI, M. T., PERUCCACCI, S., ROSSI, M., et al., 2010, “Rainfall thresholds for the possible occurrence of landslides in Italy”, *Natural Hazards and Earth System Sciences*, v. 10, n. 3, pp. 447–458.
- CAINE, N., 1980, “The rainfall intensity-duration control of shallow landslides and debris flows”, *Geografiska annaler: series A, physical geography*, v. 62, n. 1-2, pp. 23–27.
- CEMADEN, 2022. “Mapa Interativo”. Website. Disponível em www.cemaden.gov.br/mapainterativo. Acesso em: 20 de novembro 2023.
- CHAE, B.-G., PARK, H.-J., CATANI, F., et al., 2017, “Landslide prediction, monitoring and early warning: a concise review of state-of-the-art”, *Geosciences Journal*, v. 21, pp. 1033–1070.
- CHLEBORAD, A. F., BAUM, R. L., GODT, J. W., et al., 2008, “A prototype system for forecasting landslides in the Seattle, Washington, area”, *Geological Society of America*, p. 103–120.
- CHUNG, M. C., TAN, C. H., CHEN, C. H., 2017, “Local rainfall thresholds for forecasting landslide occurrence: Taipingshan landslide triggered by Typhoon Saola”, *Landslides*, v. 14, n. 1, pp. 19–33.
- CIARAPICA, L., TODINI, E., 2002, “TOPKAPI: A model for the representation of the rainfall-runoff process at different scales”, *Hydrological Processes*, v. 16, n. 2, pp. 207–229.
- COTTING, A. L. M., CANIL, K., HORITA, F., et al., 2024, “Análise das inovações metodológicas nos planos municipais de redução de riscos, 2020-2021, de São Bernardo do Campo e Franco da Rocha, estado de São Paulo, Brasil”, *Revista Ecologias Humanas*, v. 10, n. 12, pp. 65–83.
- CROZIER, M. J., 1999, “Prediction of rainfall-triggered landslides: A test of the antecedent water status model”, *Earth Surface Processes and Landforms*:

The Journal of the British Geomorphological Research Group, v. 24, n. 9, pp. 825–833.

- CRUDEN, D. M., VARNES, D. J., 1996, “Landslide types and processes”, *Landslides: investigation and mitigation, Transportation research board special report*, v. 247, pp. 36–57.
- CUNHA, M., FARAH, F., CERRI, L., et al., 1991, *Ocupação de encostas*. Publicação IPT. IPT. Disponível em: <<https://books.google.com.br/books?id=mREqmgEACAAJ>>.
- DOS SANTOS ALVALÁ, R. C., DE ASSIS DIAS, M. C., SAITO, S. M., et al., 2019, “Mapping characteristics of at-risk population to disasters in the context of Brazilian early warning system”, *International Journal of Disaster Risk Reduction*, v. 41, pp. 101326.
- EFRON, B., 1992, “Bootstrap methods: another look at the jackknife”. In: *Breakthroughs in statistics: Methodology and distribution*, Springer, pp. 569–593.
- EFRON, B., TIBSHIRANI, R. J., 1994, *An introduction to the bootstrap*. Chapman and Hall/CRC.
- FABBRI, A. G., CHUNG, C.-J. F., CENDRERO, A., et al., 2003, “Is prediction of future landslides possible with a GIS?” *Natural Hazards*, v. 30, pp. 487–503.
- FAWCETT, T., 2006, “An introduction to ROC analysis”, *Pattern recognition letters*, v. 27, n. 8, pp. 861–874.
- FIGUEIRA, C. V., 2006, *Modelos de regressão logística*. Dissertação de mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.
- FRATTINI, P., CROSTA, G., SOSIO, R., 2009, “Approaches for defining thresholds and return periods for rainfall-triggered shallow landslides”, *Hydrological Processes: An International Journal*, v. 23, n. 10, pp. 1444–1460.
- FREDLUND, D., MORGENSEN, N. R., WIDGER, R., 1978, “The shear strength of unsaturated soils”, *Canadian geotechnical journal*, v. 15, n. 3, pp. 313–321.
- GUIDICINI, G., NIEBLE, C. M., 1976, *Estabilidade de taludes naturais e de escavação*. Editora Blucher.

- GUZZETTI, F., REICHENBACH, P., CARDINALI, M., et al., 2005, "Probabilistic landslide hazard assessment at the basin scale", *Geomorphology*, v. 72, n. 1-4, pp. 272–299.
- GUZZETTI, F., PERUCCACCI, S., ROSSI, M., et al., 2007, "Rainfall thresholds for the initiation of landslides in central and southern Europe", *Meteorology and atmospheric physics*, v. 98, pp. 239–267.
- GUZZETTI, F., PERUCCACCI, S., ROSSI, M., et al., 2008, "The rainfall intensity-duration control of shallow landslides and debris flows: an update", *Landslides*, v. 5, pp. 3–17.
- HASTIE, T., TIBSHIRANI, R., FRIEDMAN, J. H., et al., 2009, *The elements of statistical learning: data mining, inference, and prediction*, v. 2. Springer.
- HE, H., GARCIA, E. A., 2009, "Learning from imbalanced data", *IEEE Transactions on knowledge and data engineering*, v. 21, n. 9, pp. 1263–1284.
- HUANG, F., CHEN, J., LIU, W., et al., 2022, "Regional rainfall-induced landslide hazard warning based on landslide susceptibility mapping and a critical rainfall threshold", *Geomorphology*, v. 408, pp. 108236.
- INSTITUTO GEOLÓGICO, 2015, *Mapeamento de Riscos de Movimentos de Massa e Inundações do Município de São Bernardo do Campo (2015): Relatório Técnico*. Relatório técnico, Instituto Geológico, São Paulo. Disponível em: <<https://rigeo.sgb.gov.br/handle/doc/15213>>.
- INSTITUTO GEOLÓGICO, 2020, *Mapeamento de Riscos de Movimentos de Massa e Inundações do Município de São Bernardo do Campo (2020): Relatório Técnico*. Relatório técnico, Instituto Geológico, São Paulo. Disponível em: <https://www.sidec.sp.gov.br/map_risco/uploads/doc1606403526.pdf>.
- IVERSON, R. M., 2000, "Landslide triggering by rain infiltration", *Water resources research*, v. 36, n. 7, pp. 1897–1910.
- KING, G., ZENG, L., 2001, "Logistic regression in rare events data", *Political analysis*, v. 9, n. 2, pp. 137–163.
- KRAWCZYK, B., 2016, "Learning from imbalanced data: open challenges and future directions", *Progress in artificial intelligence*, v. 5, n. 4, pp. 221–232.
- LUMB, P., 1975, "Slope failures in Hong Kong", *Quarterly Journal of Engineering Geology and Hydrogeology*, v. 8, n. 1, pp. 31–65.

- MARTELLONI, G., SEGONI, S., FANTI, R., et al., 2012, “Rainfall thresholds for the forecasting of landslide occurrence at regional scale”, *Landslides*, v. 9, pp. 485–495.
- MELILLO, M., BRUNETTI, M. T., PERUCCACCI, S., et al., 2015, “An algorithm for the objective reconstruction of rainfall events responsible for landslides”, *Landslides*, v. 12, pp. 311–320.
- MELILLO, M., BRUNETTI, M. T., PERUCCACCI, S., et al., 2018, “A tool for the automatic calculation of rainfall thresholds for landslide occurrence”, *Environmental Modelling & Software*, v. 105, pp. 230–243.
- MENDES, R. M., VALERIO FILHO, M., SANTORO, J., et al., 2020, “Proposição de limiares críticos ambientais para uso em sistema de alertas de deslizamentos”, *Revista do Departamento de Geografia*, v. 40, pp. 61–77.
- MIRUS, B. B., BECKER, R. E., BAUM, R. L., et al., 2018, “Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning”, *Landslides*, v. 15, pp. 1909–1919.
- NOGUEIRA, F. R., CANIL, K., 2017, “ST 8 CARTAS GEOTÉCNICAS DE AP-TIDÃO À URBANIZAÇÃO: INSTRUMENTO DE PLANEJAMENTO PARA PREVENÇÃO DE DESASTRES E PARA A GESTÃO DO USO DO SOLO”, *Anais ENANPUR*, v. 17, n. 1.
- PERROTTA, M., MARCONATO, A., TURRA, B., et al., 2006, “Mapa geológico do Estado de São Paulo: breve descrição das unidades litoestratigráficas aflorantes no Estado de São Paulo”, *Rio de Janeiro: CPRM*.
- PERUCCACCI, S., BRUNETTI, M. T., LUCIANI, S., et al., 2012, “Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy”, *Geomorphology*, v. 139, pp. 79–90.
- RAFAELLI, C. B. S. AND FERREIRA, C. J., 2022. “Cadastro de eventos geodinâmicos e desastres do Estado de São Paulo 1991-2020.” Planilha eletrônica On-line. Disponível em https://docs.google.com/spreadsheets/d/1uElUITGvYArX23E6YpycYLtObaA_YmdK8559wBCY-TI/edit?usp=sharing. Acesso em: 20 de novembro 2023.
- REICHENBACH, P., CARDINALI, M., DE VITA, P., et al., 1998, “Regional hydrological thresholds for landslides and floods in the Tiber River Basin (central Italy)”, *Environmental Geology*, v. 35, n. 2, pp. 146–159.

- REICHENBACH, P., ROSSI, M., MALAMUD, B. D., et al., 2018, “A review of statistically-based landslide susceptibility models”, *Earth-science reviews*, v. 180, pp. 60–91.
- RICHARDS, L. A., 1931, “Capillary conduction of liquids through porous media”, *physics*, v. 1, n. 5, pp. 318–333.
- ROSSI, M., PERUCCACCI, S., BRUNETTI, M., et al., 2012, “SANF: National warning system for rainfall-induced landslides in Italy”. In: *Landslide and Engineered Slopes. Protecting society through improved understanding.*, v. 2, pp. 1895–1899.
- SANTORO, J., MENDES, R. M., PRESSINOTTI, M. M. N., et al., 2010, “Correlação entre chuvas e deslizamentos ocorridos durante a operação do plano preventivo de defesa civil em São Paulo, SP”, *SIMPÓSIO BRASILEIRO DE CARTOGRAFIA GEOTÉCNICA, 7, Maringá-PR*.
- SÃO BERNARDO, D. C., 2010, “Plano Municipal de Redução de Riscos de São Bernardo do Campo. 2010b”, *SBC: PMSBC/FUNEP*.
- SCOTT, D. W., 1992, *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.
- SEGONI, S., ROSI, A., LAGOMARSINO, D., et al., 2018, “Brief communication: Using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system”, *Natural Hazards and Earth System Sciences*, v. 18, pp. 807–812.
- SELBY, M., 1993, *Hillslope Materials*. Oxford University Press, 451p.
- SILVERMAN, B. W., 1986, *Density estimation for statistics and data analysis*. Routledge.
- TATIZANA, C., OGURA, A. T., CERRI, L. D. S., et al., 1987a, “Análise de correlação entre chuvas e escorregamentos–Serra do Mar, município de Cubatão”. In: *Congresso Brasileiro de Geologia de Engenharia*, v. 5, pp. 225–236, a.
- TATIZANA, C., OGURA, A. T., CERRI, L. E., et al., 1987b, “Modelamento numérico da análise de correlação entre chuvas e escorregamentos aplicado às encostas da Serra do Mar no município de Cubatão”. In: *En Anais do Congresso Brasileiro de Geologia de Engenharia*, pp. 237–248, b.
- TOLL, D., LOURENÇO, S., MENDES, J., et al., 2011, “Soil suction monitoring for landslides and slopes”. Geological Society of London.

TOMINAGA, L. K., SANTORO, J., AMARAL, R., 2009, “Desastres naturais”, *São Paulo: Editora Instituto Geológico.*

TOMINAGA, L. K., SANTORO, J., AMARAL, R., 2015, “Desastres naturais: conhecer para prevenir”, *São Paulo: Editora Instituto Geológico.*

TSAI, T.-L., CHEN, H.-E., YANG, J.-C., 2008, “Numerical modeling of rainstorm-induced shallow landslides in saturated and unsaturated soils”, *Environmental Geology*, v. 55, pp. 1269–1277.

VALERIANO, M. D. M., 2008, “Topodata: guia para utilização de dados geomorfológicos locais”, *São José dos Campos: INPE*, v. 72.

VENABLES, W. N., RIPLEY, B. D., 2002, *Modern applied statistics with S-PLUS*. Springer Science & Business Media.

WICKI, A., LEHMANN, P., HAUCK, C., et al., 2020, “Assessing the potential of soil moisture measurements for regional landslide early warning”, *Landslides*, v. 17, pp. 1881–1896.