

**Universidade Federal de Ouro Preto**

Núcleo de Geotecnia

Programa de Pós-Graduação em Geotecnia  
PPGEO

---

Tese

---

# **PROPOSTA DE SISTEMA DE ANÁLISE DE RISCO PARA QUEDA DE BLOCOS EM TALUDES RODOVIÁRIOS**

*Larissa Regina Costa Silveira*

Ouro Preto  
2022



**Larissa Regina Costa Silveira**

**PROPOSTA DE SISTEMA DE ANÁLISE DE  
RISCO PARA QUEDA DE BLOCOS EM  
TALUDES RODOVIÁRIOS**

Orientadora: Prof.<sup>a</sup> Dra. Milene Sabino Lana

Texto apresentado ao Programa de Pós-Graduação em Geotecnia do Núcleo de Geotecnia Escola de Minas da Universidade Federal de Ouro Preto, como parte integrante dos requisitos para obtenção do título de **Doutora em Geotecnia.**

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Milene Sabino Lana

**OURO PRETO/MG**  
ESCOLA DE MINAS / UFOP  
2022

## SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

S587p Silveira, Larissa Regina Costa.  
Proposta de sistema de análise de risco para queda de blocos em taludes rodoviários. [manuscrito] / Larissa Regina Costa Silveira. - 2022. 170 f.: il.: color., gráf., tab., mapa.

Orientadora: Profa. Dra. Milene Sabino Lana.

Tese (Doutorado). Universidade Federal de Ouro Preto. Núcleo de Geotecnia da Escola de Minas. Programa de Pós-Graduação em Geotecnia.

Área de Concentração: Geotecnia.

1. Blocos de concreto - Queda de blocos. 2. Taludes (Mecânica do solo) - Taludes rodoviários. 3. Estatística - Estatística multivariada. 4. Engenharia geotécnica - Risco geotécnico. I. Lana, Milene Sabino. II. Universidade Federal de Ouro Preto. III. Título.

CDU 624.13

Bibliotecário(a) Responsável: Maristela Sanches Lima Mesquita - CRB-1716



## FOLHA DE APROVAÇÃO

**Larissa Regina Costa Silveira**

### **Proposta de sistema de análise de risco para queda de blocos em taludes rodoviários**

Tese apresentada ao Programa de Pós-Graduação em Geotecnia do Núcleo de Geotecnia da Escola de Minas da Universidade Federal de Ouro Preto, como parte integrante dos requisitos para obtenção do título de **Doutor em Geotecnia**

Aprovada em 02 de setembro de 2022

#### Membros da banca

Profª. Drª. Milene Sabino Lana - Orientadora (Universidade Federal de Ouro Preto)  
Profª. Drª. Tatiana Barreto dos Santos - (Universidade Federal de Ouro Preto)  
Prof. Dr. Pedro Manuel Alameda Hernández - (Universidade Federal de Ouro Preto)  
Prof. Dr. André Monteiro Klen - (Instituto Federal de Minas Gerais)  
Prof. Dr. Paulo André Charbel (Instituto Federal de Goiás)

O Prof. Dr. Luis de Almeida Prado Bacellar, coordenador do Programa, aprovou a versão final e autorizou seu depósito no Repositório Institucional da UFOP em 07/10/2022



Documento assinado eletronicamente por **Luis de Almeida Prado Bacellar, COORDENADOR(A) DE CURSO DE PÓS-GRADUAÇÃO EM GEOTECNIA**, em 07/10/2022, às 15:21, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [http://sei.ufop.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0409372** e o código CRC **4D8204A4**.

# DEDICATÓRIA

Dedico essa tese a Júlio, Maria do Carmo, Juliano, Lílian, Alice, Ana Júlia, Luísa, Maria  
Eduarda e Marcela.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, à Nhá Chica e à Nossa Senhora Desatadora dos Nós que me ajudaram a “desenrolar” essa tese.

Agradeço as minhas maiores fontes de inspiração, que são meus pais e meus irmãos. Obrigada pelas oportunidades, por confiarem em mim e por me apoiarem incondicionalmente. Sem vocês eu jamais teria chegado até aqui.

Às minhas sobrinhas por serem minha alegria! Tudo é mais fácil quando me lembro daqueles sorrisos lindos.

Agradeço à Marcela, por todo amor ao longo desses anos e sobretudo pela paciência nesses em que estive envolvida com o desenvolvimento da tese. Foram muitos finais de semana em casa e muitas preocupações divididas. Obrigada pela compreensão!

A minha orientadora Milene pela disponibilidade, dicas sempre pertinentes e principalmente pelo conhecimento transmitido ao longo desses anos, além claro de toda a paciência que era visível desde as aulas na graduação. Hoje posso dizer que agradeço também pela amizade.

A todos os amigos e parceiros da “Firma da Milene”, Daniele, Davidson, Otávio, Juliano, Guilherme, Allan e Tatiana, obrigada pela parceria, diversão e troca de ideias. À Tatiana em especial, que foi um dos grandes suportes no desenvolvimento desse trabalho. Muito obrigada por toda ajuda e parceria de sempre. Esse trabalho teria sido mais difícil sem você.

A todos os professores e funcionários do NUGEO-UFOP, lugar onde não só cursei meu doutorado, como também tive a honra de ser funcionária.

Aos colegas da UFBA por toda a compreensão e apoio profissional para que eu pudesse concluir essa etapa.

A CAPES pelo apoio financeiro enquanto foi necessário.

## **RESUMO**

Tem sido uma tendência na engenharia geotécnica o tratamento de problemas típicos da área, como aqueles relacionados à estabilidade de escavações em geral, com uma abordagem voltada para análise e gestão de riscos, utilizando preferencialmente sistemas quantitativos. Esta pesquisa apresenta uma proposta de um sistema quantitativo para análise de risco de queda de blocos com foco em taludes rodoviários, onde a classificação de risco é dada por uma matriz que relaciona a probabilidade (*Likelihood*) de queda de blocos e a consequência desse evento. A consequência leva em consideração a chance do bloco atingir a rodovia (consequência A) e a chance de causar acidentes (consequência B). Tanto a probabilidade quanto a consequência A são calculadas por meio de equações, cujos valores são classificados em alto, médio (zona de transição) ou baixo de acordo com a região dos respectivos gráficos dos modelos em que esses valores são plotados. Essas equações e gráficos destes modelos foram gerados utilizando Análise de Agrupamento, Análise de Componentes Principais, Análise Discriminante e Elipses de Confiança, que são técnicas de estatística multivariada. O modelo de probabilidade gerado apresentou taxa de erro igual a 9,1%, enquanto a taxa de erro média quando o modelo foi testado em 31 diferentes amostras foi de 9,5%. O modelo de consequência A apresentou taxa de erro igual a 1,5%, enquanto a taxa de erro média quando o modelo foi testado em 31 diferentes amostras foi de 2,2%. O modelo de Consequência B é baseado em análise direta dos parâmetros, visto que o banco de dados reservado para a Consequência B não era adequado para aplicação de estatística multivariada e apenas dois parâmetros foram avaliados. O banco de dados geral utilizado nesta pesquisa possui mais de 200 taludes e as principais variáveis consideradas são os parâmetros tradicionalmente utilizados nas classificações de perigo e risco de taludes de rodovias.

**PALAVRAS-CHAVE:** Queda de blocos; Taludes rodoviários; Estatística multivariada; Risco geotécnico.

## **ABSTRACT**

It has been a trend in geotechnical engineering to manage typical problems, such as those related to the stability of excavations, with an approach focused on risk analysis and management, preferably using quantitative systems. This research presents a quantitative rockfall risk analysis system, focusing on highway rock slopes, where the risk classification is given by a matrix that relates the rockfall probability (Likelihood) and its consequences. The consequence takes into account the chance of the block reach the highway (consequence A) and the chance of causing accidents (consequence B). Both probability and consequence A are calculated using equations, whose values are classified as high, medium (transition zone) or low according to the respective model graphs. These equations and graphs of these models were generated using Cluster Analysis, Principal Component Analysis, Discriminant Analysis and Confidence Ellipses, which are multivariate statistical techniques. The generated probability model presented an error rate equal to 9.1%, while the average error rate when the model was tested on 31 different samples was 9.5%. Consequence model A presented an error rate equal to 1.5%, while the average error rate when the model was tested on 31 different samples was 2.2%. The Consequence B model is based on direct analysis of the parameters, since the database was not suitable for the application of multivariate statistics and only two parameters were evaluated. The database used in this research has more than 200 slopes and the main variables considered are the parameters traditionally used in the hazard and risk classification of highway slopes.

**KEYWORDS:** Rockfall; Highway rock slopes; Multivariate statistics; Geotechnical risk.



## Lista De Figuras

Figura 2.1 - Exemplo de sistema de gestão de risco (Adaptado por Santos (2019) de STANDARDS AUSTRALIA, 2004).....	9
Figura 2.2 - Causas básicas das quedas de blocos (Carvalho, 1991).....	10
Figura 2.3 - a) Trajetórias mais comuns de blocos em queda, de acordo com a geometria do talude (Ritchie, 1963). b) Impacto das irregularidades na superfície de um talude (Pierson et al., 2001).....	11
Figura 2.4- Carta de estabilidade do Q-Slope (Bar & Barton, 2017).....	14
Figura 2.5- a) Matriz de probabilidade; b) Matriz de perigo (Adaptado de Bauer & Neumann, 2011).....	17
Figura 2.6- Gráfico de perigo (Santos, 2019).....	23
Figura 2.7 – Matriz de risco para taludes de mineração (Santos, 2019).....	23
Figura 2.8 – Árvore de classificação (Santos, 2019).....	25
Figura 2.9 – Ábaco adaptado de Maerz et al. (2005) para determinar o escore referente à inclinação do talude pelo CRHRS modificado (adaptado de Santi et al., 2009).....	35
Figura 2.10 – Feições de lançamento, onde: a) Nenhuma; b) Menor; c) Maiores; d) Extremas (modificado de Russel et al., 2008).....	36
Figura 2.11– Ábaco de Ritchie (1963) para projetar as dimensões das áreas de captação de blocos.....	37
Figura 2.12 – Distribuição anual média dos ciclos de congelamento e descongelamento para o Estado do Colorado. O mapa foi derivado de Arnold et al. (1996) (Adaptado de Russel et al., 2008).....	39
Figura 2.13 – Exemplo de um talude cristalino que pontua 81 para o grau de descalçamento no CRHRS modificado (Russel et al., 2008).....	42
Figura 2.14– Exemplo de descalçamento em talude de rocha sedimentar (Russel et al., 2008).....	43
Figura 2.15 – Classes de reação de amostras de rochas submetidas ao Jar Slake (Russel et al., 2008).....	44
Figura 2.16– Exemplo de variação litológica entre camadas de um talude sedimentar (Russel et al., 2008).....	44
Figura 2.17– À esquerda: descontinuidades persistentes ( $> 3$ m), com uma direção de mergulho favorável em relação ao talude e à estrada (pontuação de 9). À direita:	

descontinuidades persistentes ( $> 3$ m) com uma direção de mergulho adversa em relação ao talude, de acordo com Santi et al. (2009).....	46
Figura 2.18 – Gráfico de calor (VAT) mostrando a tendência de agrupamento para formação de 2 grupos (Fonte: Data Novia).....	60
Figura 2.19– Gráfico para determinar o número ótimo de clusters pelo método do cotovelo.....	62
Figura 2.20 - Representação univariada de escores Z discriminantes (HAIR et al.).....	72
Figura 2.21 - Elipses de confiança e limites entre os grupos (Santos, 2019).....	74
Figura 2.22 - Exemplo do método bootstrap para uma amostra de 12 elementos. (Vassalo, 2017).....	77
Figura 3.1 – Fluxograma resumido da metodologia utilizada para construir o sistema de análise de risco.....	78
Figura 4.1– Matriz de correlações das variáveis relacionadas à probabilidade.....	101
Figura 4.2– Matriz de correlações das variáveis relacionadas à consequência A.....	101
Figura 4.3– Matriz de correlações das variáveis relacionadas à consequência B.....	102
Figura 4.4 – Gráfico de frequência mostrando a variabilidade das variáveis relacionadas à probabilidade.....	103
Figura 4.5 - Mapa de Calor das dissimilaridades para as variáveis relacionadas à probabilidade.....	104
Figura 4.6– Número ótimo de clusters pelo Método do Cotovelo para as variáveis relacionadas à probabilidade.....	105
Figura 4.7– Número ótimo de clusters pelo Método da Silhueta para as variáveis relacionadas à probabilidade.....	105
Figura 4.8– Número ótimo de clusters pela Regra da Maioria para as variáveis relacionadas à probabilidade.....	105
Figura 4.9 – a) Agrupamento gerado pelo R considerando dois grupos; b) Agrupamento gerado pelo R considerando três grupos.....	106
Figura 4.10 – Erro para as 31 diferentes divisões amostrais.....	116
Figura 4.11– Gráfico de probabilidade gerado pela Análise Discriminante, ainda sem os limites entre classes.....	117
Figura 4.12 – Elipses de confiança e limites entre as classes de probabilidade.....	120
Figura 4.13– Gráfico de probabilidade .....	120

Figura 4.14 – Número de indivíduos por escore para cada variável relacionada à consequência A.....	122
Figura 4.15 - Mapa de Calor para as variáveis relacionadas à consequência A.....	122
Figura 4.16– Número ótimo de clusters pelo método do cotovelo para as variáveis relacionadas à consequência A.....	123
Figura 4.17 – Número ótimo de clusters pelo método da silhueta para as variáveis relacionadas à consequência A.....	123
Figura 4.18– Número ótimo de clusters pela regra da maioria para as variáveis relacionadas à consequência A.....	124
Figura 4.19 – Agrupamento gerado pelo R para as variáveis relacionadas à consequência A.....	125
Figura 4.20 – Erro para os 31 diferentes modelos gerados para cada uma das 31 divisões amostrais para a Consequência A.....	130
Figura 4.21– Gráfico de chance do bloco atingir a rodovia (Conequência A) gerado pela Análise Discriminante, ainda sem os limites entre classes.....	131
Figura 4.22– Elipses de confiança e limites entre as classes de consequência A.....	134
Figura 4.23– Gráfico para Consequência A.....	134
Figura 4.24 - Variabilidade das variáveis relacionadas à consequência B.....	137
Figura 4.25 – Matriz de classificação para consequência B.....	139
Figura 4.26 – Matriz de Consequências.....	140
Figura 4.27 – Matriz de Risco.....	141

## Lista de Tabelas

Tabela 2.1 - Descrição de classes e resultados do SMR (Romana, 1985).....	13
Tabela 2.2- parâmetros geomecânicos observados para a determinação da suscetibilidade a queda de blocos (Adaptado de Bauer & Neumann,2011).....	15
Tabela 2.3- Influências externas (Bauer & Neumann, 2011).....	16
Tabela 2.4- Pesos atribuídos aos parâmetros para a metodologia adaptada por Silveira (2017).....	18
Tabela 2.5- Caracterização dos parâmetros adotados nos sistemas de avaliação de risco e de perigo proposto por Santos (2019).....	20
Tabela 2.6– RHRS original para o estado do Oregon (EUA) (Pierson & Van Vickle, 1993).....	27
Tabela 2.7– RHRS original para o estado do Colorado (EUA) (Adaptado de Andrew, 1994).....	30
Tabela 2.8– CRHRS modificado (Adaptado de Santi et al., 2009).....	32
Tabela 2.9– Eficácia da área de captação quanto a forma da vala (Adaptado de Santi et al., 2009).....	37
Tabela 2.10– Distância de visibilidade requerida de acordo com os limites de velocidade da via (Pierson e Van Vickle, 1993).....	49
Tabela 2.11- Classificação dos valores do coeficiente de correlação linear de Pearson (Callegari-Jaques, 2003)....	52
Tabela 3.1– Pesos atribuídos às variáveis no CRHRS e no trabalho proposto.....	83
Tabela 4.1– Parâmetros e suas devidas pontuações para análise de probabilidade.....	92
Tabela 4.2 – Comparação entre o que é adotado no CRHRS modificado (Santi <i>et al.</i> , 2009) com o que é adotado nessa proposta.....	99
Tabela 4.3 – Parâmetros e suas devidas pontuações para análise da consequência A.....	99
Tabela 4.4 – Parâmetros e suas devidas pontuações para análise da consequência B.....	100
Tabela 4.5 – Resultados dos testes de esfericidade de Bartlett.....	102
Tabela 4.6 – Taludes medóides de cada grupo referente à probabilidade.....	107
Tabela 4.7– Resultado da Análise de Agrupamento para a probabilidade gerado pelo algoritmo R.....	108
Tabela 4.8– Quantidade de taludes por grupo para taludes que somam entre 17 e 22.....	108

Tabela 4.9 – Classes de probabilidade obtidas após análise final do agrupamento gerado pelo algoritmo R.....	110
Tabela 4.10– Resultado da Análise de Agrupamento para a probabilidade para k = 3.....	110
Tabela 4.11– Quantidade de taludes por grupo para taludes que somam entre 15 e 21 para k = 3.....	111
Tabela 4.12– Variância explicada pelas componentes principais geradas na análise para probabilidade.....	112
Tabela 4.13 – Valores de média e desvio padrão para cada variável relacionada à probabilidade.....	113
Tabela 4.14 – Exemplo de cálculo dos escores das componentes principais para um talude do banco de dados referente à probabilidade.....	114
Tabela 4.15 – Teste M Box e Royston para o banco de dados referente ao modelo de probabilidade.....	115
Tabela 4.16 – Resultado da validação da regra de classificação para a probabilidade.....	118
Tabela 4.17 – Exemplo de classificação quanto à probabilidade para um talude qualquer....	121
Tabela 4.18 – Medóides de cada grupo referente à consequência A.....	125
Tabela 4.19– Resultado da Análise de Agrupamento para a consequência A gerado pelo R.....	126
Tabela 4.20– Classes para consequência A obtidas após análise final do agrupamento gerado pelo R.....	126
Tabela 4.21 – Quantidade de taludes por grupo para taludes que somam 10, 11, 12 e 13.....	126
Tabela 4.22– Variância explicada pelas componentes principais geradas na análise para consequência A.....	128
Tabela 4.23 – Valores de média e desvio padrão para cada variável relacionada à consequência A.....	129
Tabela 4.24 – Exemplo de cálculo dos escores das componentes principais para um talude do banco de dados referente à consequência A.....	129
Tabela 4.25– Teste M Box e Royston para o banco de dados referente ao modelo de consequência A.....	130
Tabela 4.26 – Resultado da validação da regra de classificação para a consequência A.....	132

Tabela 4.27 – Exemplo de classificação quanto à consequência A para um talude qualquer.....	135
Tabela 4.28– Classes para consequência B obtidas após análise dos parâmetros.....	139

## **Lista de Símbolos, Nomenclatura e Abreviações**

### SÍMBOLOS E SIGLAS GERAIS

ACP – Análise de Componentes Principais

AGNES – Agglomerative Analysis: análise aglomerativa

CART – “Classification and Regression Trees”: árvores de classificação e regressão

CDOT - “Colorado Department of Transportation”: Departamento de Trânsito do Colorado.

CLARA - Clustering Large Applications: agrupamento para grandes aplicações

CRHRS – “Colorado Rockfall Hazard Rating System” – sistema de classificação de perigo de queda de blocos do Colorado

DIANA – Divisive Analysis: análise divisiva

Dm - profundidade medida da vala

Dr - profundidade recomendada pelo Ábaco de Ritchie (1963) para a vala

EXT – extensão do talude

FMD – fluxo médio diário

H – altura do talude (m)

IAT – índice referente à altura do talude

ISSMGE - Sociedade Internacional de Mecânica dos Solos e Engenharia Geotécnica

IUGS – “International Union of Geological Sciences”: União Internacional de Ciências Geológicas

Ja - índice relacionado à alteração e preenchimento das descontinuidades

Jn - índice relacionado ao número de famílias de descontinuidades

Jr - índice relacionado à rugosidade das descontinuidades

Jwice - índice relacionado às condições climáticas/ambientais

LV – limite de velocidade

MAH - métodos de agrupamento hierárquicos

MANH - métodos de agrupamento não hierárquicos

NYSDOT – “New York State Department of Transportation”: Departamento de Trânsito do Estado de Nova Iorque.

PAM – “Partitioning Around Medoids”: particionando em torno de medoides

PDV – percentual de distância de visibilidade

RHRS – “Rockfall Hazard Rating System” – sistema de classificação de perigo de queda de blocos

RMR – “Rock Mass Rating”: classificação do maciço rochoso

RMR<sub>b</sub> – “Basic Rock Mass Rating”: RMR básico

RMV – risco médio ao veículo

RQD - “Rock Quality Designation”: designação de qualidade da rocha

SMR – “Slope Mass Rating”: classificação de taludes

SRFslope – “Strenght Reduction Fator”: índice relacionado às tensões no maciço, efeito do desmonte e/ou presença de descontinuidades importantes, como falhas persistentes

VAT - “Visual Assessment of Tendency”: método visual do gráfico de calor

Wi – grau de alteração da rocha de acordo com ISRM,  $i = 1, 2, \dots, 5$

Wm - largura medida da vala;

Wr - largura recomendada pelo Ábaco de Ritchie (1963) para a vala.

WWS - within sum of square: soma dos quadrados das distâncias

## SIMBOLOGIA PARA ESTATÍSTICA UNIVARIADA E MULTIVARIADA

df: graus de liberdade

**I**: matriz identidade

**R**: matriz de correlação amostral

**S**: matriz de covariâncias amostral

**W**: estatística de Shapiro-Wilk

**X**: vetor aleatório multivariado

$X^2$ : distribuição qui-quadrado

$z$  - variável normal padronizada de  $W$

$Z_i$  – variável  $i$  padronizada

$\mu_i$  – média aritmética univariada da variável  $i$

$\sigma_i$  – desvio padrão para a variável  $i$

$\rho_{ij}$  – coeficiente de correlação linear de Pearson entre as variáveis  $i$  e  $j$

$\rho_s$  – coeficiente de correlação por postos de Spearman

$\sigma_{ij}$  - covariância entre as variáveis  $i$  e  $j$  em estatística multivariada

$\sigma_{ii}$  - variância da variável  $i$  em estatística multivariada

$\sigma_{jj}$  - variância da variável  $j$  em estatística multivariada

$v_i$  - dados de uma variável, organizados em ordem crescente de magnitude



$\bar{v}$ : média de  $v_i$

$a_i$  - coeficientes lineares não enviesados de Sarhan & Greenberg (1956)

$\lambda$  - coeficiente estimado por 50 amostras e suavizadas utilizando o polinômio  $\ln(n) - d$ , sendo  $d = 3$  para  $7 \leq n \leq 20$  e  $d = 5$  para  $21 \leq n \leq 2000$

$\bar{x}_y$  - média amostral

$s_y$  - desvio padrão amostral

$W_j$  - estatística de Shapiro-Wilk para a variável  $j$

$R_j$  - estatística do teste de normalidade multivariada

$e$  - graus de liberdade equivalentes

$\Sigma$  - matriz de covariâncias populacional

$k$  - número de populações ou de grupos

$M$  - estatística  $M$  de Box

$S_j$  - matriz de variâncias e covariâncias amostral do grupo  $j$

$D$  - matriz de dissimilaridades

$H$  - estatística de Hopkins

$Y$  - conjunto original de dados

$U$  - conjunto de  $p$  pontos gerados aleatoriamente no espaço

$Q$  - conjunto de  $p$  pontos reais de  $Y$

$q_i$  - é a distância entre o  $i$ -ésimo ponto real pertencente a  $W$  e seu vizinho mais próximo em  $Y$ .

$u_i$  - é a distância do  $i$ -ésimo ponto aleatoriamente gerado em  $U$  a seu vizinho mais próximo em  $Y$ .

$\delta_{ij}$  - coeficiente de dissimilaridade

$d_{ij}$  - distância entre as observações  $i$  e  $j$

$\tilde{x}_i$  - vetor de observações do indivíduo  $i$

$x_{ij}$  - valor da variável  $j$  no indivíduo  $i$

$V^{-2}$  - matriz diagonal dos recíprocos das variâncias

$\Sigma^{-1}$  - inversa da matriz de covariâncias populacional

$s_{ij}$  - coeficiente de similaridade

$P_{p \times p}$  - matriz de covariâncias das variáveis padronizadas

$\lambda_p$  - autovalores ordenados para  $p$  variáveis

$e_p$  - autovetores normalizados da matriz de correlações para  $p$  variáveis

$Y_i$  -  $i$ -ésima componente principal

$Var[Y_i]$  - variância da  $i$ -ésima componente principal

$\lambda_i$  – variância da i-ésima componente principal

$\hat{e}_j^t$  - j-ésimo autovetor estimado correspondente ao j-ésimo maior autovalor da matriz  $\mathbf{W}^{-1}\mathbf{B}$

$\mathbf{W}$  - matriz de soma de quadrados e produtos cruzados dentro dos grupos

$\mathbf{B}$  - matriz de soma de quadrados e produtos cruzados entre os grupos

$x_{ib}$  - vetor de observações do elemento amostral b que pertence à população i;

$\bar{\mathbf{x}}_i$  - vetor de médias amostral da população i;

$\bar{\mathbf{x}}$  - vetor de médias amostral, considerando todas as n observações.

$\hat{Y}_j$  – escores das funções discriminantes

$\hat{Y}_i$  - escores das funções discriminantes aplicadas aos vetores de médias amostrais observados para cada população

$\boldsymbol{\mu}$  é a média populacional da população i

$\alpha$  é o nível de significância

# Sumário

<b>CAPÍTULO 1.....</b>	<b>1</b>
<b>INTRODUÇÃO .....</b>	<b>1</b>
<b>1.1. CONSIDERAÇÕES INICIAIS .....</b>	<b>1</b>
<b>1.2. OBJETIVOS.....</b>	<b>3</b>
<b>1.3. JUSTIFICATIVA E RELEVÂNCIA .....</b>	<b>3</b>
<b>CAPÍTULO 2.....</b>	<b>6</b>
<b>REVISÃO BIBLIOGRÁFICA.....</b>	<b>6</b>
<b>2.1. CONCEITOS BÁSICOS SOBRE RISCO GEOTÉCNICO.....</b>	<b>6</b>
<b>2.2. QUEDAS DE BLOCOS .....</b>	<b>10</b>
<b>2.3. AVALIAÇÃO DE ESTABILIDADE, PERIGO E RISCO PARA TALUDES.....</b>	<b>12</b>
<b>2.4. AVALIAÇÃO DE PERIGO E RISCO DE QUEDA DE BLOCOS PARA TALUDES     RODOVIÁRIOS.....</b>	<b>25</b>
<b>2.4.1. CRHRS Modificado .....</b>	<b>31</b>
<b>2.5. TÉCNICAS DE ESTATÍSTICA MULTIVARIADA .....</b>	<b>50</b>
<b>2.5.1. Conceitos iniciais e viabilidade de aplicação das técnicas .....</b>	<b>50</b>
<b>2.5.2. Análise de Agrupamento .....</b>	<b>57</b>
<b>2.5.3. Análise de Componentes Principais.....</b>	<b>68</b>
<b>2.5.4. Análise Discriminante .....</b>	<b>71</b>
<b>2.5.5. Elipses de Confiança .....</b>	<b>74</b>
<b>CAPÍTULO 3.....</b>	<b>78</b>
<b>METODOLOGIA .....</b>	<b>78</b>
<b>3.1. ESTUDO PRELIMINAR DO BANCO DE DADOS .....</b>	<b>78</b>
<b>3.2. ANÁLISE E SELEÇÃO DE VARIÁVEIS .....</b>	<b>81</b>
<b>3.3. VIABILIDADE DE APLICAÇÃO DAS TÉCNICAS DE ESTATÍSTICA     MULTIVARIADA .....</b>	<b>83</b>
<b>3.4. ANÁLISE DE AGRUPAMENTO .....</b>	<b>84</b>
<b>3.5. ANÁLISE DE COMPONENTES PRINCIPAIS .....</b>	<b>86</b>
<b>3.6. ANÁLISE DISCRIMINANTE E CRIAÇÃO DO MODELO .....</b>	<b>86</b>
<b>3.7. ELIPSES DE CONFIANÇA E FINALIZAÇÃO DO MODELO .....</b>	<b>88</b>
<b>3.8. CONSTRUÇÃO DO SISTEMA DE ANÁLISE DE RISCO .....</b>	<b>88</b>
<b>CAPÍTULO 4.....</b>	<b>89</b>
<b>RESULTADOS E DISCUSSÕES .....</b>	<b>89</b>
<b>4.1. SELEÇÃO DE VARIÁVEIS E PREPARAÇÃO DOS BANCOS DE DADOS .....</b>	<b>89</b>

4.1.1.	Probabilidade.....	89
4.1.2.	Consequências.....	93
4.1.3.	Consequência A .....	95
4.1.4.	Consequência B .....	100
4.2.	TESTES ESTATÍSTICOS PRELIMINARES .....	101
4.3.	MODELO DE PROBABILIDADE DE QUEDA DE BLOCOS .....	102
4.3.1.	Análise de Agrupamento .....	102
4.3.2.	Análise de Componentes Principais.....	112
4.3.3.	Análise Discriminante .....	114
4.3.4.	Elipses de confiança e gráfico final .....	119
4.4.	MODELO PARA CONSEQUÊNCIA A .....	121
4.4.1.	Análise de Agrupamento .....	121
4.4.2.	Análise de Componentes Principais.....	128
4.4.3.	Análise Discriminante .....	130
4.4.4.	Elipses de confiança e gráfico final .....	133
4.5.	MODELO PARA CONSEQUÊNCIA B .....	136
4.6.	SISTEMA DE ANÁLISE DE RISCO POR MATRIZ DE RISCO .....	140
CAPÍTULO 5.....		144
CONCLUSÕES .....		144
REFERÊNCIAS .....		146
APÊNDICE.....		156

# **CAPÍTULO 1**

---

## **INTRODUÇÃO**

### **1.1. CONSIDERAÇÕES INICIAIS**

Qualquer atividade humana, por mais simples que seja, está sujeita a riscos, sendo que na engenharia, esses riscos geralmente estão relacionados a acidentes ou falhas que podem ter como consequência a geração de custos indevidos ou não planejados, danos à integridade física de instalações e pessoas e, no pior cenário, perda de vidas.

Diante disso, tem havido uma tendência na engenharia geotécnica de se começar a tratar das situações típicas da área, como desastres naturais e outros problemas relacionados à estabilidade de escavações em geral, com uma abordagem voltada para análise e gestão de riscos. No entanto, nem sempre os conceitos são utilizados de forma correta, o que pode levar a más interpretações em projetos e pesquisas ligados ao assunto e, conseqüentemente, perder-se muito tempo até encontrar uma solução adequada para determinado problema.

Fell *et al.* (2008), comentam que o risco é uma medida da probabilidade de um evento ocorrer e de sua gravidade. Com isso, o risco pode ser estimado como o produto “probabilidade x consequência”, sendo que a consequência, na engenharia geotécnica, é o resultado decorrente de um movimento de massa e pode ser expressa em termos de perdas em geral, danos ou ferimentos. Ainda de acordo com Fell *et al.* (2008), a análise de risco consiste no uso da informação disponível para estimar o risco à população, bens ou o ambiente, a partir do perigo. Geralmente contêm os seguintes passos: definição do escopo, identificação do perigo, e estimativa do risco.

Quando se trata de estabilidade de taludes, o primeiro passo para uma análise de risco é quantificar a probabilidade de ocorrências de rupturas e também o grau de exposição a condições perigosas através de parâmetros observados em campo, no caso, as características geotécnicas do maciço que compõe o talude, a geometria desse talude, a presença ou não de algum elemento de contenção de fragmentos rompidos, ou se há distância suficiente para que

esses fragmentos possam se depositar sem atingir o elemento exposto, por exemplo, rodovias, casas, pessoas e também vias em minerações.

De acordo com McMillan & Matheson (1998), a utilização de técnicas de desmonte não controlado para corte de taludes rodoviários deixou um legado de instabilidade nesses taludes, o que pode colocar os usuários das rodovias em situações iminentes de perigo. Isso pode ser agravado pelo aumento da malha rodoviária e, conseqüentemente, do número de usuários.

Para que haja gestão eficiente do risco, é necessário que antes esse risco seja quantificado de forma correta e detalhada, que leve em consideração todos os parâmetros que podem levar a rupturas, como tombamentos, quedas e escorregamentos. Afinal, obras de contenção são onerosas e sua aplicação deve ser orientada por análises de risco quantitativas devidamente justificadas. Por outro lado, reparos de danos em rodovias decorrentes de movimentos de massa também são caros e geram transtornos no trânsito local. Além do risco aos usuários que podem se ferir ou mesmo perder a vida, o que torna ainda mais necessário o devido controle de taludes rodoviários.

Esse trabalho tem por objetivo o desenvolvimento de uma metodologia quantitativa para análise de risco de queda de blocos em taludes rochosos rodoviários, onde a classificação de risco é dada por uma matriz que relaciona a probabilidade, ou *Likelihood*, de ocorrer queda de blocos com a consequência desse evento, onde essa consequência leva em consideração a chance do bloco atingir a rodovia, o que por si só já gera gastos e transtornos; e também a chance de causar acidentes, que é algo mais crítico. Tanto a probabilidade quanto as consequências são calculadas através de equações, cujos valores são classificados como altos, médios (zona de transição), ou baixos de acordo com a região dos respectivos gráficos gerados pelo modelo proposto.

As equações, gráficos e classificações de probabilidade de queda de blocos podem ser utilizadas para outros tipos de taludes, como de mineração ou urbanos. Já o modelo para estimar as consequências é direcionado para rodovias, embora possa ser adaptado, ou embasar uma análise de consequências em outros contextos. Todo o modelo foi gerado através de técnicas de estatística multivariada aplicadas a um banco de dados com mais de 200 taludes rochosos rodoviários, naturais e de corte, do Estado do Colorado (EUA), que foram utilizados

para o desenvolvimento do Colorado Rockfall Hazard Rating System Modificado (CRHRS Modificado) (Santi *et al.*, 2009).

É importante deixar claro que para a metodologia proposta nesse trabalho, dentre todas as variáveis presentes no banco de dados original utilizado por Santi *et al.* (2009), foram selecionadas apenas aquelas cujas características independem de fatores sazonais regionais, visto que a intenção é criar uma metodologia prática e aplicável em rodovias de qualquer local.

## **1.2. OBJETIVOS**

O objetivo geral da tese é propor um sistema quantitativo de classificação de risco de quedas de blocos para taludes rodoviários, sendo que neste trabalho, o objetivo principal é gerar esse sistema para taludes em rochas cristalinas. Como objetivos específicos, pode-se mencionar:

- Analisar preliminarmente o banco de dados completo e selecionar as variáveis geotécnicas que descrevem o fenômeno de queda de blocos em taludes rochosos, ou seja, selecionar aqueles parâmetros que influenciam na probabilidade à queda.
- Realizar o mesmo tipo de análise do banco de dados para verificar quais parâmetros se relacionam com as chances do bloco atingir a rodovia e também aqueles que têm relação com a probabilidade de acontecer acidentes. Ou seja, verificar as variáveis que serão inseridas na análise de consequências.
- Construir os bancos de dados e prepará-los para aplicação das técnicas estatísticas.
- Aplicar técnicas de estatística multivariada. Propor equações e gráficos de probabilidade, ou *Likelihood*, e de consequência para, por fim, obter a classificação de risco por meio de uma matriz de risco para taludes rodoviários em rochas cristalinas.

## **1.3. JUSTIFICATIVA E RELEVÂNCIA**

No Brasil, o sistema rodoviário ainda é o principal tipo de via de transporte, com cada vez mais usuários, se sobressaindo consideravelmente em relação às ferrovias e hidrovias. Para a construção da rodovia é geralmente necessário, em vários pontos da mesma, se fazer cortes nos maciços rochosos, que sem o devido controle, podem causar sérios problemas de

instabilidade e colocar aqueles que ali trafegam expostos a diversas situações de perigo. Além do mais, muitos trechos podem estar no pé de taludes rochosos naturais, porém instáveis.

Alguns autores propuseram sistemas de classificações de perigo para rodovias e que, portanto, não quantificam as consequências, como é o caso do Rockfall Hazard Rating System – RHRS (Pierson et al., 1990) e algumas de suas derivações; já outras derivações começam a considerar as consequências, como o Colorado Rockfall Hazard Rating System e Ontario Rockfall Hazard Rating System. Embora sejam ferramentas muito úteis, esses sistemas não atribuem pesos para as variáveis, o que leva o usuário a entender que todas elas contribuem igualmente para o fenômeno de queda de blocos. McMillan & Matheson (1998) também propuseram uma abordagem de análise de perigo de queda de blocos em taludes rodoviários através de dados de campo, onde a consequência do evento não é contabilizada.

A queda de blocos é um fenômeno rápido e difícil de prever, que pode causar diversos danos ao elemento exposto. Sendo assim, é importante que se tenha um sistema de análise de risco onde cada parâmetro observado na análise tenha atribuído a si um peso adequado, para que os erros de subestimação ou superestimação de risco sejam minimizados e também quantificados. Isso é possível aliando a geotecnia com a estatística, de forma a analisar situações diferentes, aumentando a confiabilidade das análises feitas.

Como rodovias são grandes obras lineares e muitas vezes com muitos taludes rochosos em seu entorno, é necessário que esse sistema de avaliação de risco de movimentos de massa seja simples e prático, a fim de viabilizar o trabalho do engenheiro tanto com relação ao tempo de conclusão, quanto financeiramente. Caso contrário, os trabalhos de campo seriam muito desgastantes devido à extensão das áreas, o que acarretaria erros. Também poderiam se tornar mais onerosos e mesmo inviáveis de se aplicar devido ao nível detalhe requerido nos levantamentos para aplicação de metodologias complexas.

É importante deixar claro que o objetivo da metodologia de análise de risco proposta nesse trabalho não é minimizar a relevância das já consagradas análises de estabilidade de taludes por métodos de equilíbrio limite ou por métodos numéricos, afinal, embora considerem alguns parâmetros em comum, são metodologias diferentes, com abordagens e respostas diferentes. No entanto, essas análises de estabilidade por equilíbrio limite e por métodos numéricos



requerem um minucioso levantamento de campo, ensaios de laboratório e licença de softwares específicos, o que pode tornar esse tipo de análise mais demorada e cara, sobretudo quando se tem um grande número de taludes para se avaliar.

Com isso, através de análises de risco bem sucedidas, práticas e confiáveis, é possível estabelecer rapidamente os locais com real urgência de intervenção a fim de projetar o tipo de estrutura de contenção e controle mais adequada para cada tipo de cenário, evitando gastos em locais menos críticos e, principalmente, evitando que graves acidentes ocorram devido a rupturas que poderiam ter sido previstas e gerenciadas.

# CAPÍTULO 2

---

## REVISÃO BIBLIOGRÁFICA

### 2.1. CONCEITOS BÁSICOS SOBRE RISCO GEOTÉCNICO

As análises de risco vêm ganhando destaque no meio geotécnico justamente por tratar do problema tanto no âmbito da probabilidade de ocorrência, quanto das consequências que esse evento pode trazer, sendo portanto, uma análise completa do problema. De acordo com Fell *et al.* (2008), a maioria dos estudos relacionados a risco, sobretudo quando se trata de zoneamento de áreas de risco são qualitativos. A depender da fase e da finalidade do estudo, análises qualitativas são relevantes e é o suficiente. No entanto, a depender do nível de detalhamento requerido e da complexidade do evento estudado, é importante encontrar maneiras de quantificar esse risco.

Outro problema que ainda é comum nos estudos de risco geotécnico, conforme pontuado por Cascini *et al.* (2005), é a falta de uniformidade na terminologia referente ao assunto. Com isso, pode-se ter estudos relacionados a risco geotécnico onde na verdade foi verificada apenas a suscetibilidade de ocorrência do evento, sem qualquer menção às consequências. Se não houver uniformidade na terminologia, não haverá também precisão quanto aos resultados dos diversos estudos zoneamento de áreas de risco, como foi observado por Santos (2019) com base em Cascini *et al.* (2005).

Na tentativa de diminuir esse tipo de problema e homogeneizar o entendimento de profissionais que atuam no ramo da análise de risco geológico-geotécnico, Fell *et al.* (2008), diferenciaram os principais conceitos utilizados pelos geotécnicos, como risco, probabilidade, *likelihood*, perigo, vulnerabilidade e consequência. As definições colocadas a seguir, apresentadas por Fell *et al.* (2008) são baseadas nas definições adotadas pela *International Union of Geological Sciences* (IUGS) (1997) e adotadas pela Sociedade Internacional de Mecânica dos Solos e Engenharia Geotécnica (ISSMGE):

- Suscetibilidade (*Susceptibility*): avaliação quantitativa ou qualitativa da classificação, volume (ou área) e distribuição espacial de escorregamentos em geral que existem ou

podem ocorrer em uma área. Está mais relacionada às características da área, como por exemplo, as condições geotécnicas de um talude, sem necessariamente apresentar qualquer relação temporal.

- Perigo (*Hazard*): uma condição com potencial para causar uma consequência indesejável. Deve conter a probabilidade de um determinado evento acontecer, a sua magnitude, dentro de um período de tempo.
- Probabilidade: Medida do grau de certeza de ocorrência de um evento, que deve estar entre 0 (impossibilidade de ocorrência) e 1 (certeza de ocorrência). É uma estimativa da probabilidade de ocorrência do evento futuro incerto. A probabilidade pode ser dada por meio de um “grau de crença”, ou seja, uma medida quantificada de crença, ou confiança na *likelihood* de um resultado, obtida considerando todas as informações disponíveis e confiáveis.
- *Likelihood*: termo utilizado para a descrição qualitativa da probabilidade.
- Vulnerabilidade (*Vulnerability*): grau de perda para um determinado elemento ou conjunto de elementos dentro de uma área de perigo (*hazard*). É expressa em uma escala de 0 (sem perda) a 1 (perda total). Pode ser relacionado tanto a propriedades quanto a vidas. Relacionada à exposição a um determinado evento.
- Consequência (*Consequence*): resultados ou resultados potenciais decorrentes da ocorrência de um evento expressos qualitativa ou quantitativamente, em termos de perda, desvantagem ou ganho, dano, lesão ou perda de vida.
- Risco (*Risk*): é uma medida da probabilidade e gravidade de um determinado evento. Geralmente é expresso pelo produto da probabilidade do evento/fenômeno acontecer pela consequência, caso ocorra. Uma interpretação mais geral do risco envolve uma comparação da probabilidade e consequências em uma forma diferente desse produto.
- Análise de Risco (*Risk Analysis*): uso de informações disponíveis para estimar o risco para indivíduos, população, propriedade ou meio ambiente, diante de uma situação de perigo (*hazard*). As análises de risco geralmente contêm as seguintes etapas: definição do escopo, identificação de perigos e estimativa de risco.

O principal motivo para fazer uma análise de risco é justamente para avaliar e posteriormente implantar medidas de mitigação nos locais apontados como problemáticos, sendo esse o conceito de gestão de riscos. Um exemplo de um sistema de gestão de riscos, com os devidos

passos a serem seguidos é apresentado na Figura 2.1 e foi proposto pela STANDARDS AUSTRALIA (2004).

Nesse tipo de sistema, o primeiro passo é identificar o risco em um determinado contexto, ter ciência do que pode acontecer, quando e onde para posteriormente analisar. Feita essa identificação, parte-se para a análise de risco, isto é, se determina a probabilidade de ocorrência de um dado evento e as suas consequências.

O segundo passo se trata de fazer uma avaliação do risco. É nessa etapa que se estabelece, por exemplo, as regiões prioritárias para receber algum tipo de tratamento, ou estudo mais aprofundado, levando-se em consideração o que seria uma situação de risco “tolerável” ou “não tolerável”, mais crítica. A escolha do tipo de tratamento, bem como a implementação e a avaliação final pós-mitigação são partes do terceiro passo, que é o gerenciamento de risco. Ressalta-se que monitoramento e comunicação deve fazer parte de todas as etapas.

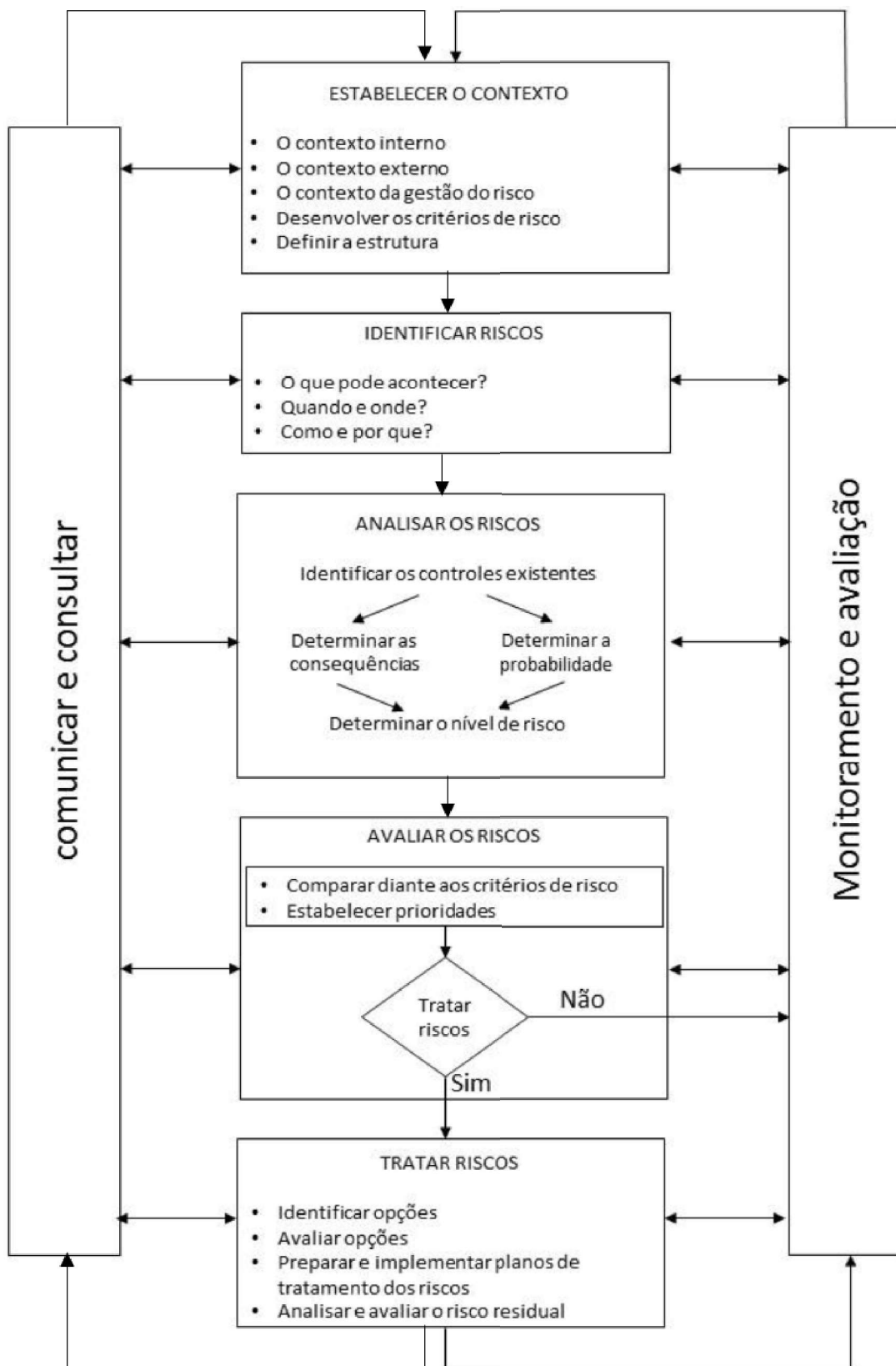


Figura 2.1 - Exemplo de sistema de gestão de risco (Adaptado por Santos (2019) de STANDARDS AUSTRALIA, 2004).

## 2.2. QUEDAS DE BLOCOS

As quedas de bloco (*rockfall*) correspondem ao destacamento de uma massa de rocha de uma encosta íngreme ou escarpada (Giani, 1992). Esse movimento de massa geralmente é rápido e difícil de prever. Existem dois grandes grupos de causas básicas que podem desencadear a queda de blocos, de acordo com Carvalho (1991). A primeira está mais relacionada ao fraturamento do maciço rochoso que acaba por formar blocos instáveis. Esses, por sua vez, ao cair, descalçam outros, os deixando salientes e instáveis (Figura 2.2a).

A segunda causa básica é mais comum em rochas sedimentares, que podem apresentar diferentes camadas intercaladas entre si. Nesse caso, se uma camada for mais intemperizável do que outra, o maciço se torna propenso à erosão diferencial. Com isso, uma camada mais alterada pode ser erodida, descalçando uma camada superior menos alterada. Essa camada saliente pode ter também a presença de descontinuidades, o que forma blocos instáveis (Figura 2.2b). Nota-se que apesar de os maciços sedimentares apresentarem uma causa básica diferente para a queda de blocos, o mecanismo se dá da mesma forma que para os demais maciços: presença de descontinuidades aliada a saliências e descalçamento de blocos.

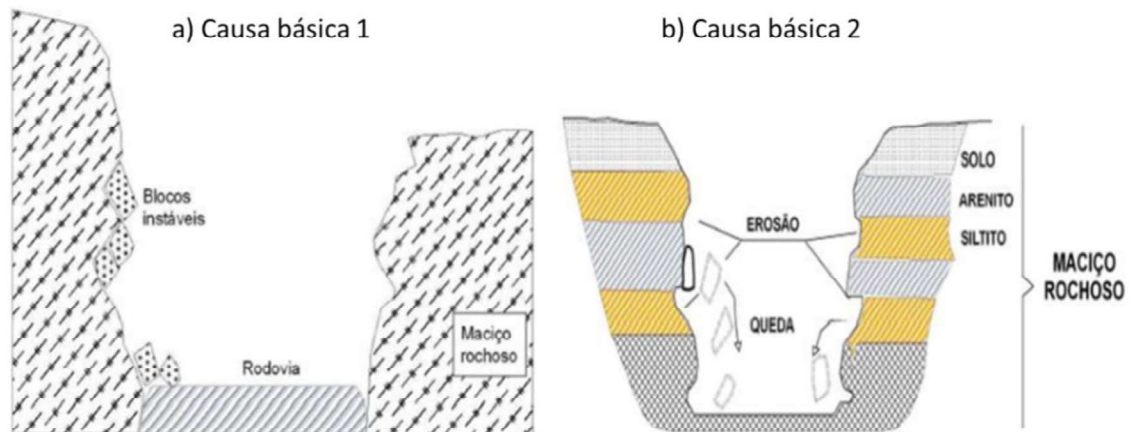


Figura 2.2 - Causas básicas das quedas de blocos (Carvalho, 1991).

Ao se falar em quedas de blocos, é importante se ter em mente, além dos fatores desencadeantes do movimento, os parâmetros que influenciam no alcance que esse bloco em queda irá tomar, pois o destino final desse bloco está intimamente relacionado com o grau de exposição em que pessoas, automóveis, equipamentos, ou qualquer elemento importante para

um determinado contexto qualquer, está inserido, sendo que o nível de danos do evento se relaciona com esse grau de exposição.

Quanto mais elevada é a posição de um bloco em um talude, maior é a sua energia potencial e, em caso de queda, maior deve ser a sua energia cinética, conseqüentemente, maior será a distância alcançada pelo bloco até parar. A inclinação do talude também influencia no movimento desse bloco e conseqüentemente no alcance. Isso porque se o talude possui inclinações entre cerca de 30° a 60°, o bloco tende a rolar e/ou saltar, o que pode fazer com que ele pare em pontos mais distantes do pé do talude. Já para taludes verticais, a depender da energia do bloco, espera-se que ele pare mais próximo do pé do talude (Figura 2.3a).

A presença de saliências e cicatrizes na superfície do talude pode fazer com que o bloco se comporte como um projétil e atinja distâncias ainda maiores, com trajetórias cada vez mais difíceis de prever. A Figura 2.3b ilustra essa situação ao mostrar as diferentes trajetórias A, B, C e D que um bloco poderia tomar de uma forma geral.

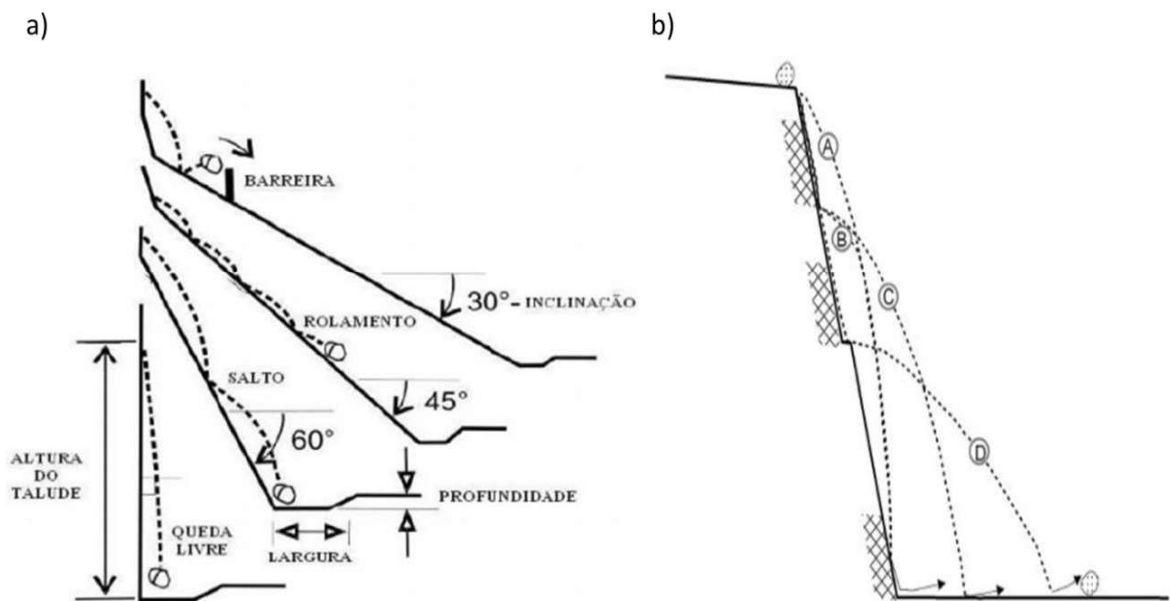


Figura 2.3 - a) Trajetórias mais comuns de blocos em queda, de acordo com a geometria do talude (Ritchie, 1963). b) Impacto das irregularidades na superfície de um talude (Pierson *et al.*, 2001).

Em A, o bloco cai em queda livre, sem colidir com a superfície do talude; com isso é mais provável que o bloco estacione nas proximidades do pé do talude. Em B e C são mostradas

duas diferentes situações que ilustram o um possível comportamento do bloco ao colidir com a superfície do talude. Caso essa superfície seja menos irregular, pode ser que a distância alcançada por ele seja menor (B); no entanto, se a colisão for entre o bloco e alguma irregularidade no talude, pode ser que ele atinja maiores distâncias (C). Em D é mostrada uma clara situação de colisão entre blocos e cicatrizes no talude, ou cristas de bancadas.

A forma do bloco e as características elásticas da rocha e da superfície de impacto também interferem no alcance, pois se os materiais que compõem o bloco e a superfície de impacto absorverem grande parte da energia mecânica durante a colisão, o bloco não continuará em movimento por muito mais tempo e não alcançará grandes distâncias. Nessas condições, entende-se que esses materiais possuem baixos coeficientes de restituição. Quanto à forma, blocos arredondados tendem a rolar e saltar, enquanto que blocos mais tabulares perdem grande parte da energia devido ao atrito entre sua maior dimensão e a superfície de impacto.

Com isso, conclui-se que a trajetória e o alcance dos blocos estão relacionados principalmente com a geometria e condições superficiais do talude, mas é também importante se levar em consideração as características dos materiais que compõem o bloco e as superfícies de impacto, bem como a forma e tamanho dos blocos, assim como já propôs Ritchie (1963) e Hoek (2000).

### **2.3. AVALIAÇÃO DE ESTABILIDADE, PERIGO E RISCO PARA TALUDES**

São diversas as metodologias propostas encontradas na literatura para avaliar as condições de estabilidade de um talude, de risco, ou de perigo de rupturas. Esses sistemas vão desde classificações empíricas cujo resultado pode fornecer indiretamente informações a respeito da suscetibilidade a rupturas, até metodologias mais complexas, que quantificam o perigo dessas rupturas e até mesmo o risco. Também é possível encontrar metodologias mais abrangentes, ou seja, aquelas que se adequam a qualquer contexto, como sistemas de análise focados em mais de uma determinada situação, como rodovias, minerações e áreas urbanas.

O *Slope Mass Rating* – SMR (Romana, 1985) é um sistema para avaliar a condição de estabilidade de um talude a partir do *Rock Mass Rating* – RMR (Bieniawski, 1989). Isso significa que no SMR é aplicado ao valor do RMR básico um ajuste baseado na orientação



das descontinuidades em relação à orientação do talude, na resistência ao cisalhamento das descontinuidades, bem como no método de escavação (Equação 2.1).

$$SMR = RMR_b + (F_1 \times F_2 \times F_3) + F_4 \quad (2.1)$$

Onde:

RMR<sub>b</sub>: RMR básico;

F<sub>1</sub>: fator relacionado ao paralelismo entre a direção de mergulho do talude e das descontinuidades;

F<sub>2</sub>: fator relacionado à resistência ao cisalhamento das descontinuidades (Romana, 1993), que para o caso de rupturas planares, depende do ângulo de mergulho das descontinuidades. Para o caso de tombamentos, adotar 1;

F<sub>3</sub>: fator que depende da relação entre os ângulos de mergulho da face do talude e das descontinuidades;

F<sub>4</sub>: fator relacionado ao método de escavação.

Sendo assim, o SMR fornece como resultado, além da classe de qualidade do maciço, a condição de estabilidade e também a probabilidade de haver ruptura (Tabela 2.1). É uma metodologia empírica para estimar a probabilidade de haver ruptura, aplicada a taludes em qualquer contexto (rodovias, minerações e qualquer outro), sem qualquer referência a perigo, consequências dessas rupturas ou riscos.

Tabela 2.1 - Descrição de classes e resultados do SMR (Romana, 1985).

Classe	V	IV	III	II	I
SMR	0 - 20	21 - 40	41 - 60	61 - 80	81 - 100
Qualidade do maciço	Muito ruim	Ruim	Moderada	Boa	Muito Boa
Estabilidade	Totalmente instável	Instável	Parcialmente estável	Estável	Muito Estável
Rupturas	Grandes rupturas – semelhantes a solo	Planares e grandes cunhas	Cunhas	Alguns blocos	Nada
Probabilidade	0,9	0,6	0,4	0,2	0

Outra metodologia, essa mais recente, para avaliação da condição de estabilidade de taludes é o Q-Slope, proposto por Bar & Barton (2017). Através desse sistema, é possível determinar se um talude é estável, instável ou se está em uma zona de incerteza. Esse resultado é obtido através de um gráfico que relaciona o valor do Q-Slope com a inclinação do talude, como mostra a Figura 2.4. O cálculo do Q-Slope (Equação 2.2) se dá de maneira semelhante, com algumas adaptações, com relação ao Q-Tunnels (Barton, 1974).

$$Q_{slope} = \frac{RQD}{J_n} \times \left(\frac{J_r}{J_a}\right)_o \times \frac{J_{wice}}{SRF_{slope}} \quad (2.2)$$

Onde:

RQD: *Rock Quality Designation* (Deere 1963; Deere et al. 1967);

J<sub>n</sub>: índice relacionado ao número de famílias de descontinuidades;

J<sub>r</sub>: índice relacionado à rugosidade das descontinuidades;

J<sub>a</sub>: índice relacionado à alteração e preenchimento das descontinuidades;

J<sub>wice</sub>: índice relacionado às condições climáticas/ambientais;

SRF<sub>slope</sub>: índice relacionado às tensões no maciço, efeito do desmonte e/ou presença de descontinuidades importantes, como falhas persistentes.

O: Fator O (*o-factor*), que é um fator de ajuste aplicado ao valor de (J<sub>r</sub>/J<sub>a</sub>) relacionado com a orientação das descontinuidades.

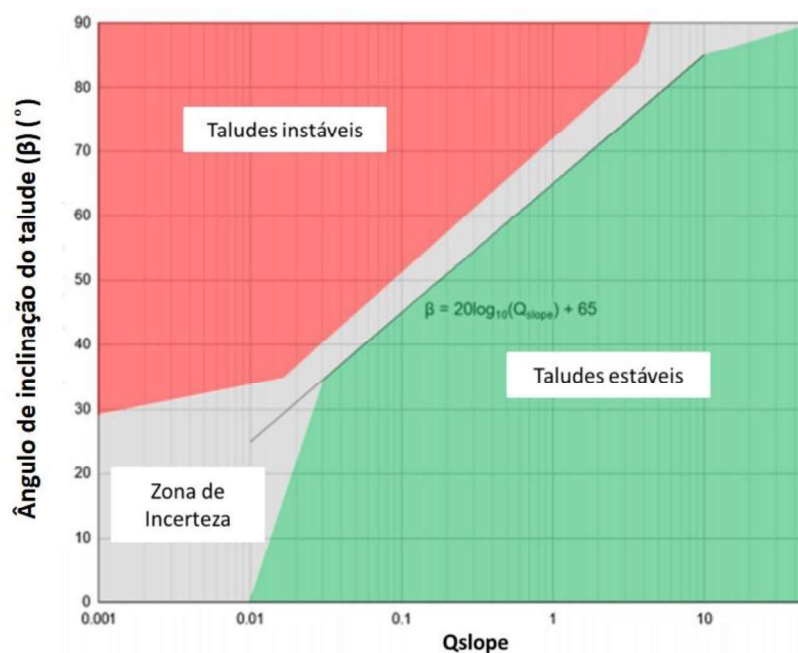


Figura 2.4- Carta de estabilidade do Q-Slope (Bar & Barton, 2017).

Ainda na Figura 2.4 é possível identificar uma equação onde o valor de  $\beta$  (ângulo de inclinação do talude) está em função do valor obtido para o Q-slope. Esse valor para  $\beta$  é o ângulo máximo de inclinação que o talude pode ter para se manter estável, dado um determinado valor para o Q-Slope. Com isso, o Q-Slope pode ser tratado também como uma ferramenta preliminar de dimensionamento de um ângulo de inclinação adequado para um talude, considerando suas condições geomecânicas.

Bauer & Neumann (2011) propuseram uma metodologia para análise de perigo de queda de blocos em áreas urbanas em que o foco era a região dos Alpes Bávicos, na Alemanha. Tendo em mãos um banco de dados com as áreas de ocorrência de quedas de blocos, foi sugerida essa metodologia de avaliação através de observações de campo. A classe de perigo é dada através da classificação da probabilidade do evento ocorrer e de sua intensidade, sendo a probabilidade relacionada com a suscetibilidade de ocorrência de quedas de blocos e influências externas, como clima e sismos. A intensidade é dada de acordo com o volume (ou massa) de rocha destacada.

A suscetibilidade é determinada através da soma dos pesos (notas) atribuídos a parâmetros geomecânicos do maciço rochoso que compõem o talude a ser avaliado, sendo esses parâmetros resumidos na Tabela 2.2, juntamente com indícios de atividade recente, como blocos soltos no maciço e caídos no pé do talude. O impacto das influências externas é dado pela soma dos pesos atribuídos a faixas de precipitação e zona sísmica em que a região está inserida (Tabela 2.3).

Tabela 2.2- parâmetros geomecânicos observados para a determinação da suscetibilidade a queda de blocos (Adaptado de Bauer & Neumann,2011).

	Parâmetro	Característica observada
Mecânica de rochas	Orientação relativa de descontinuidades e superfície de talude	Analisar se as descontinuidades são favoráveis ao movimento de queda em relação ao talude. Se existem famílias que formam blocos, ou cunhas.
	Grau de alteração	Analisar o grau de alteração, de acordo com a ISRM (1981). Quanto mais alterado, menor a resistência, coesão e ângulo de atrito no maciço. Além do mais, a alteração das rochas facilita o processo de erosão, o que pode levar a queda de blocos.
	Configuração estrutural	Compreende o grau de transecção das descontinuidades (presença de mineral de preenchimento duro, impermeável, que sela as paredes das descontinuidades), persistência, abertura e rugosidade das paredes.

Tabela 2.2- parâmetros geomecânicos observados para a determinação da suscetibilidade a queda de blocos (Adaptado de Bauer & Neumann,2011).

Mecânica das rochas	Grau de flexibilização	Parâmetro utilizado quando não é possível obter detalhadamente em campo um ou mais parâmetros citados acima (orientação, alteração ou configuração estrutural). São feitas observações gerais em campo, porém com menor grau de confiança.
Ambiente Geomecânico	Tipo de formação	Se na formação rochosa houver presença de rocha solúvel em água, ou de composição argilosa, por exemplo, o estado de tensões no maciço pode ser alterado e levar a movimentação de massa.
	Deformações profundas	Deformações podem alterar o estado de tensões no maciço em longo prazo e influenciar na estabilidade.
	Movimento no pé do talude	Movimentos de massa no pé do talude podem influenciar na estabilidade de toda a área.

Tabela 2.3- Influências externas (Bauer & Neumann, 2011).

Zona de terremoto	Zona 2	Zona 1	Zona 0
Probabilidade atribuída	Alta	Média	Baixa
Nota	50	20	0
Precipitação	> 310 mm	230 – 310 mm	< 230 mm
Probabilidade atribuída	Alta	Média	Baixa
Nota	50	20	0

Para análise da precipitação, os autores propõem considerar o evento mais intenso em um intervalo de 72h, para um período de recorrência de 100 anos. No caso, os dados utilizados nessa metodologia foram fornecidos pelo Serviço Climático Alemão. Já para a sismicidade, consideram o zoneamento previsto no Código Alemão DIN 4149, que se baseia no Mapa de risco sísmico da Suíça, Áustria e Alemanha (Silveira, 2017).

A intensidade do evento é medida a partir do volume de massa rochosa destacada, ou da dimensão característica do bloco. No caso, a intensidade é dada em duas escalas em termos de magnitude do evento: quedas em baixa escala de magnitude que se subdivide em queda de um bloco único ou pequenos blocos de até 200mm, quedas com volume até 100m<sup>3</sup> ou blocos maiores que 200mm e, por fim, quedas com volume até 10000m<sup>3</sup>; quedas em larga escala e avalanches de rochas, com volume de blocos maior que 10000m<sup>3</sup> e maior que 1000000m<sup>3</sup>, respectivamente. A Figura 2.5 mostra as matrizes de probabilidade (2.5a) e de perigo (2.5b).

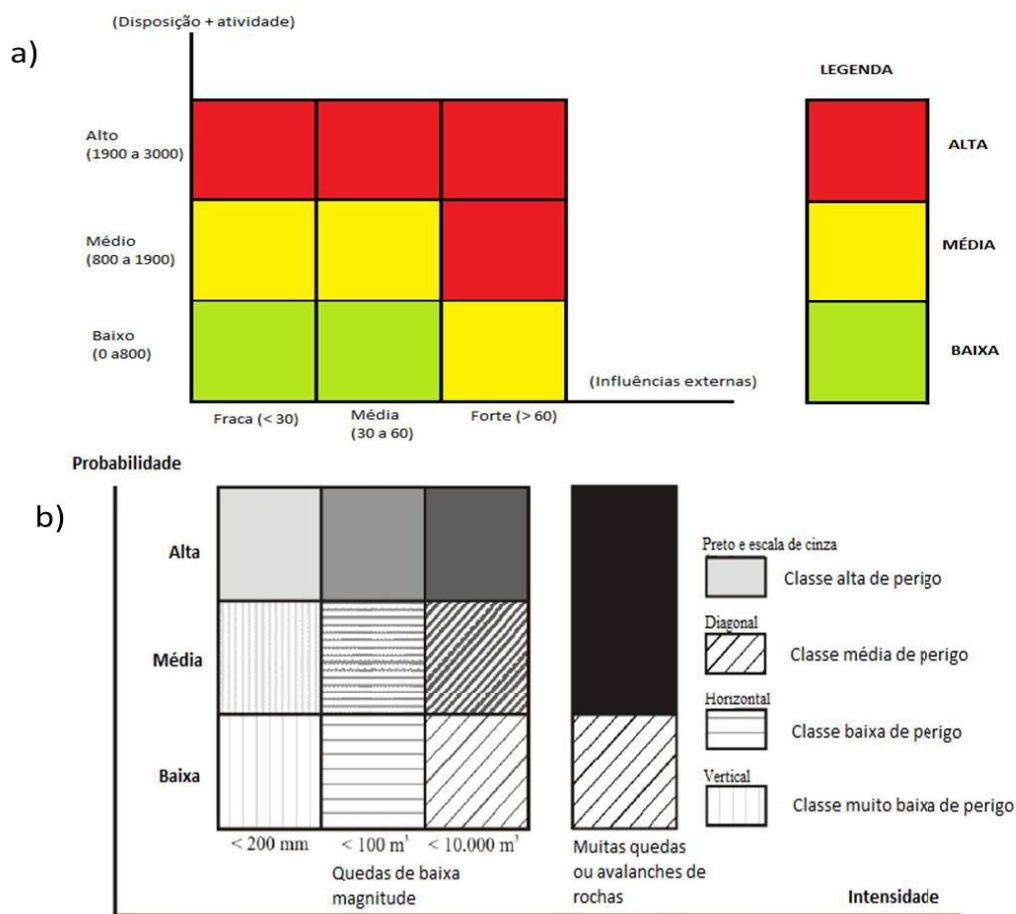


Figura 2.5- a) Matriz de probabilidade; b) Matriz de perigo (Adaptado de Bauer & Neumann, 2011).

O sistema proposto por Bauer & Neumann (2011), apesar de prático e fácil de utilizar, é apropriado apenas para quedas de blocos em áreas urbanas de regiões alpinas e, como mencionado acima, foi desenvolvido considerando as condições geológicas e climáticas das regiões dos Alpes Bávares, gerando respostas pouco representativas quando aplicado em outros locais. Diante disso, os próprios autores encorajam testes e adaptações a outras condições.

Silveira (2017) propôs algumas adaptações a essa metodologia, de modo que ela pudesse abranger não somente outras condições geológicas, sísmicas e climáticas, como também taludes em outros contextos, tais como minerações, rodovias, ferrovias e também em áreas urbanas. No caso, houve a troca de alguns parâmetros em relação à metodologia original; no entanto, a avaliação final de perigo se dá pelas mesmas matrizes (Figura 2.5) e da mesma forma que a metodologia original.

Silveira (2017) retirou da análise os parâmetros inseridos em Ambiente Geomecânico (Tabela 2.2) e passou a considerar a geometria do talude e características da área de captação de blocos, como a distância, inclinação e material que a compõe. Nessa metodologia adaptada, tanto taludes mais altos, como alguns naturais em áreas urbanas, quanto taludes mais baixos, como cortes rodoviários e bancadas de mineração, podem ser avaliados. Na Tabela 2.4 os parâmetros e os pesos adotados na metodologia proposta por Silveira (2017) são detalhados.

Tabela 2.4- Pesos atribuídos aos parâmetros para a metodologia adaptada por Silveira (2017).

<b>1 - Orientação das descontinuidades</b>	Desfavorável	Razoável	Favorável
Característica	Oposta ou paralelo/subparalela ao talude	Horizontal	Internas ou verticais
Probabilidade atribuída	Alta	Média	Baixa
Peso	500	200	0
<b>2 - Grau de alteração</b>	Forte alteração (W4 a W5)	Baixo grau de alteração (W1 a W3)	
Probabilidade atribuída	Alta	Baixa	
Peso	200	0	
<b>3 - Configuração estrutural das descontinuidades</b>			
<b>Persistência</b>	Mais de 10 m de extensão (alta).	Entre 1 e 10 m de extensão (média).	Menos de 1m de extensão (baixa).
Peso	200	100	0
<b>Grau de transecção</b>	Nenhum mineral de ligação		Existência de mineral de ligação
Peso	200		0
<b>Abertura das descontinuidades</b>	Abertas (>1 cm)	Suavemente abertas (entre 0,5mm a 1 cm).	Fechadas (<0,5 mm)
Peso	200	100	0
<b>Condição das paredes</b>	Lisas		Rugosas
Peso	200		0
<b>4) Grau de afrouxamento (alternativo a categoria 3).</b>	Indicações claras, como: juntas abertas e alteradas, lacunas entre as paredes (alto).	Poucas indicações visíveis (médio).	Nenhuma indicação (baixo).
Peso	800	400	0
<b>Total de mecânica de rochas (1+2+(3 ou 4))</b>			
<b>5) Altura do talude (H) (m)</b>	<b><math>IAT = e^{0,07996xH}</math></b>		
Peso: Pontuação atribuída (Nota = IAT)	600	200	100
<b>6) Inclinação e superfície</b>	Plano inclinado com feições de lançamento	Talude sub-vertical a vertical com feições de lançamento	Abaixo de 30° ou próximo a 90°; sem feições de lançamento
Classe atribuída	Alta	Média	Baixa
Peso	200	100	0

Tabela 2.4- Pesos atribuídos aos parâmetros para a metodologia adaptada por Silveira (2017).			
<b>7) Área de Captação</b>	Pequena distância, inclinação elevada, com pouca ou nenhuma vegetação.	Distância e/ou inclinação moderada. Pouca ou nenhuma vegetação	Grande distância, baixa inclinação, com ou sem vegetação.
Classe atribuída	Alta	Média	Baixa
Peso	200	100	0
<b>Total de geometria do talude e área de captação (5+6+7)</b>			
<b>8) Indícios de atividade</b>	Ativo		Não ativo
Peso	500		0
<b>Total de atividade (0 a 500)</b>			
<b>Total de disposição e atividade = suscetibilidade [1+2+(3 ou 4)+5+6+7+8]</b>			
Classe atribuída	Alta	Média	Baixa
Faixa de somatório de pesos	1900 a 3000	800 a 1900	0 a 800
<b>9) Sismos</b>	<b>Sismos tectônicos</b>	<b>Minerações, regiões cársticas e veículos pesados.</b>	<b>Região isenta de sismos.</b>
Classe atribuída	Alta	Média	Baixa
Nota	50	20	0
<b>10) Precipitação</b>	<b>&gt; 129 mm/6 dias</b>	<b>48,2–129 mm/6 dias</b>	<b>&lt; 48,2 mm/6 dias</b>
Classe atribuída	Alta	Média	Baixa
Nota	50	20	0
<b>Total de impacto externo (9+10)</b>			
Classificação	Forte (> 60)	Média (30 a 60)	Fraca (< 30)

É importante salientar que nessa adaptação se pretende avaliar o perigo de queda de blocos originados principalmente pelas descontinuidades do maciço rochoso e influências externas. Casos relacionados a problemas de deformabilidade, mais comuns em taludes muito altos de regiões alpinas, não são adequados para esse sistema adaptado.

Com relação às influências externas, as faixas de precipitação foram modificadas a fim de se considerar uma maior amplitude entre a precipitação máxima e mínima acumulada em 6 dias, tomando como base a região de Ouro Preto (Minas Gerais, Brasil) (Silveira, 2017). Já com relação aos sismos, Silveira (2017) considerou como a pior situação os sismos gerados por

terremotos e como situação intermediária aqueles gerados por detonações e tráfego de veículos pesados.

Santos (2019) propôs um sistema de análise de perigo e de risco quantitativo para taludes de mineração. Para isso, utilizou um banco de dados com 88 taludes de minas de várias partes do mundo, cuja situação de estabilidade era conhecida. As variáveis utilizadas nesse método são mostradas na Tabela 2.5.

Tabela 2.5- Caracterização dos parâmetros adotados na proposta de Santos (2019).

Variável							Tipo
Resistência à compressão uniaxial da rocha intacta (MPa) - P <sub>1</sub>							Métrica
Espaçamento (m) - P <sub>2</sub>							Métrica
Persistência da descontinuidade principal (m) - P <sub>3</sub>							Métrica
Abertura da descontinuidade principal (mm) - P <sub>4</sub>							Métrica
Rugosidade da descontinuidade principal -P <sub>5</sub>							
Parâmetro	Muito rugosa	Rugosa	Levemente rugosa	Lisa	Polida	Qualitativa	
Peso	5	4	3	2	1		
Preenchimento da descontinuidade principal -P <sub>6</sub>							
Parâmetro	Nenhum	Preenchimento duro		Preenchimento macio		Qualitativa	
Peso	5	<5 mm	>5 mm	<5 mm	>5 mm		
		4	3	2	1		
Alteração da rocha -P <sub>7</sub>							
Parâmetro	Não alterada (W1)	Ligeiramente alterada (W2)	Moderadamente alterada (W3)	Muito alterada (W4)	Decomposta (W5)	Qualitativa	
Peso	5	4	3	2	1		
Condição de água subterrânea -P <sub>8</sub>							
Parâmetro	Seco	Úmido	Encharcado	Gotejando	Com fluxo	Qualitativa	
Peso	5	4	3	2	1		
Orientação da descontinuidade principal* -P <sub>9</sub>							
Parâmetro *	$\beta d > \beta s$ $ad - as > 30^\circ$	$\beta d > \beta s$ $ad - as < 30^\circ$	$0 \leq \beta d \leq \beta s / 4$ $ad - as > 30^\circ$	$\beta s / 4 \leq \beta d \leq \beta s / 2$ $ad - as < 30^\circ$	$\beta s / 2 \leq \beta d \leq \beta s$ $ad - as < 30^\circ$	Qualitativa	
Descrição	Muito favorável	Favorável	Razoável	Desfavorável	Muito desfavorável		
Peso	5	4	3	2	1		
* $\beta d$ = mergulho da descontinuidade; $\beta s$ = mergulho do talude; $ad$ = direção de mergulho da descontinuidade; $as$ = direção de mergulho do talude.							
Método de desmonte -P <sub>10</sub>							
Parâmetro	Pré-fissuramento	Pós-fissuramento	Desmonte macio/amortecido	Desmonte de produção modificado	Desmonte regular / mecânico	Qualitativa ordinal	
Peso	5	4	3	2	1		
Altura do talude global (m) -P <sub>11</sub>							Métrica
Inclinação do talude global (°) - P <sub>12</sub>							Métrica



As técnicas de estatística multivariada utilizadas para gerar o sistema de Análise de Perigo foram Análise de Componentes Principais (ACP) por matriz de correlação, Análise Discriminante e Elipses de Confiança. A Análise Discriminante, por ser uma técnica de discriminação e classificação, foi utilizada para gerar o modelo. Para isso, é necessário que o banco de dados contenha variáveis independentes quantitativas e que a variável dependente seja qualitativa, ou seja, classificatória.

A Análise de Componentes Principais foi, portanto, aplicada ao banco de dados com o objetivo de quantificar as variáveis, visto que variáveis independentes quantitativas é um pré-requisito para aplicação da Análise Discriminante e, como se pode observar na Tabela 2.5, algumas variáveis do banco de dados são qualitativas. A escolha pela matriz de correlação se deu pela necessidade de padronizar as variáveis, que possuem pesos e unidades de medidas diferentes.

Com isso, a Análise Discriminante foi aplicada após a ACP, sendo as variáveis independentes os escores das componentes principais. A variável dependente foi a condição de estabilidade dos 88 taludes analisados, onde foi possível verificar três condições: estabilidade, ruptura inter-rampa e de bancada e, por fim, ruptura global. As Elipses de Confiança foram utilizadas com o intuito de traçar matematicamente os limites entre as classes no gráfico dos escores discriminantes, gerado pela Análise Discriminante.

Portanto, para saber em qual classificação de perigo um talude qualquer se encontra, deve-se primeiramente calcular os escores das componentes principais ( $Cp_n$ ) através das Equações 2.3 a 2.14, utilizando as variáveis P1 a P12 padronizadas (Equação 2.15), onde  $\mu_n$  é a média e  $\sigma_n$  o desvio padrão das variáveis ( $P_n$ ).

$$Cp_1 = 0,24Z_1 + 0,33Z_2 - 0,29Z_3 - 0,34Z_4 + 0,36Z_5 + 0,03Z_6 + 0,41Z_7 + 0,36Z_8 - 0,32Z_9 + 0,21Z_{10} + 0,15Z_{11} + 0,12Z_{12} \quad (2.3)$$

$$Cp_2 = -0,46Z_1 + 0,06Z_2 + 0,10Z_3 - 0,10Z_4 + 0,05Z_5 - 0,04Z_6 + 0,02Z_7 + 0,35Z_8 - 0,30Z_9 - 0,52Z_{10} + 0,07Z_{11} - 0,52Z_{12} \quad (2.4)$$

$$Cp_3 = -0,08Z_1 + 0,41Z_2 + 0,17Z_3 + 0,22Z_4 + 0,01Z_5 - 0,40Z_6 + 0,15Z_7 - 0,04Z_8 + 0,38Z_9 + 0,19Z_{10} + 0,57Z_{11} - 0,22Z_{12} \quad (2.5)$$

$$Cp_4 = -0,34Z_1 + 0,17Z_2 - 0,49Z_3 - 0,24Z_4 - 0,14Z_5 + 0,39Z_6 - 0,35Z_7 - 0,32Z_8 - 0,05Z_9 + 0,19Z_{10} + 0,33Z_{11} - 0,12Z_{12} \quad (2.6)$$

$$Cp_5 = 0,01Z_1 - 0,08Z_2 - 0,14Z_3 - 0,57Z_4 - 0,65Z_5 - 0,33Z_6 + 0,16Z_7 + 0,15Z_8 + 0,23Z_9 - 0,06Z_{10} - 0,05Z_{11} + 0,05Z_{12} \quad (2.7)$$

$$Cp_6 = -0,18Z_1 - 0,18Z_2 - 0,49Z_3 + 0,06Z_4 + 0,33Z_5 - 0,69Z_6 - 0,28Z_7 - 0,04Z_8 - 0,06Z_9 + 0,06Z_{10} - 0,16Z_{11} + 0,03Z_{12} \quad (2.8)$$

$$Cp_7 = -0,22Z_1 + 0,65Z_2 + 0,03Z_3 + 0,07Z_4 - 0,07Z_5 - 0,05Z_6 - 0,17Z_7 - 0,01Z_8 + 0,00Z_9 - 0,34Z_{10} - 0,20Z_{11} + 0,58Z_{12} \quad (2.9)$$

$$Cp_8 = 0,09Z_1 + 0,16Z_2 - 0,52Z_3 - 0,32Z_4 - 0,02Z_5 + 0,22Z_6 + 0,25Z_7 + 0,15Z_8 + 0,48Z_9 - 0,16Z_{10} - 0,34Z_{11} - 0,29Z_{12} \quad (2.10)$$

$$Cp_9 = 0,49Z_1 + 0,25Z_2 + 0,10Z_3 - 0,13Z_4 - 0,01Z_5 - 0,02Z_6 - 0,68Z_7 + 0,33Z_8 + 0,07Z_9 + 0,08Z_{10} - 0,09Z_{11} - 0,28Z_{12} \quad (2.11)$$

$$Cp_{10} = 0,36Z_1 + 0,27Z_2 - 0,00Z_3 - 0,15Z_4 + 0,00Z_5 - 0,18Z_6 + 0,15Z_7 - 0,70Z_8 - 0,24Z_9 - 0,27Z_{10} - 0,17Z_{11} - 0,32Z_{12} \quad (2.12)$$

$$Cp_{11} = 0,35Z_1 - 0,13Z_2 - 0,30Z_3 + 0,43Z_4 - 0,37Z_5 - 0,06Z_6 - 0,05Z_7 + 0,12Z_8 - 0,35Z_9 - 0,31Z_{10} + 0,44Z_{11} + 0,15Z_{12} \quad (2.13)$$

$$Cp_{12} = -0,17Z_1+0,22Z_2+0,03Z_3+0,31Z_4-0,41Z_5-0,09Z_6+0,09Z_7+0,08Z_8-0,44Z_9 \\ +0,54Z_{10}-0,35Z_{11}-0,18Z_{12} \quad (2.14)$$

Onde:

$$Z_n = \frac{(P_n-\mu_n)}{\sigma_n} \quad (2.15)$$

Obtidos os escores das componentes principais ( $Cp_n$ ), é possível então conhecer a classe de perigo dos taludes, plotando os valores obtidos pelas funções discriminantes, LD1 e LD2 (Equações 2.16 e 2.17), no gráfico de perigo (Figura 2.6). O perigo baixo está relacionado aos taludes estáveis; médio, aos taludes que apresentam ruptura em bancos ou inter-rampas; alto, aos taludes que apresentam ruptura global.

$$LD_1 = -1,00Cp_1-0,75Cp_2-0,23Cp_3-0,19Cp_4-0,29Cp_5-0,11Cp_6-0,15Cp_7 \\ -0,08Cp_8-0,02Cp_9+0,39Cp_{10}+0,32Cp_{11}+0,31Cp_{12} \quad (2.16)$$

$$LD_2 = 0,26Cp_1-0,01Cp_2-0,39Cp_3+0,09Cp_4-0,03Cp_5+0,36Cp_6+0,07Cp_7 \\ -0,51Cp_8+0,37Cp_9-0,20Cp_{10}-0,05Cp_{11}+0,01Cp_{12} \quad (2.17)$$

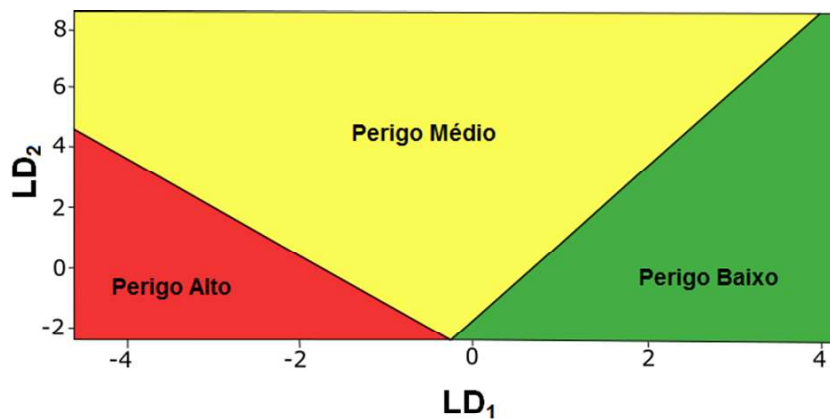


Figura 2.6- Gráfico de perigo (Santos, 2019).

Já o sistema de análise de risco proposto por Santos (2019) é dado através de uma matriz de risco (Figura 2.7), onde a classificação relativa à probabilidade (susceptibilidade) de ruptura se cruza com a classificação relativa às consequências dessa ruptura.

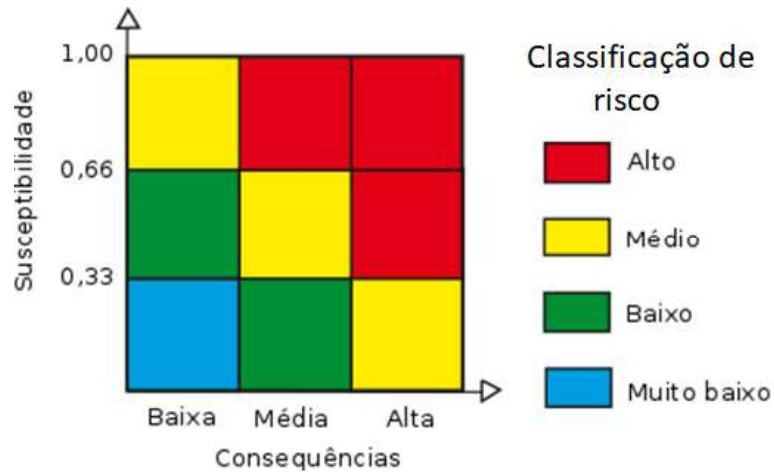


Figura 2.7 – Matriz de risco para taludes de mineração (Santos, 2019).

O modelo de probabilidade de ruptura foi gerado através de regressão logística, onde a variável dependente era a condição de estabilidade do talude, sendo: estável ou instável. As variáveis independentes foram, primeiramente, aquelas mostradas na Tabela 2.5. Posteriormente, o teste de Wald foi conduzido para se determinar as variáveis realmente significativas no modelo logístico gerado, ou seja, aquelas que realmente devem ser consideradas e entrar na equação de determinação da probabilidade. Utilizando um nível de confiança igual a 90%, as variáveis significativas foram: rugosidade (P5), preenchimento (P6), alteração do maciço rochoso (P7), condição de água subterrânea (P8) e orientação da família de descontinuidades principal (P9). O modelo logístico para calcular a probabilidade do talude ser instável é dado através da Equação 2.18.

$$\hat{p}(instável) = \frac{e^{38,86-1,84P5-2,25P6-2,29P7-6,77P8+2,41P9}}{1+e^{38,86-1,84P5-2,25P6-2,29P7-6,77P8+2,41P9}} \quad (2.18)$$

Em relação às consequências, Santos (2019) pondera que rupturas em taludes de mina estão diretamente atreladas à escala de ruptura do talude, visto que rupturas em grande escala, como as globais por exemplo, tendem a gerar danos maiores. Para se mensurar as consequências da ruptura foi utilizada a técnica de árvore de classificação através do algoritmo *Classification and Regression Trees* (CART) proposto por Breiman et al. (1984). A Árvore de Classificação é mostrada na Figura 2.8.

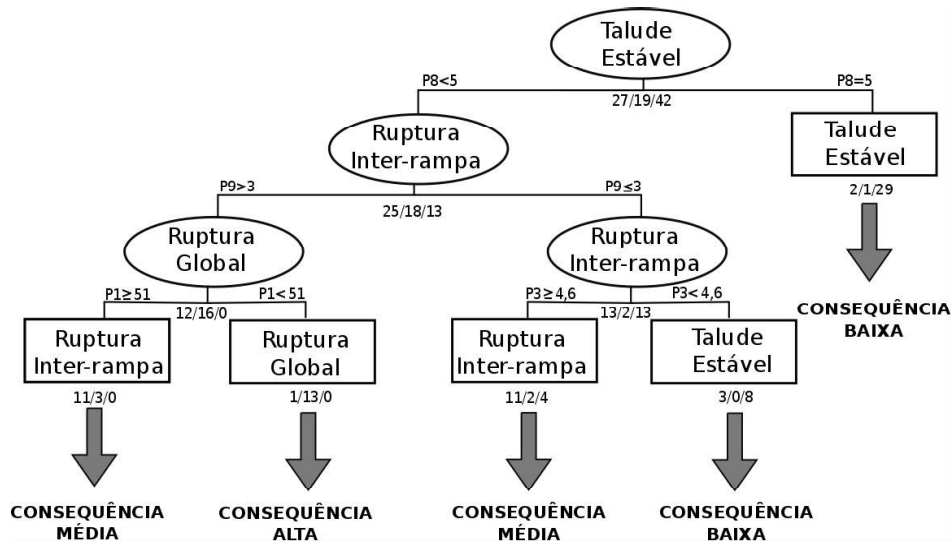


Figura 2.8 – Árvore de classificação (Santos, 2019).

De acordo com Santos (2019), em cada nó da árvore se observa três números onde o primeiro número é a quantidade de taludes que apresentaram ruptura inter-rampa, o segundo representa aqueles que apresentaram ruptura global e o terceiro, os taludes estáveis. Daí, cada nó é rotulado de acordo com a classe que possui maior número de indivíduos, sendo que os rótulos são: talude estável é a consequência baixa; ruptura inter-rampa é consequência média; ruptura global, consequência alta.

A metodologia proposta por Santos (2019) é particularmente interessante, pois ao aliar a estatística com a geotecnia, a autora conseguiu propor equações para quantificar o perigo e o risco. Essas foram geradas considerando a atuação de todos os parâmetros em conjunto para uma determinada condição de estabilidade conhecida e, com isso, o impacto de cada parâmetro no resultado final é melhor contabilizado, em vez de simplesmente considerar que todas as variáveis são igualmente importantes para a condição final de estabilidade. No entanto, são sistemas de perigo e de risco apropriado para taludes de mina e a aplicação em outros locais requer mais estudos para possíveis adaptações.

## 2.4. AVALIAÇÃO DE PERIGO E RISCO DE QUEDA DE BLOCOS PARA TALUDES RODOVIÁRIOS

A partir dos procedimentos adotados por Wyllie (1987) para determinar problemas com quedas de blocos em taludes de rodovias, outras classificações surgiram com o intuito de auxiliar no controle desse tipo de evento nas estradas (Silva, 2016). Uma das metodologias

mais difundidas para avaliar problemas de quedas de blocos em rodovias é o Rockfall Hazard Rating System (RHRS) desenvolvido por Pierson *et al.* (1993) e implantado no Estado do Oregon (EUA). O RHRS avalia o nível de perigo ao qual uma determinada rodovia está exposta utilizando parâmetros importantes para esse tipo de análise (Tabela 2.6). Para as características geológicas, são consideradas as características para taludes em rochas cristalinas (1) e taludes em rochas sedimentares (2).

Esses nove parâmetros recebem escores (notas) que vão de 3 a 81, variando exponencialmente. No caso, quando um determinado parâmetro possui uma característica favorável a segurança/estabilidade, recebe 3. Essa nota aumenta à medida que a situação se torna mais desfavorável, até 81. Os escores de todos os parâmetros são somados, e quanto mais alta a pontuação maior o perigo, sendo possível hierarquizar os locais com maior urgência de intervenção. Além do mais, é possível classificar o perigo da seguinte maneira: se o somatório for menor que 300, o perigo é baixo; se for maior que 500, é alto e possui prioridade de intervenção (Pierson & Van Vickle, 1993; Hoek, 2007).

Tabela 2.6– RHRS original para o estado do Oregon (EUA) (Pierson & Van Vickle, 1993).

		Pontuação			
Parâmetro		3	9	27	81
Altura do talude (m)		7,5	15	22,5	30
Efetividade da Área de Captação		Boa captura	Moderada	Limitada	Sem captura
Risco Médio ao Veículo		25% do tempo	50% do tempo	75% do tempo	100% do tempo
% de Distância de Visibilidade		100%; distância de visibilidade adequada	80%; distância moderada	60%; distância limitada	40%; muito limitada
Largura da Rodovia (m)		13,2	10,8	8,4	6,0
Características Geológicas	Condição estrutural (1)	Fraturas descontínuas e orientação favorável	Fraturas descontínuas e orientação aleatória	Fraturas descontínuas e orientação adversa	Fraturas contínuas e orientação adversa
	Rugosidade (1)	Rugosa, irregular	Ondulações suaves	Planas	Estriadas ou preenchimento argiloso
	Condição estrutural (2)	Pouca influência de erosão diferencial	Erosão ocasional	Muitos focos de erosão	Grandes e muitos focos de erosão
	Erosão diferencial (2)	Pouca diferença	Diferença moderada	Grande diferença	Diferença extrema
Clima e presença de água		Precipitação baixa a moderada; sem períodos de congelamento; sem presença de água no talude	Precipitação moderada; ou curtos períodos de congelamento; ou presença intermitente de água no talude	Precipitação alta; ou longos períodos de congelamento; ou presença contínua de água no talude	Precipitação alta e longos períodos de congelamento; ou presença de água no talude com longos períodos de congelamento
Histórico de quedas		Poucas	Ocasionais	Muitas	Recorrentes
Tamanho do bloco (m)/ Volume de queda (m <sup>3</sup> )		0,3/2,3	0,6/4,6	0,9/6,9	1,2/9,2

O fato de os parâmetros considerados no RHRS serem basicamente obtidos através de observações de campo e informações sobre o tráfego torna o método uma ferramenta rápida e prática para se avaliar o perigo que um determinado talude oferece aos usuários de uma

rodovia quanto à queda de blocos. Visto que uma rodovia é uma obra linear, pode ser que ao longo dos vários quilômetros percorridos existam muitos taludes que necessitam de análise; sendo assim, rapidez e praticidade são primordiais em qualquer ferramenta gerada para esse fim e isso pode ter sido um fator que contribuiu para a popularidade do método, que possui variantes de acordo com as condições do local onde é aplicado.

O RHRS foi adotado por 18 estados diferentes dos EUA, sendo que alguns deixaram o sistema como está, enquanto outros o modificaram para se adequar às suas condições locais (Bateman, 2003). Conforme citado por Russel et al. (2008), os estados que não modificaram o sistema incluem Califórnia, Kentucky, Nevada, Pensilvânia, Virgínia (Witteman et al. 1992), Wyoming, Carolina do Norte, Utah (Pack & Boie, 2002) e West Virginia (Szwilski, 2002). Os estados que modificaram incluem Colorado (Stover, 1992), Arizona, Nova Jersey, Vermont (Eliassen & Ingraham, 2000), Nova York (NYSDOT, 1996), Tennessee (Vandewater et al. 2005), Idaho (Miller, 2003), Ohio (Shakoor, 2005) e New Hampshire (Fish & Lane, 2001).

Em Ohio as principais modificações em relação ao RHRS do Oregon foram feitas para se adequar às condições geológicas locais. Ohio é caracterizada por ter muitas rochas sedimentares, sendo assim, o problema de queda de blocos tem forte relação com a erosão diferencial entre as camadas, o que pode levar ao descalçamento das camadas menos erodíveis. Com isso, o Departamento de Trânsito de Ohio (DOT Ohio) passou a realizar testes de durabilidade (*Slake Durability Test*) para incluir no RHRS (Russel et al., 2008).

A outra mudança se refere à eficácia da área de captação; para o RHRS de Ohio, a eficácia da área de captação é medida ao se comparar as dimensões reais, medidas em campo, com aquelas recomendadas pelo ábaco proposto por Ritchie (1963), que usa a altura e o ângulo do talude para projetar a área de captação (Shakoor, 2005; Ritchie, 1963).

Conforme já foi citado por Silva (2016), o Departamento de Transito do Estado de Nova York (NYSDOT, 1996) também modificou o RHRS do Oregon, porém, com a finalidade de estabelecer um provável risco, incorporando na análise fatores relacionados à seção rodoviária (SF) e de exposição humana (HEF), além dos fatores geológicos (GF). Ao se multiplicar esses três fatores, se tem o Risco Relativo Total (TRR) (NYSDOT, 1996).



O Fator Geológico considera, entre outros fatores relacionados às discontinuidades, dois tipos de taludes separadamente: cristalino e sedimentar. Encostas cristalinas são classificadas com base na persistência, número e orientação das discontinuidades. Encostas sedimentares são classificadas com base no mergulho das camadas e de outras discontinuidades, além da erosão diferencial (NYSDOT, 1996).

O Fator de Seção rodoviária avalia a possibilidade dos blocos chegarem à estrada através da comparação direta com os critérios de projeto de área de captação de Ritchie. Já o fator de exposição humana avalia as chances de ocorrência de acidentes devido à queda de blocos, com base em parâmetros como comprimento do talude, distância de visibilidade para tomada de decisão, número de pistas, tráfego médio diário, limite de velocidade e distância de parada disponível (NYDOT, 1996).

Nesse sistema também foi incluído um parâmetro de Redução de Risco (RR) para taludes, caso haja medidas de mitigação. Para os vários tipos de mitigações, são atribuídos valores distintos de RR. Com isso, o Risco Relativo Total (TRR) é reduzido por este valor RR, resultando em um valor de Risco Residual. ( $\text{Risco Residual} = \text{TRR} - \text{RR}$ ) (Silva, 2016).

De acordo com Russel *et al.* (2008), o RHRS utilizado no Colorado (CRHRS) é uma ligeira modificação do RHRS do Oregon. A origem do CRHRS se deu quando Stover (1982), junto com o *Colorado Geological Survey*, prepararam um sistema para que o Departamento de Trânsito do Colorado (CDOT) pudesse avaliar o perigo de queda de blocos em rodovias em todo o estado. O banco de dados utilizado era referente a acidentes devido à queda de blocos em estradas estaduais do Colorado (CRASH), onde os dados sobre os acidentes foram combinados com outros parâmetros, como largura da estrada, tráfego, entre outros. Em 1994, o Colorado Rockfall Hazard Rating System foi finalizado e incluiu os parâmetros mostrados na Tabela 2.7.

Em 1997 o CDOT modificou esse modelo proposto e mostrado na Tabela 2.7, incluindo na análise a eficácia da área de captação, a distância de visibilidade para tomada de decisão e tráfego médio diário. Outra modificação ocorreu em 2003, substituindo o tráfego médio diário pelo risco médio aos veículos (Russel *et al.*, 2008).

Tabela 2.7– RHRS original para o estado do Colorado (EUA) (Adaptado de Andrew, 1994).

Fator	Parâmetro	Pontuação			
Perfil do Talude		3	9	27	81
	Altura do talude (ft)	25 – 50	50 – 75	75 – 100	> 100
	Extensão (ft)	0 - 250	250 - 500	500 – 750	> 750
	Inclinação do talude (°)	15 a 25	25 a 35	35 a 50	> 50
	Condições da superfície	Possíveis feições de lançamento	Poucas e pequenas feições	Muitas feições de lançamento	Muitas e grandes feições
Características geológicas	Tamanho médio do bloco	6 a 12 in	1 a 2 ft	2 a 5 ft	5 ft
	Quantidade de eventos	1 ft <sup>3</sup> a 1 yd <sup>3</sup>	1 a 3 yd <sup>3</sup>	3 a 10 yd <sup>3</sup>	> 10 yd <sup>3</sup>
	Condição estrutural (1)	Fraturas descontínuas e orientação favorável	Fraturas descontínuas e orientação aleatória	Fraturas descontínuas e orientação adversa	Fraturas contínuas e orientação adversa
	Rugosidade (1)	Rugosa, irregular	Ondulações suaves	Planas	Estriadas; preenchimento argiloso
	Condição estrutural (2)	Pouca influência de erosão diferencial	Erosão ocasional	Muitos focos de erosão	Grandes e muitos focos de erosão
	Erosão diferencial (2)	Pouca diferença	Diferença moderada	Grande diferença	Diferença extrema
Clima e presença de água		Precipitação baixa a moderada; sem períodos de congelamento; sem presença de água no talude	Precipitação moderada; ou curtos períodos de congelamento; ou presença intermitente de água no talude	Precipitação alta; ou longos períodos de congelamento; ou presença contínua de água no talude	Precipitação alta e longos períodos de congelamento; ou presença de água no talude com longos períodos de congelamento
Histórico de quedas		Poucas	Ocasionais	Muitas	Recorrentes
Número de acidentes		0 a 5	5 a 10	10 a 15	> 15

### 2.4.1. CRHRS Modificado

Após as modificações no CRHRS original de 1994 citadas anteriormente, Santi *et al.* (2009) ainda propuseram uma outra versão modificada do CRHRS onde foram adicionados alguns fatores geológicos e climáticos que, de acordo com os autores, contribuem para a queda de blocos na região do Colorado. Esse novo CRHRS foi então utilizado tanto em taludes de corte, como nos taludes naturais acima do corte, que os autores chamam de “taludes totais”.

Ao todo, 355 taludes foram avaliados, sendo: para rochas cristalinas, 147 taludes de corte e 77 taludes totais; para rochas sedimentares, 43 taludes de corte e 31 taludes totais; para taludes em matriz de solo com blocos de rocha encrustados, foram 45 taludes de corte e 12 taludes totais. Os parâmetros inseridos no CRHRS modificado estão na Tabela 2.8 e são todos detalhados adiante, de acordo com Santi *et al.* (2009).

Destaca-se aqui a importância de se explicar cada parâmetro do CRHRS modificado, visto que o banco de dados que deu origem ao sistema de análise de risco proposto nesse trabalho é justamente parte desse com 355 taludes, que foi gentilmente cedido por Paul Santi. No entanto, as discussões a respeito das variáveis utilizadas de fato e importantes em um contexto fora dos limites do Colorado, além das mudanças realizadas, serão feitas mais adiante, nos capítulos reservados à explicação da metodologia aplicada e dos resultados obtidos.

Tabela 2.8– CRHRS modificado (Adaptado de Santi et al., 2009).

Grupo	Parâmetro	Pontuação			
		3	9	27	81
Talude	Altura do talude (m)	7,5 a 15	15 a 23	23 a 30	> 30
	Frequência de queda	> 2 anos	1 a 2 anos	Sazonal	O ano todo/ Eventos severos
	Escore da inclinação do talude	0 a 2	2 a 4	4 a 8	> 8
	Feições de lançamento	Nenhuma (talude suave)	Menores	Maiores	Extremas
	Área de Captação	95 a 100% / Classe 1	65 a 94% / Classe 2	30 a 64% / Classe 3	< 30% / Classe 4
	Clima	Precipitação anual (mm)	< 250	250 a 500	500 a 900
Ciclos de gelo/degelo anuais		1 a 5	6 a 10	11 a 15	> 16
Presença de água		Seco	Úmido	Gotejando	Com fluxo
Aspecto do talude		N	E, W, NE, NW	SE, SW	S
Características geológicas  (1) Taludes cristalinos  (2) Taludes Sedimentares  (*) Descontinuidades: maciços cristalinos e sedimentares  (3) Taludes em matriz de solo com blocos de rocha	Características do maciço (1)	Homogêneo/ maciço	Pequenas falhas/veios fortes	Xistosidades e zonas de cisalhamento (< 15 cm)	Pegmatitos/ micas/zonas de cisalhamento (>15 cm)
	Grau de descalçamento (1)	0 a 0,3 m	0,3 a 0,6 m	0,6 a 1,2 m	> 1,2 m
	Alteração (1)	Fresca	Descoloração superficial	Levemente alterado/ amolecido	Alterações internas
	Grau de descalçamento (erosão diferencial) (2)	0 a 0,3 m	0,3 a 0,6 m	0,6 a 1,2 m	> 1,2 m
	Jar slake (2)	6	5	3 a 4	1 a 2
	Camadas (2)	1 a 2 camadas fracas < 15 cm	1 a 2 camadas fracas > 15 cm	> 2 camadas fracas < 15 cm	> 2 camadas fracas >15 cm

		1	1 + aleatórias	2	> 2
	Nº de famílias (*)	1	1 + aleatórias	2	> 2
	Tamanho de bloco/volume (*)	<0,3m/ <0,75m³	0,3 a 0,6m/ 0,75 a 2,3m³	0,6 a 1,5m/ 2,3 a 7,6m³	> 1,5m/ > 7,6m³
	Persistência e orientação (*)	< 3m mergulhando para dentro do talude	> 3m cortando o talude	< 3m mergulhando para fora do talude	> 3m mergulhando para fora do talude
	Abertura (*)	0	0,1 a 1 mm	1 a 5 mm	> 5mm
	Alteração/preenchimento (*)	Fresca	Descoloração superficial	Preench. granular	Preench. argiloso
	Rugosidade (*)	Rugosa	Ondulada	Plana	Estriada
	Tamanho do bloco (3) [x 3]	0 a 0,3 m	0,3 a 0,6 m	0,6 a 1,2 m	> 1,2 m
	Forma do Bloco (3) [x 3]	Tabular	Maciço, compacto	Compactos angulares	Arredondado
	Vegetação (3) [x 3]	Totalmente vegetado	Desigual	Plantas isoladas	Nenhuma
	Escore total do perigo (somatório até aqui)				
Tráfego	Distância de visibilidade	> 80%	60 a 80%	40 a 60%	< 40%
	Risco médio ao veículo	0 a 24%	25 a 49%	50 a 74%	> 74%
	Nº de acidentes	0 a 2	3 a 5	6 a 8	> 8
Escore total do risco (somatório total)					

- **Altura do talude**

Este parâmetro mede a altura do talude de corte ou a altura total do talude, considerando também a porção natural acima do corte, a depender de onde está a possível fonte de queda de blocos. Se houver uma zona de origem de queda de blocos no alto da encosta, além do corte, a altura total da encosta é medida. Se apenas o talude de corte estiver sendo avaliado, a altura máxima do corte é considerada.

- **Frequência de queda**

Este parâmetro pontua a frequência de queda de blocos que ocorre em um determinado local e baseia a pontuação na ocorrência do evento por tempo específico, ou seja, de quanto em quanto tempo esse evento ocorre. Para o Estado do Colorado, essas informações podem ser

obtidas com equipes de manutenção do Departamento de Trânsito (DOT), monitorando quantas vezes as rochas precisam ser removidas da estrada em um local específico. É importante destacar a dificuldade de se obter esse parâmetro com exatidão em locais cujo monitoramento desse tipo de problema não é constante e usual.

- **Escore da inclinação do talude**

Santi *et al.* (2009) modificaram este parâmetro, deixando de adotar o que era considerado no CRHRS original (Tabela 2.7), que pontuava o ângulo de inclinação com base na premissa de que ângulos maiores geram maiores problemas de instabilidade (Andrew, 1994; CDOT, 1997).

Maerz *et al.* (2005) conduziram algumas pesquisas utilizando simulações de quedas de blocos em softwares específicos e concluíram que quedas de blocos originadas em taludes com inclinações entre 30° e 45° e próximas a 85° alcançaram maiores distâncias, o que pode ter consequências mais severas por ter mais chances de atingir a rodovia.

Portanto, a pontuação do ângulo de talude foi alterada para refletir as chances dos blocos de rocha atingirem a estrada, dada a trajetória do bloco para vários ângulos de talude simulados por Maerz *et al.* (2005). Ou seja, enquanto o CRHRS original tratava o ângulo de inclinação do talude como um fator que gera instabilidades, o CRHRS modificado passou a tratar o ângulo de inclinação como um fator que se relaciona com o alcance dos blocos. Quanto à relação com a instabilidade, o ângulo de inclinação é tratado indiretamente na orientação e persistência das discontinuidades, como será explicado mais adiante.

Essa pontuação é dada de acordo com um “ranking de consequência”, que nesse caso é a chance do bloco atingir a via, conforme o ábaco adaptado de Maerz *et al.* (2005), mostrado na Figura 2.9. Semelhante ao parâmetro de altura do talude, os ângulos do talude de corte ou do talude total (natural, acima do corte) devem ser medidos de acordo com a localização da fonte potencial de queda de blocos. Se houver fontes potenciais tanto no corte quanto no talude acima dele, considera-se a pior situação, ou seja, o ângulo que gerar o maior escore no CRHRS.

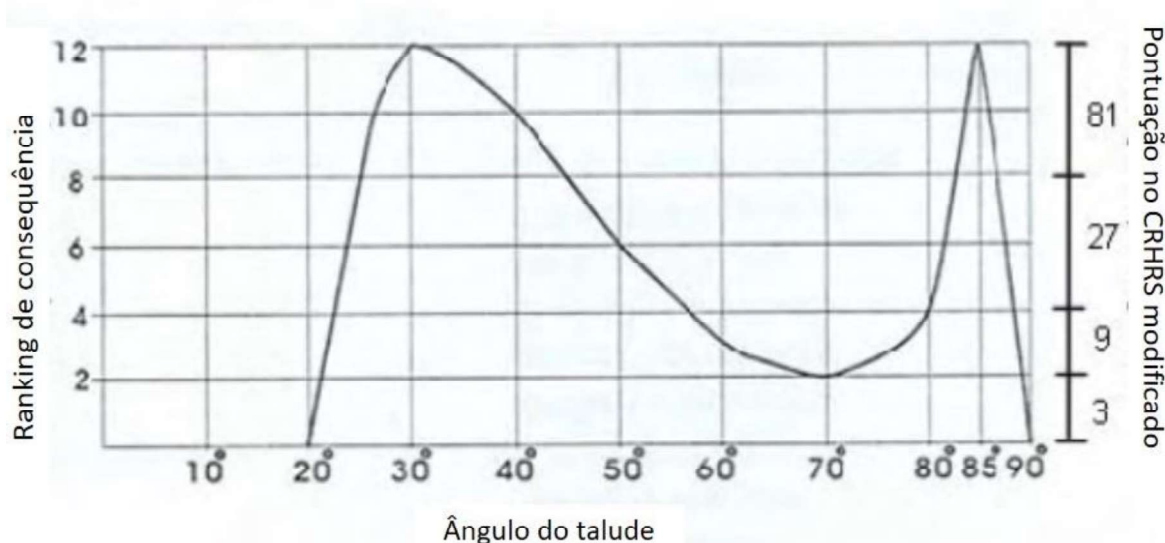


Figura 2.9 – Ábaco adaptado de Maerz et al. (2005) para determinar o escore referente à inclinação do talude pelo CRHRS modificado (adaptado de Santi et al., 2009).

- **Feições de lançamento**

Esse parâmetro mede basicamente a presença e magnitude das irregularidades na superfície do talude, o que pode modificar a trajetória de um bloco fazendo com que o mesmo se projete para maiores distâncias ao colidir com essas feições. Abaixo segue a descrição do que consta na Tabela 2.8 e, logo após, algumas imagens para exemplificar (Figura 2.10).

- Nenhum: talude suave, com pouca ou nenhuma variação topográfica ao longo da superfície.
- Menores: pequenas variações topográficas no perfil do talude que podem causar o lançamento de rochas, como a presença de cicatrizes ou bancos que se estendem até 0,6 m da superfície do talude.
- Maiores: variações topográficas maiores no perfil do talude que podem causar o lançamento de blocos, como a presença de cicatrizes ou bancos que se estendem entre 0,6m a 1,8 m da superfície do talude.
- Extremas: talude altamente irregular com grandes feições (cristas ou bancos) que se estendem por mais de 1,8 m da superfície do talude.



Figura 2.10 – Feições de lançamento, onde: a) Nenhuma; b) Menor; c) Maiores; d) Extremas (modificado de Russel et al., 2008).

- **Área de captação**

O RHRS modificado avalia a eficácia da área de captação tanto em termos de dimensões, quanto em relação à forma da mesma. Além do mais, a avaliação é feita por critérios matemáticos, menos subjetivos do que o RHRS original, onde a análise da eficácia era dada de forma basicamente qualitativa. A eficácia da área de captação quanto às dimensões é avaliada comparando as dimensões reais da estrutura com as dimensões recomendadas por Ritchie (1963), através de seu ábaco (Figura 2.11), utilizando a Equação 2.19. O resultado da equação é expresso em porcentagem.

$$Eficácia da área de captação = 100 \times \frac{(Dm+Wm)}{(Dr+W_r)} \quad (2.19)$$

Onde:

Dm: profundidade medida da vala;



Wm: largura medida da vala;

Dr: profundidade recomendada pelo Ábaco de Ritchie (1963) para a vala;

Wr: largura recomendada pelo Ábaco de Ritchie (1963) para a vala.

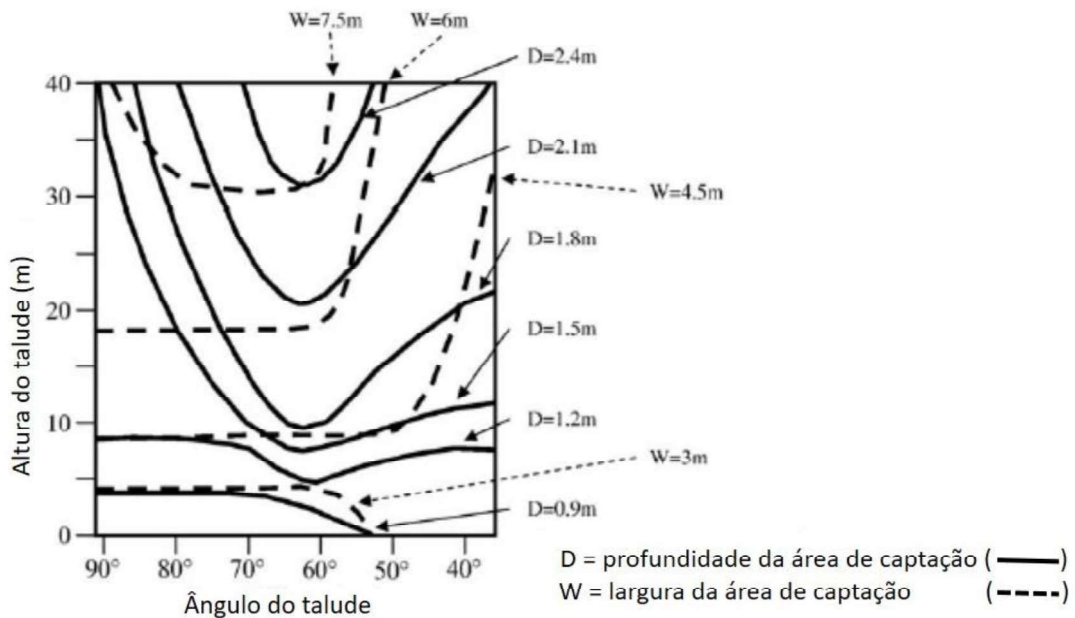


Figura 2.11– Ábaco de Ritchie (1963) para projetar as dimensões das áreas de captação de blocos.

Já a eficácia da área de captação quanto à forma é avaliada com base no ângulo de inclinação da vala ou na presença de quaisquer barreiras (barreira Jersey, cercas de contenção de quedas de blocos e guarda-corpo) entre a vala e a estrada. A Tabela 2.9 mostra a classe e pontuação correspondente à forma da área de captação. Nota-se que quanto maior o ângulo de inclinação para dentro da vala, menor a chance do bloco não ser contido caso caia nessa área de captação, logo menor o escore atribuído.

Tabela 2.9– Eficácia da área de captação quanto a forma da vala (Adaptado de Santi et al., 2009).

Eficácia da área de captação quanto à forma				
Classe	1	2	3	4
Inclinação da vala/ barreira	> 30° ou presença de barreira	21° a 30°	11° a 20°	0° a 10°
Escore CRHRS	3	9	27	81

Avaliados e atribuídos os escores para as dimensões e para a forma da área de captação, deve-se tomar aquele que for maior, ou seja, a situação mais desfavorável à segurança. Ainda nesse aspecto, Barrett e White (1991), sugerem que taludes íngremes e severamente irregulares do Colorado podem tornar os projetos de áreas de captação baseados no ábaco de Ritchie inaplicáveis. Portanto, Santi *et al.* (2009) consideram para o CRHRS modificado que se o escore para as feições de lançamento for 81 para um determinado talude em análise, a eficácia da área de captação deve receber automaticamente uma pontuação de 81.

- **Precipitação anual**

Parâmetro autoexplicativo, a precipitação anual para o CRHRS modificado especifica as quantidades reais de chuva e neve ao longo de um ano, a fim de remover a subjetividade dos critérios “baixo, moderado, alto” originalmente utilizados pelo CRHRS.

- **Ciclos de gelo e degelo anuais**

Santi *et al.* (2009) incluíram o número de ciclos de congelamento/descongelamento no CRHRS modificado devido aos seus efeitos documentados sobre queda de blocos e estabilidade de taludes.

O índice de gelo/degelo é definido por Lienhart (1988) como o produto da porcentagem mensal de dias com precipitação superior a 0,25 mm e o número de dias do ciclo de congelamento. Os dias do ciclo de congelamento são definidos como o número médio anual de dias em que a temperatura diária oscila acima e abaixo do ponto de congelamento (Lienhart, 1988).

Arnold *et al.* (1996) registraram o número de ciclos de gelo/degelo usando o índice de Lienhart (1988) em 5000 locais, tomando a temperatura média, precipitação média e condições de umidade acima de 0,25mm durante um período de um ano. Esta informação foi usada para classificar a distribuição geográfica dos ciclos de gelo/degelo anuais para Colorado (Figura 2.12). As divisões de pontuação no RHRS modificado são baseadas nesses dados.

Ciclos de gelo/degelo anuais médios para o Colorado.

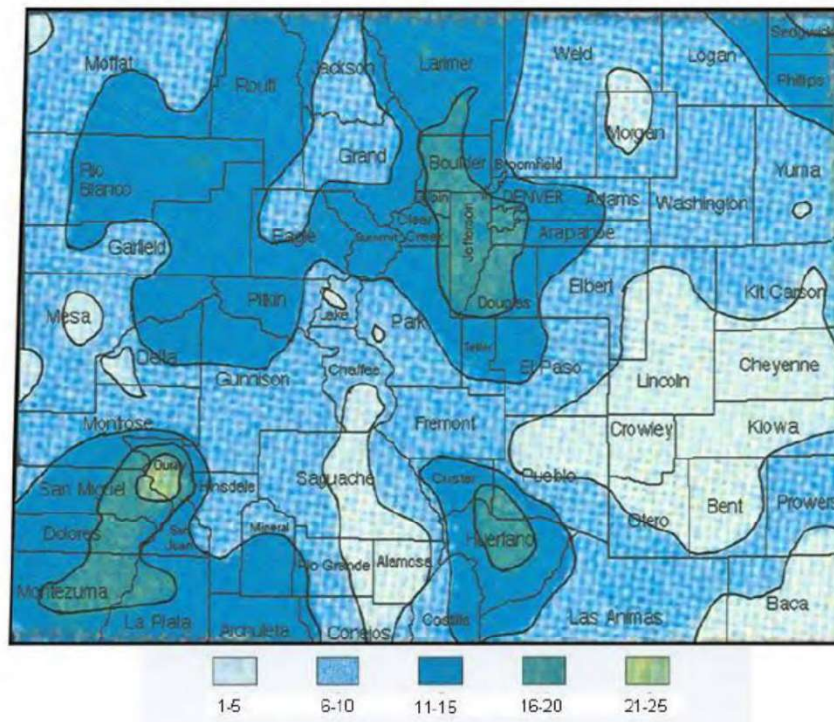


Figura 2.12 – Distribuição anual média dos ciclos de congelamento e descongelamento para o Estado do Colorado. O mapa foi derivado de Arnold et al. (1996) (Adaptado de Russel *et al.*, 2008).

- **Presença de água**

Também é um parâmetro autoexplicativo. A infiltração, ou a presença de água no talude, foi alterada para adjetivos mais descritivos, em vez do original "nenhum, algum, moderado, alto". No caso de a classificação ser realizada durante uma estação seca, as taxas de infiltração ainda podem ser estimadas observando faixas de água e zonas de descoloração nas faces do maciço, o que significa que a infiltração sazonal ocorre.

- **Aspecto do talude**

De acordo com Santi *et al.* (2009), o “aspecto do talude” é um parâmetro novo no CRHRS e é baseado na evidência de que as encostas do Colorado voltadas para o sul (S) experimentam muito mais ciclos de congelamento e degelo anualmente do que as encostas voltadas para o norte (N). As encostas voltadas para o norte ficam sombreadas a maior parte do tempo, portanto, sofrerão menor variação de temperatura ao longo de um determinado dia. Com isso, terão menos problemas relacionados a ciclos de gelo e degelo, recebendo o escore mais baixo, que é 3.

Os taludes voltados para leste (E), oeste (W), nordeste (NE) e noroeste (NW) recebem um pouco mais de sol durante o dia do que os taludes voltados para norte, recebendo então o segundo menor escore (9). Já os taludes voltados para sudeste (SE) e sudoeste (SW) recebem mais sol que os citados anteriormente, o que começa a deixar a situação mais crítica quanto aos ciclos de gelo e degelo. Sendo assim, esses taludes recebem o escore 27 para esse parâmetro de “aspecto do talude”. Encostas voltadas diretamente para o sul experimentam o maior número de variações de temperatura em um período de 24 horas e, portanto, recebem o escore 81.

- **Características geológicas**

Esse grupo de parâmetros considera três tipos de taludes que devem ser analisados separadamente devido às suas características particulares que influenciam no processo de queda de blocos. São eles: taludes em rochas cristalinas, onde a heterogeneidade do maciço e as fraturas tendem a controlar a queda de blocos; taludes em rochas sedimentares, cujo processo de queda de blocos tende a ocorrer influenciado pela erosão diferencial das camadas; taludes de blocos em matriz de solo (ou colúvio, detritos, depósitos glaciais, etc.), onde a erosão da matriz acaba provocando a movimentação dos blocos de rocha.

Como pode ser visto na Tabela 2.8, cada um desses três tipos de taludes possui três parâmetros individuais, ou seja, parâmetros que somente são observados e pontuados para esse determinado tipo de talude (ex: o parâmetro *jar slake* é pontuado apenas quando se trata de taludes em rochas sedimentares). Além desses três parâmetros característicos e próprios de cada tipo de talude, também são observados seis parâmetros relacionados às condições das discontinuidades. Esses parâmetros são observados apenas para taludes cristalinos e sedimentares; não são incluídos para aqueles cujos blocos de rocha estão em matriz de solo.

Sendo assim, quanto às características geológicas, tanto os taludes cristalinos quanto os sedimentares consideram e pontuam nove parâmetros (três próprios para cada tipo e seis para as discontinuidades). Já os taludes de blocos em matriz de solo, somente três parâmetros são pontuados quanto à geologia. Com isso, para classificar os tipos de talude igualmente, os escores dos parâmetros geológicos dos taludes de blocos em matriz são multiplicados por três (identificado na Tabela 2.8 como [x3]). Este multiplicador é usado para evitar que encostas de

rocha tenham mais peso do que encostas de bloco em matriz, quando ambas poderiam ter o mesmo potencial de queda de rochas (Santi *et al.*, 2009).

- **Taludes cristalinos: características do maciço**

De acordo com Santi *et al.* (2009) a queda de blocos em rochas cristalinas é amplamente controlada pela homogeneidade geral do maciço rochoso. Esse parâmetro lida com características específicas de rochas metamórficas e ígneas do Colorado. Abaixo segue a descrição do que é mostrado na Tabela 2.8.

Um maciço homogêneo implica que há pouca variação mineralógica ou litológica. Um maciço rochoso com pequenas falhas e/ou veios fortes é autoexplicativo, no entanto, deve-se notar que os veios incluídos nesta categoria causam pouca perda de resistência do maciço rochoso, ou seja, não são zonas importantes de fraqueza, pois existem em uma forma cristalizada. Zonas de xisto e/ou cisalhamento < 15 cm significa que há zonas de xistosidade ou outras texturas no maciço que podem contribuir para que haja instabilidades. Essas pequenas zonas de cisalhamento enfraquecem o maciço e criam caminhos para aumentar o intemperismo.

Maciços com forte presença de pegmatitos, ou micas fracas, ou zonas de cisalhamento > 15 cm parece ser o fator dominante de causas de instabilidade em grande escala e rupturas em taludes de rocha cristalinas em rodovias do Colorado com base em eventos anteriores. Por exemplo, a ruptura de 2005 na U.S. 6 perto de Golden, CO, resultando no fechamento da rodovia por vários meses, foi o resultado dessa condição (Santi *et al.*, 2009).

É possível notar certo cuidado dos autores ao se referirem a essas características, deixando claro sempre que são condicionadas às características geológicas dos maciços rochosos cristalinos do Colorado. Isso se deve ao fato de que o CRHRS modificado foi proposto e testado por eles apenas nas rodovias do Colorado. No entanto, é importante ressaltar que ao observar as características descritas nesse parâmetro, pode-se perceber que são generalistas, ou seja, que mostra aspectos comuns a rochas ígneas e metamórficas, não somente para o Colorado, o que permite adotar essas características em outros locais, desde que haja discernimento, critério e bom senso, sobretudo se for preciso fazer alguma adaptação.

- **Taludes cristalinos: grau de descalçamento**

Quedas de blocos frequentes do talude podem gerar feições salientes, descalçando blocos e, portanto, criando novas condições de instabilidade. A erosão diferencial de zonas de cisalhamento e xistos mais fracos também podem causar descalçamento em rochas cristalinas. Portanto, o grau de descalçamento é avaliado para taludes cristalinos no CRHRS modificado, tomando como base as dimensões desse descalçamento, semelhante ao que é analisado nas feições de lançamento, porém aqui, o problema está relacionado à instabilização do talude, visto que novos blocos salientes ficarão suspensos, não ao alcance dos blocos (Figura 2.13).



Figura 2.13 – Exemplo de um talude cristalino que pontua 81 para o grau de descalçamento no CRHRS modificado (Russel *et al.*, 2008).

- **Taludes cristalinos: grau de alteração**

Definições específicas e autoexplicativas estão na Tabela 2.8 como critérios de observação do grau de alteração em rochas cristalinas. Deve-se notar que este parâmetro leva em consideração o grau de alteração da matriz rochosa, não o grau de alteração ao longo das descontinuidades. Essas são classificadas em outra categoria.

- **Taludes sedimentares: grau de descalçamento**

Parâmetro muito semelhante ao grau de descalçamento para as rochas cristalinas, no entanto, agora originado pela erosão diferencial, que é uma das causas básicas de problemas de quedas de blocos em taludes de rochas sedimentares. Assim como nos taludes cristalinos, os escores são dados de acordo com as dimensões desse descalçamento, visto que quanto maior a porção instável, mais a queda do bloco depende de outras condições estruturais; ou maiores as chances desse bloco que se comporta como uma “laje suspensa” se desprender. A Figura 2.14 mostra um exemplo.



Figura 2.14– Exemplo de descalçamento em talude de rocha sedimentar (Russel *et al.*, 2008).

- **Taludes sedimentares: *jar slake***

Diante da influência do grau de descalçamento em problemas com quedas de blocos em taludes sedimentares e sabendo que a causa básica é o fato de se ter uma camada mais fraca intercalada com uma camada mais forte em relação ao intemperismo, é recomendável que se faça testes de durabilidade nessas camadas (*Slake Durability Test*) (Senior, 1999; Shakoor, 2005). No entanto, esse ensaio não é passível de execução rápida em campo.

Com isso, Santi *et al.* (2009) sugerem que testes simples de 30 minutos do *Jar Slake* podem ser realizados, ou seja, ao mergulhar amostras de rochas em água por 30 minutos é possível observar se há degradação da amostra e o quanto se degradou através da classificação de 1 a 6 mostrada na Figura 2.15. A pontuação para esse parâmetro no CRHRS modificado (Tabela 2.9) é dada de acordo com essa classificação.



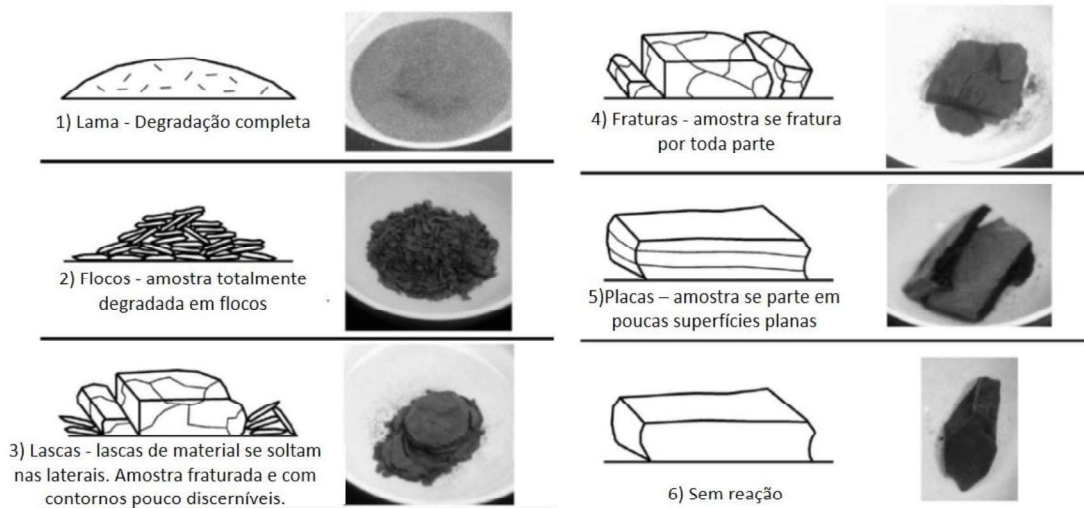


Figura 2.15 – Classes de reação de amostras de rochas submetidas ao Jar Slake (Russel *et al.*, 2008).

- **Taludes sedimentares: camadas**

De acordo com Vandewater *et al.* (2005), a severidade da queda de blocos e o tipo de queda é amplamente dependente do grau de variação litológica e da espessura dessas camadas em um talude rochoso sedimentar. Com isso, Santi *et al.* (2009) incluíram grau de variação litológica (parâmetro “camadas” da Tabela 2.8) no CRHRS modificado. As principais características consideradas são o número de camadas fracas no talude e suas respectivas espessuras. A Figura 2.16 ilustra a situação.

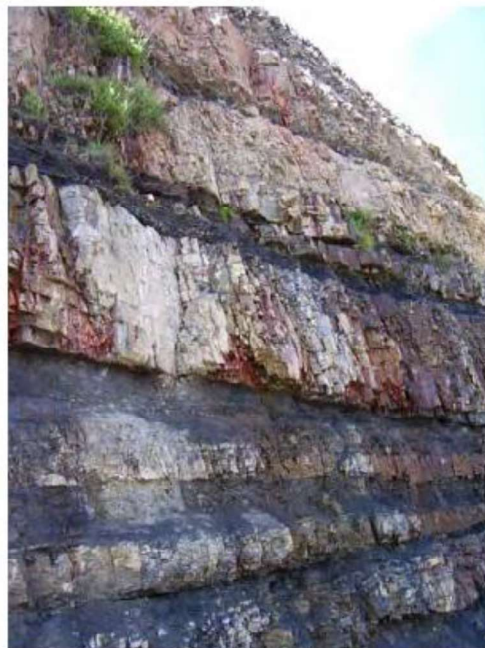


Figura 2.16– Exemplo de variação litológica entre camadas de um talude sedimentar (Russel *et al.*, 2008).



- **Descontinuidades: número de famílias**

O número famílias de descontinuidades foi adicionado ao CRHRS modificado devido ao fato de que quanto mais fraturado é o maciço, maior tende a ser a influência de infiltrações de água (ou gelo também, a depender das regiões), o que reflete diretamente nas condições de alteração e preenchimento dessas fraturas. Além do mais, quanto mais famílias, maiores também as chances de elas se cruzarem, a depender da persistência, e formarem blocos instáveis.

- **Descontinuidades: tamanho de blocos**

O tamanho do bloco fornece informações sobre o espaçamento das descontinuidades, uma vez que o tamanho do bloco é controlado por esta característica (Brown, 1981). Eventos de queda de blocos podem ser caracterizados por blocos únicos, com determinadas dimensões, ou por um volume de material de tamanhos variados. A categoria apropriada deve ser usada para o evento que parece ocorrer com mais frequência ou é mais provável de ocorrer em um determinado local. No entanto, embora o volume de queda de blocos esteja incluído nesta classificação (Tabela 2.8), o CRHRS modificado tem por objetivo, a princípio, classificar eventos individuais de queda de blocos e não deslizamentos de grandes volumes.

O parâmetro de tamanho/volume do bloco é importante também devido à relação com o grau de gravidade de um evento de queda de blocos, visto que blocos maiores possuem mais energia cinética quando caem e, portanto, são mais propensos a rolar mais longe quando alcançam a base da encosta e acabar na estrada, sobretudo em taludes que não são verticais (Senior, 1999; Maerz et al., 2005; Pierson e Van Vickle, 1993). Blocos maiores também causarão mais danos durante uma colisão com um veículo. Além disso, blocos maiores em queda têm maior probabilidade de instabilizar outros blocos e resultar em uma queda secundária ao colidir com a superfície do talude.

- **Descontinuidades: persistência e orientação**

Neste parâmetro é analisado o quanto as famílias de descontinuidades presentes no talude possuem uma orientação adversa em relação à orientação do talude. É observado aqui se as descontinuidades presentes se cruzam entre si formando blocos instáveis. Com isso, indiretamente se leva em consideração também a orientação do talude. Além da orientação, é

necessário também verificar a persistência, visto que descontinuidades menos persistentes terão menos chances de se cruzarem mesmo que a orientação seja adversa.

A classificação de persistência e orientação de descontinuidades foi alterada no CRHRS modificado anexando números de Pierson e Van Vickle (1993) e Brown (1981), usados para definir persistência contínua ou descontínua (maior que 3 m e menor que 3 m, respectivamente). A descrição da orientação foi esclarecida para definir “adverso” como descontinuidades que mergulham para fora da face do talude e “favorável” como aquelas que mergulham para dentro da face do talude (Santi *et al.*, 2009). A Figura 2.17 mostra à esquerda um talude que pontua 9 no CRHRS modificado e à direita um que pontua 81, de acordo com as análises de Russel *et al.* (2008).



Figura 2.17– À esquerda: descontinuidades persistentes ( $> 3$  m), com uma direção de mergulho favorável em relação ao talude e à estrada (pontuação de 9). À direita: descontinuidades persistentes ( $> 3$  m) com uma direção de mergulho adversa em relação ao talude (Russel *et al.*, 2008).

Embora a princípio sejam feitas basicamente observações em campo, não há impedimentos para que sejam feitas também análises cinemáticas posteriores para comprovar o quanto essas famílias possuem orientações desfavoráveis à estabilidade. No entanto deve-se ressaltar que sistemas como o CRHRS são metodologias pensadas para se ter uma resposta rápida quanto ao perigo de queda de blocos, devido ao fato de que muitas vezes serão analisados muitos quilômetros de rodovias.

- **Descontinuidades: abertura**

De acordo com Santi *et al.* (2009), vários autores, como Senior (1999) e Maerz *et al.* (2005), recomendam incluir a abertura das descontinuidades no RHRS, pois quanto maior a abertura, maiores as chances de ocorrer infiltração de água, e alterações associadas. Valores numéricos simples e mensuráveis são usados para esta categoria baseado, em parte, em Brown (1981). Cabe destacar também que se as juntas estão abertas, o fator de atrito será inexistente ou diminuído, seja devido ao vazio criado, ou devido a materiais argilosos como preenchimento, por exemplo.

- **Descontinuidades: alteração e preenchimento**

A resistência das paredes das descontinuidades tem uma grande influência no potencial de queda de blocos e o intemperismo químico reduz as propriedades mecânicas ao longo das superfícies de descontinuidade, reduzindo a coesão e o atrito (Santi *et al.*, 2009). Além disso, evidências de intemperismo químico indicam presença de água ao longo das superfícies de descontinuidade, o que aumenta a poropressão e contribui para a ocorrência de queda de blocos (Anderson *et al.*, 1999).

Esse parâmetro é semelhante ao que é analisado no parâmetro “alteração em rochas cristalinas”; porém agora, as observações são nas paredes das descontinuidades e não na matriz rochosa em si. Devido ao fato de que o preenchimento tem forte relação com a alteração dessas paredes, esses dois fatores são analisados juntos, pois espera-se que se há uma quantidade considerável de preenchimento, é porque houve alteração. Se as paredes estão sãs, a probabilidade de haver preenchimento é menor.

- **Descontinuidades: rugosidade**

No CRHRS modificado, Santi *et al.* (2009) incluíram o atrito das descontinuidades (rugosidade) como uma categoria separada e modificadas de Brown (1981). As superfícies rugosas têm uma textura áspera quando sentidas à mão. Superfícies onduladas são aquelas em que as rugosidades são mais arredondadas e lisas. Planar significa que não existem rugosidades, a superfície é lisa ao toque. Superfícies estriadas são aquelas resultantes de movimentações passadas, o que acabou resultando em estrias nas paredes das descontinuidades. Nesse caso, a resistência ao cisalhamento é reduzida potencialmente ao valor residual.

- **Blocos em matriz: tamanho de blocos**

Embora blocos maiores representem mais perigo, é provável que blocos menores caiam com muito mais frequência em taludes cujos blocos estão encrustados em matriz de solo e/ou detritos (blocos em matriz), pois uma quantidade menor de material que compõe a matriz precisa ser erodida para desalojar um bloco de, por exemplo, 0,3 m de dimensão média do que um bloco de cerca de 1,5 m. Com isso, as pontuações não devem ser atribuídas simplesmente na existência de blocos maiores na encosta, mas sim atribuídas aos maiores blocos observados como os que aparentemente possuem maior potencial de serem desalojados da matriz.

- **Blocos em matriz: forma do bloco**

A forma do bloco tem um efeito direto sobre a capacidade de rolamento das rochas e, conseqüentemente, o alcance das mesmas. Uma matriz que contenha grandes blocos tabulares não representará um perigo tão grande quanto outra contendo rochas arredondadas (Vandewater et al., 2005). Sendo assim, baseado no que foi adotado na classificação do Tennessee (Vandewater et al., 2005), Santi *et al.* (2009) incluíram esse parâmetro no CRHRS modificado, para taludes do tipo blocos em matriz.

- **Blocos em matriz: vegetação**

O Idaho DOT (Miller, 2003) inclui um parâmetro para a vegetação em seu sistema RHRS ao classificar o potencial de deslizamento ou grandes fluxos de detritos nas encostas do solo, visto que a presença de vegetação melhora a estabilidade de taludes de solo ao reduzir a erosão. O CRHRS modificado também contabiliza esse parâmetro, semelhante ao Idaho-RHRS, para avaliar os efeitos da vegetação nesses materiais (Santi et al., 2009), conforme é mostrado na Tabela 2.8, onde quanto mais vegetado é o talude, menor é o escore atribuído.

- **Escores de perigo, tráfego e escores de risco**

Até então, foram descritos os parâmetros do CRHRS modificado que, somados, geram a pontuação (escore) de perigo para um determinado talude analisado. Sendo assim, o perigo de queda de blocos é avaliado nesse sistema considerando, além das condições geológicas e geotécnicas do talude, as condições geométricas do talude, bem como a presença e eficácia da área de captação.

Ao inserir os parâmetros relacionados às condições de tráfego, se pode ter uma ideia da exposição dos usuários às essas condições de perigo, sendo uma forma de se mensurar a consequência do evento, que no caso, são os acidentes causados por quedas de blocos. De acordo com Santi *et al.* (2009), esses parâmetros permanecem inalterados em relação à atualização anterior do CRHRS, com exceção da classificação de risco médio do veículo em vez do tráfego médio diário. Com isso, o risco de queda de blocos é dado pela soma do perigo com os parâmetros de tráfego do CRHRS e explicados em seguida.

- **Tráfego: distância de visibilidade**

O percentual de distância de visibilidade é dado pela relação (divisão) entre a distância de visibilidade medida para tomada de decisão, pela distância requerida para tomada de decisão. Essa relação é multiplicada então por 100. A distância de visibilidade medida para tomada de decisão é definida como a distância em que um objeto de 15 cm colocado na beira da estrada é visível para um motorista (Pierson e Van Vickle, 1993). A distância de visibilidade requerida está relacionada com o limite de velocidade da via e é mostrada na Tabela 2.10.

Tabela 2.10– Distância de visibilidade requerida de acordo com os limites de velocidade da via (Pierson e Van Vickle, 1993).

Limite de velocidade (km/h)	Distância requerida (m)
40	114
48	137
56	160
64	183
72	206
80	229
89	267
97	305
105	1320

- **Tráfego: risco médio ao veículo (RMV)**

Esse parâmetro está relacionado com a quantidade de tempo em que um veículo está exposto à uma condição perigosa. Leva em consideração o fluxo médio diário (FMD) em carros/dia, o limite de velocidade da via (LV) em km/h e a extensão do talude (EXT) em km, como mostra a equação 2.20. O resultado é expresso em porcentagem.

$$RMV = 100 \times \frac{(FMD \times EXT)/24}{LV} \quad (2.20)$$

A equação para RMV pode gerar números maiores que 100, o que indica que há mais de um veículo na extensão considerada do talude em um determinado momento. Deve-se deixar claro que a equação acima não leva em consideração veículos que viajam significativamente mais devagar do que o limite de velocidade da via, como pode acontecer em mau tempo ou durante condições de tráfego pesado.

- **Tráfego: número de acidentes**

Este é um parâmetro utilizado no CRHRS modificado e foi obtido através de um banco de dados do departamento de trânsito do Colorado (CDOT) que listou todos os acidentes causados por "rochas na estrada" de 1º de janeiro de 1976 a 31 de dezembro de 2004, permitindo um somatório dos acidentes relacionados com queda de blocos em um determinado local.

No entanto, de acordo com Russel *et al.* (2008), esse parâmetro é problemático, não possui acurácia, visto que esses acidentes relatados como relacionados a queda de blocos podem ter sido causados tanto devido à instabilidade de taludes, quanto por rochas que caíram de caminhões, por exemplo. Ou seja, não se sabe com certeza a origem dos blocos que causaram esses acidentes.

## **2.5. TÉCNICAS DE ESTATÍSTICA MULTIVARIADA**

### **2.5.1. Conceitos iniciais e viabilidade de aplicação das técnicas**

A estatística multivariada trata de um conjunto de métodos estatísticos onde mais de uma variável é avaliada de forma simultânea, para cada indivíduo de um determinado grupo amostral, ou banco de dados. De acordo com Mingoti (2013), conforme também citado por Pereira (2015), as variáveis consideradas para análises multivariadas devem ser aleatórias e interrelacionadas, de modo que seus efeitos sejam mais evidentes quando elas são avaliadas em conjunto.

Antes de se aplicar qualquer técnica de estatística multivariada, deve-se analisar a adequabilidade do banco de dados para tal, ou seja, é necessário fazer uma análise prévia desse banco, onde um dos primeiros requisitos desejáveis é que o número de amostras deve ser ao menos cinco vezes maior que o número de variáveis (Pereira, 2015). Além do mais, o banco de dados deve apresentar variabilidade e as variáveis devem ser correlacionadas (Mingoti, 2013).

Alguns testes podem ser aplicados ao banco de dados para verificar a adequabilidade para a aplicação das técnicas de estatística multivariada em geral. Outros testes complementares podem ser aplicados para verificar se o banco de dados atende a certos requisitos para uma determinada técnica específica, que pode exigir ou recomendar a aplicação quando há normalidade multivariada dos dados ou homocedasticidade, por exemplo. Dentre esses testes, pode-se citar o Teste de Esfericidade de Bartlett, Teste de Normalidade Multivariada e Teste M Box. Quanto à variabilidade dos dados, essa pode ser verificada através de análise descritiva das variáveis.

O teste de esfericidade de Bartlett (Bartlett, 1951) é aplicado com a finalidade de se verificar se existem correlações significativas entre as variáveis, sendo a correlação uma medida padronizada entre -1 e 1. Se o valor é próximo de 0, significa que a correlação entre as variáveis é baixa; se for próximo de 1, a relação é fortemente positiva; se for próximo de -1 a relação é fortemente negativa entre as variáveis. Dentre os tipos de coeficientes de correlação, destacam-se os Coeficientes de Correlação Linear de Pearson e Coeficientes de Correlação Linear de Spearman.

O Coeficiente de Correlação Linear de Pearson é indicado quando as variáveis são quantitativas e foi proposto por Francis Galton e Karl Pearson em 1897. O estimador desse coeficiente ( $\rho_{ij}$ ) pode ser calculado através da Equação 2.21 e, de acordo com Callegari-Jaques (2003), a classificação da correlação entre duas variáveis pode ser avaliada conforme a Tabela 2.11.

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} \sigma_{jj}}} \quad (2.21)$$

Onde:

$\sigma_{ij}$ : covariância entre as variáveis  $i$  e  $j$ .

$\sigma_{ii}$ : variância da variável  $i$ .

$\sigma_{jj}$ : variância da variável  $j$ .

Tabela 2.11- Classificação dos valores do coeficiente de correlação linear de Pearson (Callegari-Jaques, 2003).

Coeficiente de Correlação	Classificação
$0 < \rho < 0,30$	Fraca correlação linear
$0,30 < \rho < 0,60$	Moderada
$0,60 < \rho < 0,90$	Forte
$0,90 < \rho < 1,00$	Muito forte

É importante destacar, apenas para conhecimento, que na estatística univariada, o símbolo  $\sigma$  é usual para designar o desvio padrão, enquanto a variância é normalmente designada por  $\sigma^2$ . No entanto, na estatística multivariada, é comum utilizar o símbolo  $\sigma_{ii}$  para representar a variância de uma variável  $i$  e o símbolo  $\sigma_{ij}$  para representar a covariância entre as variáveis  $i$  e  $j$ .

Para o caso de variáveis ordinais, se usa o Coeficiente de Correlação por Postos de Spearman, visto que para esse tipo de variável as correlações não podem ser calculadas de modo análogo às correlações de Pearson (Bunchaft & Kellner, 1999). De acordo com Lira (2004), o Coeficiente de Correlação de Spearman usa a ordem das observações, não o valor observado. Deste modo, este coeficiente não é sensível a assimetrias na distribuição, nem à presença de outliers, não exigindo, portanto, que os dados sejam normais. O estimador do coeficiente ( $\rho_s$ ) é mostrado na Equação 2.22.

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)} \quad (2.22)$$

Onde:

$d_i$ : diferença entre as ordenações (postos de  $x_i$  – postos de  $y_i$ );

$n$ : número de pares de ordenações ( $x_i, y_i$ ).

Ainda de acordo com Lira (2004), a princípio essas correlações não mostram tendência linear, mas podem ser consideradas como índices de monotonicidade, que significa que para valores



positivos da correlação, aumentos no valor de X correspondem a aumentos no valor de Y, e para coeficientes negativos vale o oposto.

Conhecidas as formas de se calcular o coeficiente de correlação tanto para variáveis quantitativas (Pearson) quanto para variáveis qualitativas (Spearman), é possível então aplicar o Teste de Esfericidade de Bartlett (Bartlett, 1951), que consiste basicamente em comparar a matriz com esses coeficientes de correlação (matriz de correlação das variáveis) com a matriz identidade (Equação 2.23).

$$\begin{bmatrix} 1 & \cdots & \rho_{1p} \\ \vdots & \ddots & \vdots \\ \rho_{p1} & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.23)$$

Onde:

$\rho_{ij}$ , em que  $i=1,2,\dots,p$  e  $j=1,2,\dots,p$ : correlação entre as variáveis  $i$  e  $j$ ,

$p$ : número de variáveis.

Nota-se, pela Equação 2.23 que o teste parte do pressuposto que a matriz de correlações entre as variáveis ( $\mathbf{R}$ ) é igual a matriz identidade ( $\mathbf{I}$ ) e essa se torna a hipótese básica ( $H_0$ ) que, se obedecida, indica que não há correlação entre as variáveis, visto que  $\rho_{ij}$  seria igual a 0. Caso contrário, ou seja, se essa hipótese  $H_0$  for rejeitada, indica que existe correlação suficiente e que o banco de dados é adequado para aplicação das técnicas de estatística multivariada. Em geral, é recomendável que as correlações sejam no mínimo 0,30.

A Equação 2.24 mostra a estatística do teste, em que  $n$  é o tamanho da amostra,  $p$  o número de variáveis e  $|\mathbf{R}|$  é o determinante da matriz de correlações. A estatística do teste apresenta distribuição qui-quadrado ( $X^2$ ) com graus de liberdade (df) conforme a Equação 2.25. Para que haja correlações significativas na matriz de correlação, com 95% de confiança, o p-valor deve ser inferior a 0,05.

$$X^2 = - \left[ (n - 1) - \frac{2p+5}{6} \right] \ln|\mathbf{R}| \quad (2.24)$$

$$df = \frac{p(p-1)}{2} \quad (2.25)$$

Diferente da presença de correlação entre as variáveis, requisito para a aplicação de qualquer técnica de análise estatística multivariada, a normalidade multivariada não é exigida para todas as técnicas. No entanto, para algumas, como a análise discriminante, essa normalidade é desejável, embora não seja obrigatória. Sendo assim, é recomendável que essa análise também seja feita durante a fase prévia à análise dos dados.

Royston (1983) propôs uma generalização do teste de normalidade univariada de Shapiro-Wilk para normalidade multivariada, onde se estima a estatística  $W$  de Shapiro-Wilk para cada variável e, no final, os valores são somados.

Considerando um banco de dados multivariado, representado pelo vetor aleatório  $\mathbf{X}$ , com  $n$  indivíduos e  $p$  variáveis, com matriz amostral de covariâncias  $\mathbf{S}$ , o teste de Shapiro-Wilk, define  $W$  de acordo com a Equação 2.26.

$$W = \frac{(\sum a_i v_i)^2}{\sum (v_i - \bar{v})^2} \quad (2.26)$$

Em que:

$v_i$ , onde  $i=1,2,\dots,n$ : representam os dados de uma variável, organizados em ordem crescente de magnitude,

$$\bar{v}: \sum \frac{v_i}{n}$$

$a_i$ , onde  $i=1,2,\dots,n$ : são os melhores coeficientes lineares não enviesados de Sarhan & Greenberg (1956).

Para amostras de tamanho  $n$  entre 7 e 2000,  $W$  pode ser transformada em uma variável normal padronizada,  $z$ , de acordo com Royston (1983) (Equação 2.27). O valor de  $n$  entre 7 e 2000 se justifica porque grandes valores positivos de  $z$  são evidência de não normalidade.

$$z = \frac{(1-W)^\lambda - \bar{x}_y}{s_y} \quad (2.27)$$

Onde:

$\lambda$ : coeficiente estimado por 50 amostras e suavizadas utilizando o polinômio  $\ln(n) - d$ , sendo  $d = 3$  para  $7 \leq n \leq 20$  e  $d = 5$  para  $21 \leq n \leq 2000$ ;

$\bar{x}_y$ : média amostral e

$s_y$ : desvio padrão amostral.

Com isso, para testar a normalidade multivariada dos dados, deve-se calcular a estatística univariada de Shapiro-Wilk,  $W_j$ , de modo análogo à Equação 2.27, mas para cada variável  $j$ , tal que  $j=1, 2, \dots, p$  e depois normalizá-la (Equação 2.28).

$$z = \frac{(1-W_j)^{\lambda} - \bar{x}_j}{s_j} \quad (2.28)$$

Feito isso, calcula-se posteriormente a estatística do teste de normalidade multivariada dada por  $R_j$ , tal que  $j=1, 2, \dots, p$  (Equação 2.29), onde  $\phi(x)$  é dado pela equação 2.30. A estatística do teste de Royston (1983) é finalmente dada pela Equação 2.31.

$$R_j = \left\{ \phi^{-1} \left[ \frac{1}{2} \phi(-z_j) \right] \right\}^2 \quad (2.29)$$

$$\phi_x = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x \exp\left(-\frac{1}{2}p^2\right) dt \quad (2.30)$$

$$H = \frac{e \sum R_j}{p} \quad (2.31)$$

Onde:

$e$ : constante denominada graus de liberdade equivalentes, onde  $e < p$ .

A estatística  $H$  apresenta distribuição qui-quadrado ( $\chi^2$ ) com graus de liberdade ( $df$ ) igual a  $e$ . A hipótese básica ( $H_0$ ) do teste é que os dados são normais multivariados, portanto, para aceitar essa hipótese  $H_0$  com 95% de confiança, o  $p$ -valor deve ser superior a 0,05.

Assim como a normalidade multivariada, a homocedasticidade pode ser requerida ou desejável em certas técnicas de estatística multivariada, como nas análises discriminantes lineares. Homocedasticidade é quando duas matrizes de covariâncias das variáveis independentes são iguais (homogêneas) para as populações consideradas ( $\Sigma_k$ ). O teste

utilizado para testar a homocedasticidade entre matrizes de covariância é o M Box, proposto por Box (1949).

Supondo a existência de  $k$  populações e que a hipótese básica ( $H_0$ ) a ser testada seja que as respectivas matrizes de covariância das populações sejam estatisticamente iguais, tem-se:

$H_0: \Sigma_1 = \Sigma_2 = \dots = \Sigma_k$ , existe homocedasticidade entre os dados;

$H_1: \Sigma_1 \neq \Sigma_2 \neq \dots \neq \Sigma_k$ , não existe homocedasticidade entre os dados.

O valor da estatística  $M$  é dada pela equação 2.32.

$$M = (n - k) \ln |\mathbf{S}| - \sum_{j=1}^k (n_j - 1) \ln |\mathbf{S}_j| \quad (2.32)$$

Onde:

$n$ : número total de informações;

$k$ : número de populações, com  $k=1,2,\dots,j$ ;

$n_j$ : uma observação do grupo  $j$ ;

$\mathbf{S}$ : matriz comum de covariâncias amostrais igual a  $\frac{\sum_{j=1}^k (n_j - 1) \mathbf{S}_j}{n - k}$ ;

$\mathbf{S}_j$  é a matriz de variâncias e covariâncias do grupo  $j$ .

A estatística do teste apresenta distribuição qui-quadrado ( $\chi^2$ ) com graus de liberdade igual a  $df$  e é definida pela Equação 2.33. Para que haja homocedasticidade entre as matrizes de covariância, com 95% de confiança, o p-valor deve ser superior a 0,05.

$$M(1 - c) \sim \chi^2(df) \quad (2.33)$$

Onde:

$$c = \frac{2p^2 + 3p - 1}{6(p+1)(k+1)} \left( \sum_{j=1}^m \frac{1}{n_j - 1} - \frac{1}{n - k} \right);$$

$$df = \frac{p(p+1)(k-1)}{2};$$

$p$  é o número de variáveis.

### 2.5.2. Análise de Agrupamento

A análise de agrupamentos, ou análise de *clusters*, se trata de um método estatístico utilizado para agrupar indivíduos cuja classificação, ou divisão em grupos, ainda é desconhecida, baseando-se apenas nas similaridades e dissimilaridades entre eles, sem definir qualquer critério previamente e sem requerer um conhecimento prévio de formação de grupos. Basicamente, é uma técnica de estatística multivariada que tenta organizar um conjunto de indivíduos em grupos relativamente homogêneos (Pereira, 2015).

Os métodos de agrupamentos podem ser classificados em diferentes categorias de métodos, sendo as mais populares e encontradas na literatura os métodos de agrupamento hierárquicos (MAH) e métodos de agrupamento não hierárquicos (MANH) (Tinsley & Steven, 2000). Os métodos hierárquicos (MAH) podem ser aglomerativos ou divisivos. Os MANH são também conhecidos como métodos de otimização.

No caso dos métodos aglomerativos, a princípio cada indivíduo pertence a um grupo individual e a partir daí, os indivíduos semelhantes se unem formando grupos, até a formação de um único grupo. Já nos métodos divisivos, o caminho é o oposto: parte-se de um dos extremos de classificação, que é considerar todos os indivíduos como pertencentes a um único grupo. A partir daí os indivíduos menos similares vão sendo desagrupados até que se atinja o outro extremo, que é cada indivíduo pertencer a um grupo individual.

Tanto os métodos aglomerativos como divisivos apresentam suas fusões (ou divisões) ao longo do agrupamento que pode ser apresentado através de um gráfico chamado dendrograma, sendo que essas fusões e divisões são baseadas em uma matriz de dissimilaridades **D**. Para aplicação do agrupamento pelos MAH, não é necessário que se conheça o número de grupos *k a priori* e uma vez que uma observação é aglomerada, ela nunca é realocada.

Os métodos não-hierárquicos (MANH) produzem uma partição em um número fixo de grupos, de modo iterativo, com sucessivas alocações e realocações das observações, com base em uma matriz de dissimilaridades entre as observações. Nos MANH as observações podem

ser realocadas entre os grupos ao longo do processo e, na maioria dos métodos, é necessário definir previamente o número de grupos  $k$ .

Cada indivíduo é inserido em um dos  $k$  grupos, de modo que as realocações desses indivíduos é cessada quando uma regra de parada pré-especificada com uma “função objetivo” for contemplada e isso ocorre quando é obtido o melhor agrupamento possível, de acordo com a regra predefinida. Por isso os MANH são considerados como métodos de otimização e não simplesmente de partição (Nazareth, 2020). O método de agrupamento utilizado nesse trabalho é não hierárquico, portanto, os MANH são mais detalhados adiante.

- **Tendência de Agrupamento**

Antes de se realizar uma análise de agrupamento, a depender dos objetivos da mesma, é prudente verificar se há tendência real de agrupamento, porque a partir do momento em que se inicia a análise a partir de um algoritmo qualquer, a mesma pode trazer alguma solução que não necessariamente seja um grupo natural, visto que de qualquer forma haverá medidas de proximidade entre os dados que podem particionar os indivíduos em grupos estatisticamente naturais ou não, que podem ser tecnicamente explicáveis ou não (Everitt *et al.*, 2011).

Ainda de acordo com Everitt *et al.* (2011), uma vez que não se sabe a priori a estrutura natural dos dados ou a quantidade de grupos, há o risco de se interpretar todas as soluções da análise de agrupamentos como sendo a existência de grupos (naturais) distintos, forçando uma estrutura não intrínseca aos dados ao invés de descobrir os agrupamentos naturais presentes.

Um método para identificar se existe tendência à formação de agrupamentos significativos, ou seja, em estruturas não aleatórias, é a Estatística de Hopkins (H) (Hopkins & Skellam, 1954). A estatística de Hopkins é usada para avaliar a tendência de agrupamento de um conjunto de dados medindo a probabilidade de que um determinado conjunto de dados seja gerado por uma distribuição uniforme de dados. Em outras palavras, ele testa a aleatoriedade espacial dos dados (Lawson e Jurs 1990).

Conforme explicado por Campos (2019), para a aplicação do teste, inicialmente são obtidas duas amostras,  $U$  e  $Q$ , a partir do conjunto original de dados  $Y$  de tamanho  $n$ , ambas contendo  $p$  elementos, de modo que  $p < n$ . No caso,  $U$  é composta por  $p$  pontos aleatoriamente gerados

no espaço, enquanto que  $\mathbf{Q}$  possui  $p$  pontos reais do conjunto original de dados. Com isso, calcula-se-se as distâncias entre os  $p$  elementos das amostras e seus respectivos vizinhos mais próximos no conjunto original  $\mathbf{Y}$ . A Estatística Hopkins ( $H$ ) para um determinado banco de dados com  $n$  unidades é calculada de acordo com a Equação 2.34.

$$H = \frac{\sum_{i=1}^p q_i}{\sum_{i=1}^p u_i + \sum_{i=1}^p q_i} \quad (2.34)$$

Onde:

$q_i$ : é a distância entre o  $i$ -ésimo ponto real pertencente a  $\mathbf{Q}$  e seu vizinho mais próximo em  $\mathbf{Y}$ .

$u_i$ : é a distância do  $i$ -ésimo ponto aleatoriamente gerado em  $\mathbf{U}$  a seu vizinho mais próximo em  $\mathbf{Y}$ .

Na Equação 2.34, para cada objeto, são consideradas as distâncias  $q_i$  e  $u_i$ . Quando os objetos estão em grupos bem definidos, coesos e bem separados, a distância média entre os objetos é pequena. Nesse caso o somatório de  $q_i$  tende a ser próximo de 0 e, conseqüentemente,  $H$  é próximo de 0. Alternativamente, pode-se colocar  $u_i$  no numerador e, nesse caso, a interpretação é inversa. Já em instâncias em que os objetos estão dispersos no espaço, os somatórios de  $q_i$  e  $u_i$  são próximos e o valor de  $H$  é próximo a 0,5. Em termos gerais: valores de  $H$  próximos de 0,5 indica a formação de agrupamentos aleatórios e não necessariamente significativos.

É importante salientar que a Estatística Hopkins diz respeito à tendência de clusters estatisticamente significativos, ou seja, cuja estrutura é não aleatória. No entanto, apenas o especialista que conduz a análise é capaz de rotular os grupos formados e julgar se esses grupos são tecnicamente explicáveis de acordo com sua área de atuação e com os objetivos da aplicação do agrupamento.

Um outro método para avaliar a tendência de agrupamento é o método visual do gráfico de calor (VAT) (Bezdek and Hathaway, 2002). De acordo com Kassambara (2017) e conforme descrito no site *Data Novia* também pelo professor Kassambara, o algoritmo segue os seguintes passos: primeiro é calculada a matriz de dissimilaridade entre os objetos utilizando a distância Euclidiana; depois essa matriz é reordenada de modo que os objetos similares fiquem próximos uns dos outros, dando origem à matriz de dissimilaridades ordenada; essa

matriz de dissimilaridades ordenada é exibida em forma de imagem, que é o VAT. Um exemplo de VAT é mostrado na Figura 2.18.

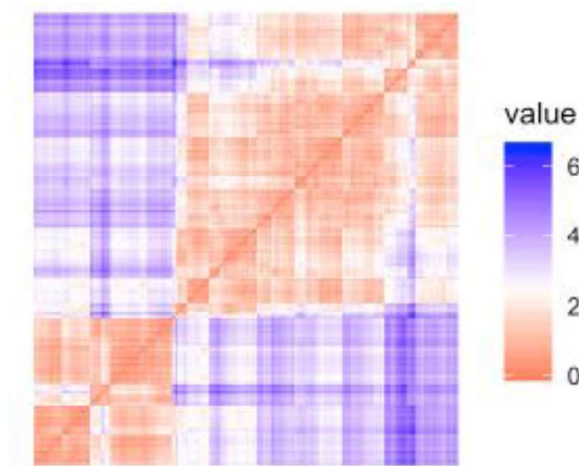


Figura 2.18 – Gráfico de calor (VAT) mostrando a tendência de agrupamento para formação de 2 grupos (Fonte: Data Novia).

Na Figura 2.18 é possível perceber dois retângulos vermelhos na diagonal. Esses retângulos indicam os dados ordenados cuja similaridade entre os objetos é alta. Isso indica que esses dois retângulos podem representar grupos. A cor azul significa baixa similaridade. O nível de cor é proporcional ao valor da dissimilaridade entre as observações, onde vermelho puro significa que a distância entre as observações é 0.

- **Número ótimo de clusters**

Dentre as diversas metodologias existentes para encontrar o número ótimo de clusters, neste trabalho serão apresentadas apenas as que foram utilizadas. Elas estão descritas de forma bastante didática, inclusive com os algoritmos apropriados para análise pelo software R, também utilizado nesse trabalho, no site *Data Novia*, pelo professor Alboukadel Kassambara, baseado em Kassambara (2017). Nazareth (2020) também explicou de forma objetiva e de fácil entendimento alguns desses métodos. Outras metodologias e índices são descritos também por Ferreira (2018) e Everitt et al. (2011).

É importante salientar que o número ótimo de clusters é subjetivo, pois além de depender do método utilizado, depende também do analista de dados, pois é este quem no final, a depender



dos objetivos da aplicação da análise de agrupamento, é que vai decidir se esse número ótimo apontado estatisticamente é capaz de gerar grupos explicáveis tecnicamente.

Os métodos para avaliar o número ótimo de clusters são aplicáveis tanto para os métodos hierárquicos aglomerativos (AGNES) e divisivos (DIANA), quanto para os métodos não hierárquicos, como k-means e k-medoides (PAM). Eles se dividem em métodos diretos, que consistem na otimização de um critério, como a soma dos quadrados dentro do cluster ou a silhueta média, sendo os métodos de cotovelo (*elbow*) e silhueta, respectivamente; e métodos estatísticos, que consistem em comparar a evidência com a hipótese nula. Um exemplo é Estatística Gap.

Além dos métodos de cotovelo, silhueta e Gap, existem mais de trinta outros índices e métodos que foram publicados para identificar o número ideal de clusters, conforme foi pontuado anteriormente. Diante disso, pode-se utilizar códigos em R para calcular todos esses índices, a fim de decidir o melhor número de clusters usando a "regra da maioria". Esse código está também disponível no site *Data Novia*. Nesse projeto, para determinar o número ótimo de clusters, foram utilizados os métodos do cotovelo, silhueta média e a regra da maioria.

O método do cotovelo relaciona o total da soma dos quadrados das distâncias intergrupos das observações aos respectivos centroides (*total within sum of square - WSS*), com número de grupos ( $k$ ). O número de grupos deve ser escolhido de modo que a adição de outro grupo não aumente consideravelmente o WSS. Para isso, primeiramente se roda um algoritmo de agrupamento nos dados para diferentes valores de  $k$ . Daí, para cada  $k$  é calculada a soma dos quadrados das distâncias (WSS) e então é gerada a curva WSS *versus*  $k$ . A localização da dobra, ou cotovelo, indica o número ótimo de grupos (Figura 2.19). Devido ao fato de que às vezes podem surgir dúvidas sobre qual dobra do gráfico se deve considerar, é recomendado utilizar alguma outra metodologia complementar.

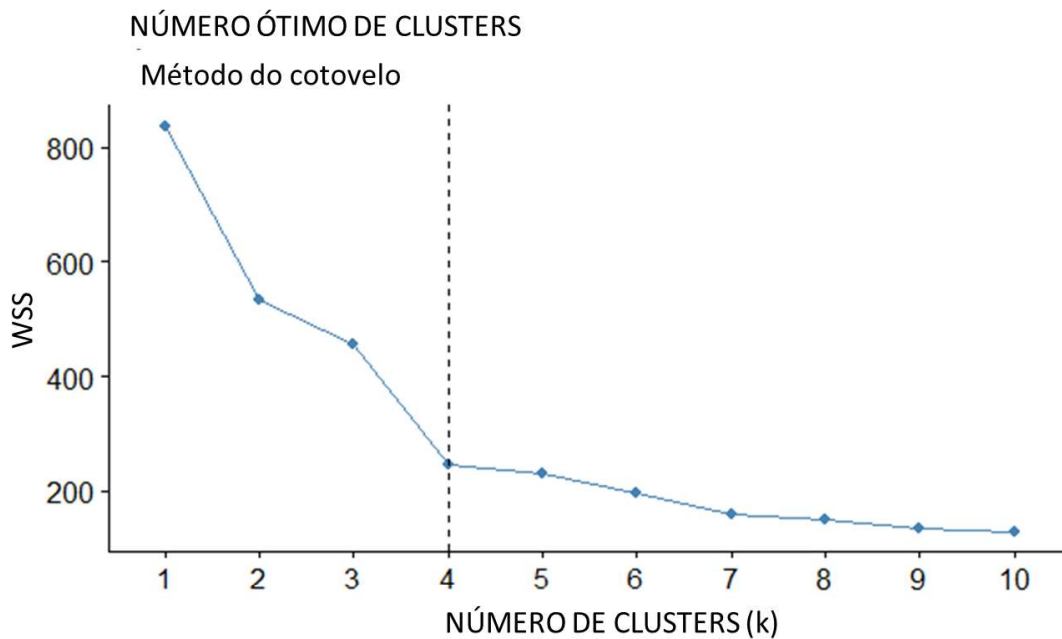


Figura 2.19– Gráfico para determinar o número ótimo de clusters pelo método do cotovelo.

A silhueta, de uma maneira geral, mede a qualidade de um agrupamento. Ou seja, ele determina o quão bem cada objeto se encontra em seu cluster. Para cada elemento  $i$  de um determinado grupo, é medida a menor distância média intercluster e intracluster. Também são medidas as maiores distâncias intercluster e intracluster. O valor da silhueta para um elemento  $i$  ( $S_i$ ) é dado através da Equação 2.35. O coeficiente de silhueta médio para um agrupamento com  $k$  grupos é dado pela média do valor de  $S_i$  sobre todos os  $n$  objetos do conjunto.

$$S_i = \frac{b_i - a_i}{\max(b_i, a_i)} \quad (2.35)$$

Onde  $b_i$  é a menor distância média entre o elemento  $i$  e todos os pontos dos demais grupos (menor distância média interclusters);  $a_i$  é a menor distância média entre o elemento  $i$  e os pontos dentro de seu grupo (distância intracluster);  $\max(a_i, b_i)$  é a distância máxima medida entre grupos e dentro do mesmo grupo.

Uma largura média alta de silhueta (próxima de 1) indica um bom agrupamento. Uma largura baixa (próxima de 0), indica que um indivíduo pode pertencer a dois grupos. Indivíduos com silhueta negativa podem ter sido alocados incorretamente em um determinado grupo.

Com isso, levando em consideração que a silhueta mede a qualidade do agrupamento, ela também pode indicar o número ótimo de grupos, observando a silhueta média para cada agrupamento com diferentes  $k$  grupos. Daí, o número ótimo de clusters  $k$  é aquele que maximiza a silhueta média em uma faixa de valores possíveis para  $k$  (Kaufman e Rousseeuw, 1990).

Assim como no método do cotovelo, também se roda um algoritmo de agrupamento para diferentes valores de  $k$ . Para cada um, se calcula a silhueta média e se gera um gráfico relacionando essa silhueta média com o valor de  $k$ . O  $k$  ótimo é aquele que possui a maior silhueta média.

Quanto à “regra da maioria”, foi utilizada a função `fviz_nbclust()`, do pacote `NbClust` (Charrad, et al., 2015) do software R, onde o número ótimo de grupos ( $k$ ) escolhido é aquele apontado como o ótimo pela maioria, dentre os 30 índices considerados nessas funções. O pacote varia todas as combinações de número de grupos, medidas de distâncias e métodos de agrupamentos possíveis.

- **Medidas de similaridade e dissimilaridade**

Em uma análise de agrupamento, os grupos são formados levando em consideração as semelhanças entre os indivíduos pertencentes a um mesmo grupo, ou considerando as diferenças entre indivíduos pertencentes a grupos distintos. Essa semelhança (ou diferença) é geralmente dada por uma medida de proximidade.

Segundo Ferreira (2018), observações muito semelhantes são ditas similares e aquelas menos semelhantes são denominadas dissimilares. Com isso, essas medidas de proximidade ou diferença podem ser dadas por medidas de similaridade e dissimilaridade, respectivamente. Para as medidas de similaridade, quanto maior o valor calculado, maior será a semelhança entre as observações e vice-versa. Já para medidas de dissimilaridade vale o oposto, ou seja, quanto maior a dissimilaridade, menor a semelhança entre as observações.

As medidas de dissimilaridade possuem diferentes formas de serem mensuradas. Em contrapartida, as medidas de similaridade são complementares às distâncias. Deste modo, quanto menor for a distância entre duas observações, maior será a medida de similaridade e

menor a de dissimilaridade entre elas. Como as medidas são complementares é possível migrar de um coeficiente para outro (Nazareth, 2020).

Adiante serão apresentados alguns coeficientes para se calcular a dissimilaridade e/ou a similaridade para o caso em que todas as variáveis são quantitativas (ou contínuas, ou métricas), visto que esse trabalho utiliza variáveis ordinais, que foram tratadas como quantitativas, conforme será explicado com mais detalhes no item que trata da metodologia desenvolvida. Para bancos de dados contendo variáveis qualitativas e variáveis mistas, as formas de se determinar o coeficiente de proximidade são descritas por Pereira (2015) e Nazareth (2020), que inclusive sugerem também outras referências importantes.

- **Medidas de dissimilaridade para dados quantitativos**

Para o caso em que todas as variáveis são contínuas, as medidas de dissimilaridade ( $\delta_{ij}$ ) são as mais usuais e para isso, se calcula as distâncias entre as observações. Uma medida  $d_{ij}$  representa uma distância se, e somente se:

- a) é positivamente definida,  $d_{ij} \geq 0$ , para todas as observações  $i$  e  $j$ ;
- b) é nula apenas para pontos coincidentes,  $d_{ij} = 0$ , se e somente se,  $i = j$ ;
- c) é simétrica,  $d_{ij} = d_{ji}$ , para todo  $i$  e  $j$ ;
- d) Obedece à desigualdade triangular,  $d_{ij} \leq d_{ik} + d_{kj}$ , para quaisquer observações  $k$ ,  $i$  e  $j$ .

Na literatura foram propostas várias medidas para determinar a matriz de dissimilaridades a partir de um conjunto de dados multivariados contínuos. Nazareth (2020) sumarizou em seu trabalho alguns dos mais utilizados. Ainda de acordo com Nazareth (2020), uma lista mais extensa pode ser encontrada em, Jajuga, *et al.* (2003), apud Everitt, *et al.* (2011), e Gower & Legendre (1986).

Apesar de as distâncias serem mais usuais para medidas de dissimilaridade, medidas de correlação também são encontradas na literatura. Para os casos em que os dados não estão em uma mesma escala, não se deve utilizar o coeficiente de correlação como medida de similaridade entre observações, mas considerar a possibilidade de normalização, se o interesse for agrupar variáveis (Nazareth, 2020).

Dentre as medidas de distâncias mais utilizadas para dados métricos, se destaca a distância Euclidiana, que se trata basicamente da distância geométrica entre os pontos. Seja  $\tilde{x}_i = [x_{i1}, \dots, x_{ip}]^t$ ,  $i = 1, 2, \dots, n$ , o vetor de observações do indivíduo  $i$ , no qual  $x_{ij}$  representa o valor da variável  $j$  no indivíduo  $i$ . Com isso, a distância Euclidiana ( $d_{ij}$ ) pode ser calculada conforme a Equação 2.36.

$$d_{ij} = \sqrt{(\tilde{x}_i - \tilde{x}_j)^t (\tilde{x}_i - \tilde{x}_j)} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (2.36)$$

De acordo Ferreira (2018), este tipo de distância não é a mais indicada para grupos de variáveis que não possuem escalas similares, pois variáveis com maior variabilidade irão dominar a classificação das distâncias.

Outras medidas de distâncias podem ser calculadas tomando como base a Distância Euclidiana. No caso, se tem a Distância Euclidiana Generalizada (Equação 2.37), onde  $\mathbf{W}$  é uma matriz positiva definida. Se  $\mathbf{W} = \mathbf{V}^{-2}$ , onde  $\mathbf{V}^{-2}$  é a matriz diagonal dos recíprocos das variâncias, tem-se a Distância Euclidiana Ponderada, que corresponde à Distância Euclidiana para os dados normalizados. Se  $\mathbf{W} = \text{diag}(1/p)$ , tem-se a Distância Euclidiana Média. Se  $\mathbf{W} = \mathbf{\Sigma}^{-1}$  (inversa da matriz de covariâncias), tem-se a Distância de Mahalanobis, que é invariante à mudança de escalas.

$$d_{ij} = \sqrt{(\tilde{x}_i - \tilde{x}_j)^t \mathbf{W} (\tilde{x}_i - \tilde{x}_j)} \quad (2.37)$$

De acordo com Pereira (2015), a Distância Euclidiana Ponderada leva em consideração apenas a diferença de variabilidade que pode haver entre as variáveis. Já a Distância de Mahalanobis leva em consideração na ponderação as possíveis diferenças de variâncias e as relações entre as variáveis, medidas em termos de covariâncias.

Outra medida de distância usual para dados métricos é a Distância de Minkowski, que quando na forma da métrica de Manhattan (Equação 2.38) é menos afetada pela presença de valores discrepantes, ou seja, é menos sensível à presença de *outliers*, mas não é aconselhável quando as  $p$  variáveis possuem diferentes unidades de medida.

$$d_{ij} = \left( \sum_{k=1}^p |x_{ik} - x_{jk}|^\lambda \right)^{1/\lambda} \quad (2.38)$$

Para a Equação 2.38, quando  $\lambda = 1$ , se tem a Métrica de Manhattan, também conhecida como *city block*. Se  $\lambda = 2$ , se tem a Distância Euclidiana.

Conforme foi mencionado anteriormente, é possível através do coeficiente de dissimilaridade ( $d_{ij}$ ), obter o coeficiente de similaridade ( $s_{ij}$ ), através da Equação 2.39.

$$s_{ij} = 1 - d_{ij}^0 \quad (2.39)$$

Onde:

$$d_{ij}^0 = \frac{d_{ij} - \min(D)}{\max(D) - \min(D)},$$

sendo  $\min(D)$  e  $\max(D)$  o menor e o maior valor de distância observados na matriz de distâncias  $\mathbf{D}_{n \times n}$ , sem levar em consideração os elementos da diagonal principal dessa matriz.

- **Métodos Não Hierárquicos**

Na maioria dos métodos não hierárquicos (MANH), é necessário definir a princípio o número de grupos  $k$ , sendo que o objetivo é agrupar os  $n$  indivíduos em  $k$  grupos de modo que cada grupo seja o mais homogêneo possível internamente e o mais heterogêneo possível externamente, otimizando então esse agrupamento. Sendo assim, esses métodos investigam algumas partições na busca de se encontrar a partição ótima. Para isso, alguns métodos foram propostos, sendo o *k-means* um dos mais simples e popular. Segundo Jain (2010) e Vela *et al.* (2012), sua popularidade se deve à sua facilidade de implementação, ao fato de não impor restrições ao conjunto de dados e de sempre convergir para uma solução.

O *k-means* ou k-médias é também conhecido como o método das médias ou centroides móveis, onde cada grupo é representado pelo centroide, que nesse caso é a média dos dados pertencentes a cada grupo. Com isso, deve ser aplicado em variáveis numéricas e é sensível a *outliers*, afinal, o agrupamento é baseado na escolha final, ou “ótima”, do centroide, que é a

média dos indivíduos pertencentes a cada grupo. Sendo assim, dados anômalos irão influenciar no cálculo dessa média.

Existem outros algoritmos de agrupamento por MANH que são menos sensíveis à presença de outliers. Nesses casos, pode-se citar o *k-medoides*, ou *PAM (Partitioning Around Medoids)* (Kaufman & Rousseeuw, 1990) e o *CLARA (Clustering Large Applications)*, que é basicamente uma extensão do *PAM* para bancos de dados maiores. No *k-medoides*, o grupo é representado por um dos objetos pertencentes a ele, não a média do grupo, como no *k-means*. Alboukadel Kassambara explica objetivamente o algoritmo *PAM* no site *Data Novia*, conforme segue abaixo, de acordo com referências como Kassambara (2017) e Kaufman & Rousseeuw (1990).

Considerando um número  $k$  de grupos, o algoritmo PAM é baseado na busca por  $k$  objetos representativos ou medóides entre as observações do conjunto de dados. Depois de encontrar um conjunto de  $k$  medóides, os clusters são construídos atribuindo cada observação ao medóide mais próximo. A seguir, cada medóide  $m$  selecionado e cada indivíduo não medóide são trocados e a função objetivo é calculada. A função objetivo corresponde à soma das dissimilaridades de todos os objetos e seu medóide mais próximo.

A etapa *SWAP* tenta melhorar a qualidade do agrupamento trocando objetos selecionados (medóides) e objetos não selecionados. Se a função objetivo pode ser reduzida trocando um objeto selecionado com um objeto não selecionado, então a troca é realizada. Isso é continuado até que a função objetivo não possa mais ser diminuída. A intenção é encontrar  $k$  objetos representativos que minimizem a soma das dissimilaridades das observações ao seu objeto representativo mais próximo. Em resumo, o algoritmo PAM segue os cinco passos abaixo, sendo os três primeiros referentes à fase de construção e os últimos à fase de troca (SWAP):

- a) Seleciona-se os  $k$  objetos para se tornarem medóides ou, caso esses objetos tenham sido fornecidos, usa-se como medóides;
- b) Calcula-se a matriz de dissimilaridade se ela não foi fornecida;
- c) Atribui-se cada objeto ao seu medóide mais próximo;

- d) Para cada grupo deve-se observar se algum dos objetos do cluster diminui o coeficiente de dissimilaridade médio; em caso afirmativo, seleciona-se a entidade que diminui mais este coeficiente como o medóide para este cluster;
- e) Se pelo menos um medóide mudou, deve-se retornar para a etapa (c); caso contrário, encerra-se o algoritmo.

Devido ao fato de não ser necessário calcular a média dos grupos e considerá-la como o centroide, o *PAM* pode ser mais indicado também quando se tem um banco de dados contendo variáveis ordinais tratadas como quantitativas. Como as variáveis utilizadas nesse trabalho são dessa natureza, o algoritmo *PAM* foi o escolhido para as análises, que foram feitas no *freeware* R (R Core Team, 2020). O algoritmo *PAM* está inserido na função *eclust()* do pacote *factoextra* (Kassambara & Mundt, 2020).

Para calcular a matriz de dissimilaridades, a função *eclust()* pode utilizar as seguintes métricas: Distância Euclidiana, Distância de Minkowski (incluindo a métrica de Manhattan), Camberra, além das medidas de correlação. De acordo com Kassambara (2017), na prática é usual a obtenção de resultados semelhantes, seja utilizando a Distância Euclidiana ou Manhattan. No entanto, se no banco de dados contiverem valores discrepantes (*outliers*), a Distância de Manhattan deve fornecer resultados mais robustos, enquanto a Euclidiana seria mais sensível a esses valores incomuns. Nesse trabalho, foi utilizada a Distância Euclidiana.

### **2.5.3. Análise de Componentes Principais**

De acordo com Mingoti (2013) e conforme citado por Pereira (2015), a análise de componentes principais (ACP ou PCA – *Principal Component Analysis*) é uma técnica de estatística multivariada em que um conjunto de variáveis correlacionadas entre si é transformado em outro conjunto de variáveis não correlacionadas (ortogonais) através de combinações lineares, chamadas componentes principais. A finalidade dessas combinações lineares é explicar a estrutura de variância e covariância de um vetor aleatório composto por  $p$  variáveis. Essa técnica foi fundamentada no artigo de Hotelling (1933).

O número de componentes principais geradas é igual ao número de variáveis ( $p$  variáveis geram  $p$  componentes). Essas componentes são calculadas em ordem decrescente de importância, ou seja, a primeira componente é mais importante que a segunda e assim



sucessivamente. Essa importância se refere à quantidade de informação contida em cada componente, isto é, o quanto da variância dos dados originais é explicada por cada componente. Isso significa que a primeira componente explica A% da variância dos dados originais; a segunda componente, B%, sendo  $B < A$  e a componente p explica C%, sendo essa a menor proporção. O somatório dessa porcentagem da variância explicada por cada componente é 100%.

Sabendo disso, pode-se entender com mais clareza o principal objetivo da aplicação dessa técnica: a redução da dimensionalidade dos dados. Se um banco de dados possui p variáveis, essa técnica pode ser aplicada para reduzi-lo. Ou seja, extrair as k componentes principais, onde  $k < p$ , de modo que k tenha uma % da variância explicada que seja satisfatória para a análise em questão. Além desse objetivo, a ACP pode ser aplicada também quando se deseja quantificar variáveis qualitativas, para aplicação posterior de alguma outra técnica de estatística multivariada.

Como o objetivo do uso da ACP nesse trabalho é para quantificar as variáveis e não para reduzir a dimensionalidade dos dados, os critérios para a escolha de quantas componentes devem ser retidas na análise não serão detalhados, visto que fogem do escopo nesse caso. No entanto, como referências, sugere-se os critérios estabelecidos por Kaiser (1970) e Catell (1966), além do livro de Mingoti (2013).

A ACP não exige normalidade multivariada para ser aplicada. Portanto, é aplicável a qualquer conjunto de variáveis, desde que a adequabilidade do banco de dados para a aplicação de técnicas de estatística multivariada esteja comprovada. Além do mais, as variáveis devem ser quantitativas ou qualitativas ordinais. A técnica pode ser aplicada tanto em matriz de covariâncias ou de correlações. Quando as variáveis apresentam variabilidades muito diferentes, ou magnitude de valores discrepantes, ou com diferentes unidades de medida, deve-se utilizar a matriz de correlações, para que não haja viés na análise.

No caso desse trabalho, mesmo que as variáveis sejam ordinais e não tenham diferenças de magnitude ou de ordem de grandeza, optou-se por extrair as componentes principais por matriz de correlação a fim de assegurar que de fato nenhuma variável com maior variabilidade exerça maior influência nas análises.

- **Componentes principais extraídas por matriz de correlações**

Santos (2019) utilizou em sua tese as componentes principais extraídas por matriz de correlações para posterior aplicação de análise discriminante nessas componentes. Com isso, descreveu de forma objetiva e prática o procedimento de extração por matriz de correlação, conforme se segue.

Seja  $\mathbf{X}$  um vetor aleatório com  $p$  variáveis. Uma variável qualquer  $X_i$  pertencente ao vetor  $\mathbf{X}$  pode ser padronizada conforme a Equação 2.40, com  $i = 1, 2, \dots, p$ , em que  $\mu_i$  é a média e  $\sigma_i$  é o desvio padrão da variável  $i$ .

$$Z_i = \frac{(X_i - \mu_i)}{\sigma_i} \quad (2.40)$$

A matriz  $\mathbf{P}_{p \times p}$  é a matriz de covariâncias e variâncias das variáveis padronizadas  $Z_i$ , mas também é a matriz de correlação das variáveis  $X_i$ .  $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p$  são os autovalores ordenados e  $e_1, e_2, e_3, \dots, e_p$  os respectivos autovetores normalizados da matriz de correlações.

A  $i$ -ésima componente principal, tal que  $i=1, 2, 3, \dots, p$ , é dada pela Equação (2.41), onde  $Y_i$  é a  $i$ -ésima componente principal,  $e_i$  é o  $i$ -ésimo autovetor normalizado e  $\mathbf{Z}$  é o vetor aleatório padronizado.

$$Y_i = \mathbf{e}_i^t \mathbf{Z} = [e_1 \quad e_2 \quad e_3 \quad \dots \quad e_p] \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix} = e_{i1}Z_1 + e_{i2}Z_2 + \dots + e_{ip}Z_p \quad (2.41)$$

A variância de  $Y_i$  é dada pela Equação (2.42).

$$Var[Y_i] = Var[\mathbf{e}_i^t \mathbf{Z}] = \mathbf{e}_i^t Var[\mathbf{Z}] \mathbf{e}_i = \mathbf{e}_i^t \sum \mathbf{e}_i = \mathbf{e}_i^t \lambda_i \mathbf{e}_i = \mathbf{e}_i^t \mathbf{e}_i \lambda_i = \lambda_i \quad (2.42)$$

Onde:  $\sum \mathbf{e}_i = \lambda_i \mathbf{e}_i$ .

Portanto o autovalor  $\lambda_i$  é a variância associada à componente principal  $Y_i$ . A proporção da variância total de  $\mathbf{X}$  que é explicada pela  $i$ -ésima componente principal é dada pela Equação (2.43).

$$\frac{Var[Y_i]}{Variância\ total\ de\ X} = \frac{\lambda_i}{traço(\Sigma\ pxp)} = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \quad (2.43)$$

#### 2.5.4. Análise Discriminante

Enquanto que a análise de agrupamento tenta enquadrar indivíduos em grupos homogêneos, com base nas suas similaridades e dissimilaridades, sem requerer um conhecimento prévio desses grupos, a análise discriminante é uma técnica estatística que visa gerar uma regra, um modelo de classificação de novos indivíduos, através de funções matemáticas, com base em um conjunto de variáveis que são responsáveis por essa classificação desses indivíduos em um grupo.

Sendo assim, para gerar essa regra de classificação é necessário que se tenha um banco de dados multivariado, onde se conhece *a priori* em qual grupo cada indivíduo está inserido. Por exemplo, se deseja gerar uma regra de classificação que irá dizer, com base em um conjunto de variáveis, se um talude é estável ou instável. Para gerar esse modelo, é necessário que se tenha um banco de dados com as variáveis importantes para o estudo de caso e também que se conheça a condição de estabilidade ou instabilidade de cada talude. Com isso, o modelo irá identificar a contribuição de cada variável e gerar uma função matemática que poderá então ser utilizada em outros taludes, para prever sua condição de estabilidade.

A Figura 2.20 apresenta uma representação univariada de escores discriminantes de duas populações A e B com distribuição normal. Na Figura 2.20-a é possível observar que os centroides estão mais distantes e, portanto, a função discriminante foi mais eficaz e gerou uma área de confusão menor (destacada em cinza) do que na Figura 2.20-b onde os centroides estão mais próximos (Santos, 2019).

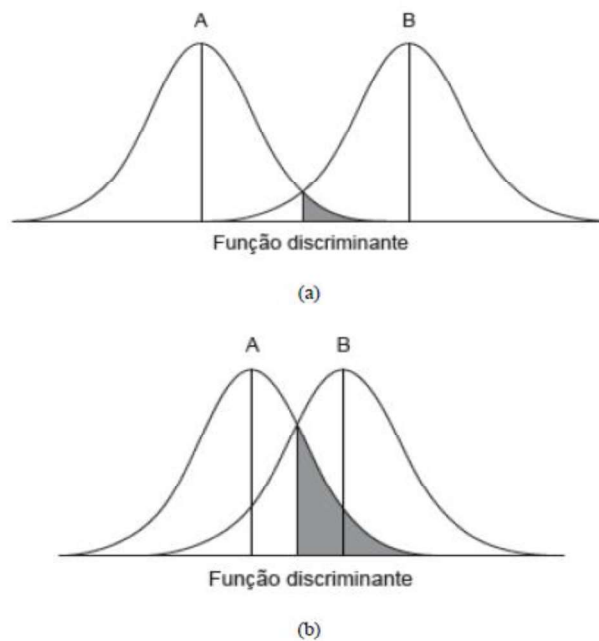


Figura 2.20 - Representação univariada de escores Z discriminantes (HAIR et al.,2009).

Para aplicação da análise discriminante, é necessário que as  $p$  variáveis independentes sejam quantitativas, enquanto que a variável dependente é qualitativa (ou não métrica), visto que normalmente essa variável representa um rótulo (ex: bom e ruim; estável ou instável; alto, médio ou baixo). Outros pressupostos são que as observações devem ser independentes entre si, ou seja, uma variável independente não deve ser dependente de outra; e as observações devem pertencer a um único grupo.

A construção da regra de classificação pode ser feita através da função discriminante linear de Fisher (Fisher, 1936) para duas populações ou através das funções discriminantes canônicas de Fisher, para mais de duas populações. Deve-se destacar que as funções lineares e canônicas de Fisher (Fisher, 1936) ainda possuem outro pressuposto para aplicação, visto que são indicadas para dados homocedásticos. Se estes não forem homocedásticos, a função discriminante quadrática é mais indicada.

No entanto, conforme citado por Santos (2019), a função quadrática, devido à sua construção ser baseada em populações com distribuições normais, é muito sensível à normalidade multivariada dos dados. Com isso, a normalidade se torna um pré-requisito importante para a sua aplicação. Portanto, mesmo que os dados não sejam homogêneos, ainda pode ser mais eficaz aplicar a função discriminante linear quando não houver a normalidade multivariada,

do que a quadrática. Assim como na tese de Santos (2019), apenas a função canônica de Fisher é apresentada nesse trabalho, uma vez que ela foi utilizada para a construção do sistema de análise de probabilidade e consequências.

Sejam  $X_j$ , com  $j = 1, 2, \dots, p$ , variáveis aleatórias e  $k$  populações. Se as matrizes de covariâncias dessas populações são iguais, tem-se que essas populações são homocedásticas. A partir daí, pode-se construir as combinações lineares denominadas funções discriminantes canônicas (Equação 2.44) (Mingoti, 2013).

$$Y_i = \hat{e}_j^t \mathbf{X}_{px1} \quad (2.44)$$

Onde  $\hat{e}_j^t$  é o  $j$ -ésimo autovetor estimado correspondente ao  $j$ -ésimo maior autovalor da matriz  $\mathbf{W}^{-1}\mathbf{B}$ , sendo  $\hat{e}_j^t \mathbf{W} \hat{e}_j = 1$ . As matrizes  $\mathbf{W}$  e  $\mathbf{B}$  são definidas, respectivamente pelas Equações 2.45 e 2.46, em que a matriz  $\mathbf{W}$  é a matriz de soma de quadrados e produtos cruzados dentro dos grupos e  $\mathbf{B}$  é a matriz de soma de quadrados e produtos cruzados entre os grupos.

$$\mathbf{W}_{pxp} = \sum_{i=1}^k \sum_{b=1}^{n_i} (x_{ib} - \bar{x}_i) (x_{ib} - \bar{x}_i)^t \quad (2.45)$$

$$\mathbf{B}_{pxp} = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x}) (\bar{x}_i - \bar{x})^t \quad (2.46)$$

Onde;

$k$ : número de populações (classes);

$n_i$ : número de elementos da população  $i$ ;

$x_{ib}$ : vetor de observações do elemento amostral  $b$  que pertence à população  $i$ ;

$\bar{x}_i$ : vetor de médias amostral da população  $i$ ;

$\bar{x}$ : vetor de médias amostral, considerando todas as  $n$  observações.

Deve-se destacar que assim como na ACP, as funções discriminantes também possuem uma ordenação em termos de importância. Ou seja, a primeira função discriminante tem maior poder de discriminação que a segunda e assim sucessivamente, se for o caso.

Após a construção dessas funções, para cada elemento amostral são gerados os escores nessas funções, denotados por  $\hat{Y}_j$ . Além desses, também se tem os escores das funções discriminantes aplicadas aos vetores de médias amostrais observados para cada população, os

$\hat{Y}_i$ . A partir daí, se calcula a distância Euclidiana entre os vetores  $\hat{Y}_j$  e  $\hat{Y}_i$  (Equação 2.47), onde os indivíduos são classificados na população cuja distância Euclidiana entre o indivíduo e a média da população for menor.

$$d = \sqrt{(\hat{Y}_j - \hat{Y}_i)^t (\hat{Y}_j - \hat{Y}_i)} \quad (2.47)$$

### 2.5.5. Elipses de Confiança

Após realizar a análise discriminante, se tem as funções discriminantes e também os gráficos dos escores gerados pela técnica. Por esse gráfico, é possível perceber a divisão entre os indivíduos nos grupos analisados, sem, no entanto, observar um limite claro entre um grupo e outro. Para se determinar esse limite, um artifício é traçar elipses de confiança em cada grupo, onde, através dos pontos de interseção entre essas elipses, se traça uma reta que é o próprio limite entre um grupo e outro (Figura 2.21).

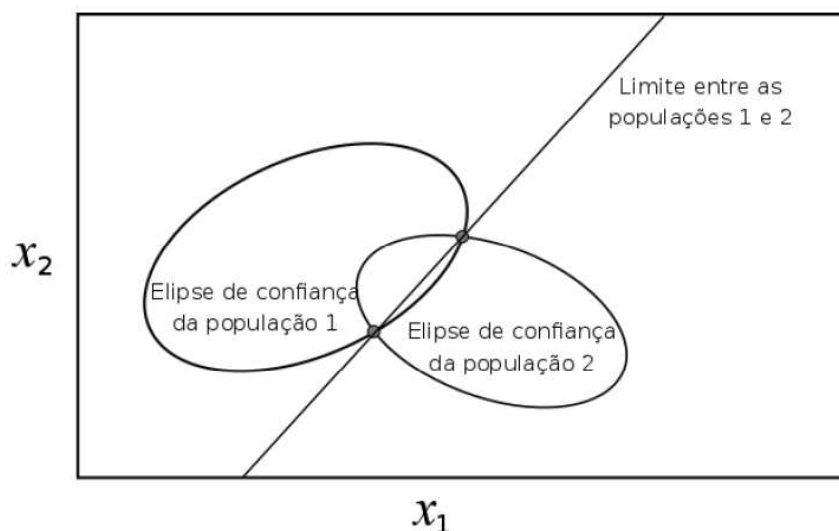


Figura 2.21 - Elipses de confiança e limites entre os grupos (Santos, 2019).

Elipses de confiança são construídas utilizando a distância estatística de Mahalanobis, onde a delimitação dos pontos de um mesmo grupo se dá quando essas distâncias são menores ou iguais do que os semi-eixos da elipse (Hair *et al.*, 2009).

As elipses de confiança são construídas de modo que se tenha dentro de cada elipse uma probabilidade  $(1-\alpha)\%$ . Ou seja, para um nível de significância ( $\alpha$ ) de 0,05, a probabilidade de

que os dados de determinada população estejam dentro da elipse que a delimita é igual a 95%. Quando os dados são normais multivariados, a distribuição pode ser aproximada por uma distribuição qui-quadrada e a elipse é dada pela Equação 2.48, conforme já havia pontuado Santos (2019).

$$(\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \leq X_2^2(\alpha) \quad (2.48)$$

Onde:

$\mathbf{x}$  é o ponto de observação;

$\boldsymbol{\mu}$  é a média da população  $i$ ;

$\boldsymbol{\Sigma}$  é a matriz de variâncias e covariâncias populacional;

$X_2^2$  é a distribuição qui-quadrada com dois graus de liberdade;

$\alpha$  é o nível de significância.

Para os casos em que os dados não apresentam distribuição normal multivariada, não se pode aproximar as distâncias de Mahalanobis pelo valor da distribuição quiquadrado para um determinado nível de significância. Inclusive esse é um erro que alguns pesquisadores cometem, ou seja, utilizam elipses de confiança para delimitar as classes em um determinado projeto utilizando a aproximação mostrada na Equação 2.48, sem verificar se as variáveis seguem uma distribuição normal. Suorineni *et al.* (2001) e Papaivanou & Suorineni (2015) são exemplos de pesquisadores que não verificaram a normalidade das variáveis antes de aplicar o conceito de elipses de confiança conforme explicado acima.

Se os dados não são normais, a solução para se aplicar as elipses de confiança de forma correta é através de métodos de reamostragem, de modo que a distribuição amostral empírica é obtida após a etapa das reamostragens. O método Bootstrap não-paramétrico (Efron, 1979), é uma técnica de reamostragem em que a distribuição de um conjunto de dados pode ser aproximada pela distribuição empírica dos próprios dados após sucessivas amostragens com reposição dentro do próprio conjunto de dados.

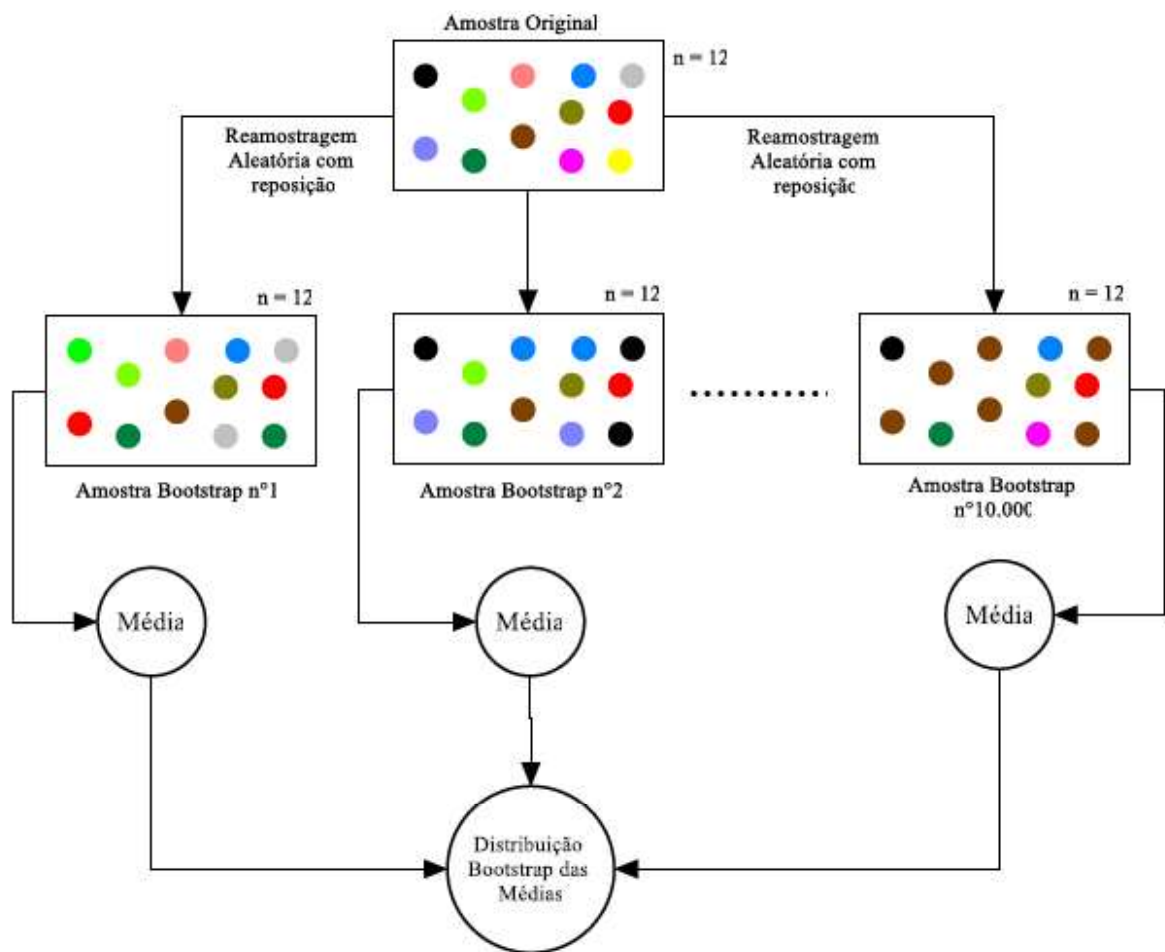
Para isso, o método de Efron (1979) considera a amostra sendo a própria população e a partir dela são realizadas sucessivas reamostragens aleatórias coletando várias amostras do mesmo tamanho da população, que são chamadas amostras bootstrap. O processo é realizado ao se

coletar aleatoriamente cada indivíduo da amostra bootstrap, com reposição desse indivíduo antes de se retirar o próximo, de modo que o objetivo final de cada processo é conseguir uma amostra bootstrap com mesmo tamanho da população, mas não necessariamente com os mesmos valores.

Esse ciclo deve ser repetido um determinado número de vezes, tal que seja possível obter dados suficientes para calcular a distribuição amostral empírica de alguma variável estatística da amostra tomada inicialmente (população). Sugere-se que o processo se repita em torno de 10000 vezes. Essas distribuições são chamadas de distribuição bootstrap (Efron,1992). Conforme citado por Vassalo (2017), a ideia do bootstrap é simular diversos cenários a partir de uma única amostra e desta forma tentar construir um espaço amostral e realizar inferências sobre parâmetros de interesse.

A Figura 2.22 ilustra o processo considerando uma amostra com 12 elementos. Nesse exemplo, cada amostra bootstrap é gerada a partir da amostra, ou população inicial através da retirada aleatória de um elemento e sua reposição antes da retirada do próximo, conforme explicado acima. Assim segue até se completar 12 elementos. Em seguida a média da amostra bootstrap é calculada. O processo se repete até atingir 10000 amostras bootstrap, número suficiente para possibilitar o cálculo da distribuição bootstrap das médias amostrais bootstrap.





**Figura 2.22** - Exemplo do método bootstrap para uma amostra de 12 elementos. (Vassalo, 2017).

A distribuição gerada pelo método Bootstrap é mostrada na Equação 2.49.

$$(\mathbf{x} - \bar{\mathbf{x}})^t \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}) \leq \text{percentil da distribuição bootstrap dos percentis } (1 - \alpha)\% \quad (2.49)$$

Onde:

$\mathbf{x}$  é o ponto de observação;

$\bar{\mathbf{x}}$  é a média amostral da população i;

$\mathbf{S}$  é a matriz de variâncias e covariâncias amostral;

$\alpha$  é o nível de significância.

# CAPÍTULO 3

## METODOLOGIA

A metodologia utilizada para a construção do sistema de análise de risco proposto nesse trabalho está resumida no fluxograma mostrado na Figura 3.1 e é detalhada nos tópicos seguintes.

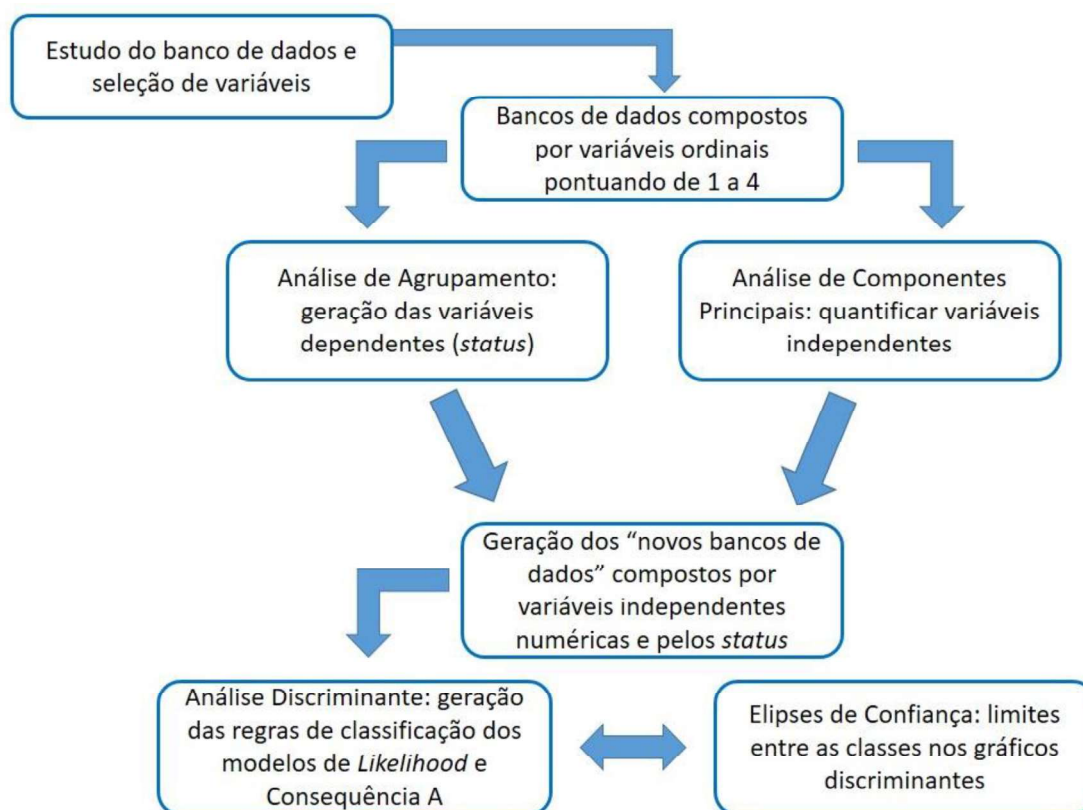


Figura 3.1 – Fluxograma resumido da metodologia utilizada para construir o sistema de análise de risco.

### 3.1. ESTUDO PRELIMINAR DO BANCO DE DADOS

O banco de dados utilizado neste trabalho é parte do mesmo banco utilizado por Santi *et al.* (2009) para gerar e aplicar o CRHRS modificado nos taludes rodoviários do Colorado (EUA) e foi gentilmente cedido por Paul M. Santi. Originalmente, esse arquivo é composto por 355 taludes rodoviários, de corte e naturais, sendo 224 em rochas cristalinas, 74 em rochas sedimentares e 57 taludes formados em solo ou rocha muito alterada, erodível, mas com a presença de blocos de rocha mais rígida (*block-in-matrix*).

Antes de se verificar a viabilidade da aplicação de qualquer técnica estatística, é necessário definir uma estratégia de trabalho que permita, primeiramente, se ter ideia do que é possível de ser feito dentro de um determinado conjunto de dados disponível e da resposta que se pretende obter no final, que nesse caso é um sistema de análise de risco para quedas de blocos em taludes rodoviários.

Sabendo que risco é dado matematicamente pela multiplicação da probabilidade de um evento ocorrer pelas consequências desse evento, o primeiro passo desse trabalho foi observar os parâmetros presentes no banco de dados completo e ver quais teriam mais relação com a probabilidade, ou *likelihood*, de ruptura, se existem variáveis relacionadas à exposição das vias e dos motoristas e se existe algum fator que pode ser definido como *status* para probabilidade, perigo, exposição, consequência, ou mesmo risco.

Conforme foi explicado anteriormente, no banco de dados consta, para cada talude, a pontuação atribuída aos parâmetros avaliados no CRHRS Modificado, variando de 3 a 81, da situação mais favorável para a mais desfavorável à segurança. Isto é, foram pontuadas as variáveis relacionadas às condições geológicas e geotécnicas, às condições climáticas, à geometria do talude e da área de captação e ao tráfego. Todas as variáveis consideradas no CRHRS modificado foram detalhadas no subitem 2.4.1 da Revisão Bibliográfica reservado apenas para essa finalidade.

Também foi explicado no subitem 2.4.1 que para os três tipos de taludes, os parâmetros ligados às condições climáticas, geometria e tráfego são os mesmos. Já quanto às características geológicas e geotécnicas, cada tipo possui três variáveis próprias. Tanto taludes em rochas cristalinas como em rochas sedimentares possuem 6 variáveis relacionadas às discontinuidades, já os taludes *block-in-matrix* não. Isso significa que os taludes em rochas cristalinas e sedimentares possuem 6 variáveis a mais que os taludes de solo com blocos de rocha.

Só o fato de os taludes rochosos apresentarem 6 variáveis a mais que os taludes *block-in-matrix* já indica que esse último não deve ser analisado juntamente com os primeiros, pois além de ter menos variáveis, nenhuma delas é observada nos taludes cristalinos e

sedimentares. Conforme foi mostrado subitem 2.4.1, os únicos parâmetros relacionados às características do maciço para os taludes *block-in-matrix* são vegetação, tamanho de blocos e forma de blocos.

O tamanho e a forma dos blocos não são exatamente variáveis que descrevem se pode haver maior ou menor suscetibilidade à queda, mas sim sobre a energia cinética desses blocos durante a queda e o alcance; afinal, o evento pode ocorrer para blocos tabulares, arredondados, grandes ou pequenos; a magnitude e a velocidade é que serão diferentes de acordo com essas características. Além do mais, as causas e gatilhos para queda de blocos nesses taludes são muito diferentes das causas em taludes rochosos. Como o foco desse trabalho são os taludes rochosos, os taludes de solo com blocos de rocha não serão analisados.

Quanto aos taludes rochosos, apesar de os cristalinos e sedimentares terem em comum as variáveis relacionadas às descontinuidades, eles também possuem três variáveis próprias e isso também faz com que eles não possam ser avaliados juntos em análises que envolvem os parâmetros relacionados ao maciço rochoso. Embora algumas variáveis próprias até apresentem certa semelhança e levem ao mesmo efeito, esses são causados por situações diferentes e não devem ser consideradas da mesma forma.

Por exemplo: o grau de descalçamento para rochas cristalinas é semelhante e causa o mesmo tipo de problema que o grau de descalçamento para rochas sedimentares. O mecanismo de queda de blocos é similar. No entanto, para rochas cristalinas isso ocorre principalmente devido a quedas anteriores e problemas estruturais no maciço, enquanto que para rochas sedimentares isso é causado pela erosão diferencial. Com isso, não se pode admitir em um modelo ou equação de probabilidade de queda, que o grau de descalçamento para rochas cristalinas terá o mesmo peso que para rochas sedimentares, que considera também outros fatores relacionados com a erosão diferencial. Sendo assim, o ideal é que sejam criados modelos de probabilidade diferentes para cada tipo de talude.

Neste trabalho o modelo de probabilidade, ou *likelihood*, de ocorrência de queda de blocos foi gerado considerando apenas os taludes em rochas cristalinas, porque constituem a maioria dos taludes no banco de dados. Com base no modelo para esses taludes e nas discussões e

conclusões obtidas, o modelo de probabilidade para taludes sedimentares poderá ser gerado em outro trabalho.

É importante salientar que os modelos de consequências valem para ambos, visto que não consideram parâmetros geológicos e geomecânicos, portanto não sofrerão mudanças para os taludes sedimentares. Sendo assim, e tendo em vista as similaridades do mecanismo de queda de blocos em taludes cristalinos e sedimentares, o modelo de consequência foi gerado considerando todos os taludes rochosos do banco de dados, ou seja, tanto os taludes cristalinos quanto sedimentares. Os modelos e as variáveis incluídas nas análises de probabilidade e consequências serão detalhados em seguida.

É importante destacar a adequação do uso do termo *likelihood* em conjunto com probabilidade nesse trabalho. Isso se deve ao fato de que no final, são gerados índices numéricos em escala não percentual relacionados à probabilidade do bloco cair, como poderá ser verificado nos resultados. Esse índice será então plotado em um gráfico de classificação da probabilidade, ou seja, probabilidade alta, média, ou baixa, sendo que o termo *likelihood* se refere justamente à probabilidade em termos qualitativos.

### **3.2. ANÁLISE E SELEÇÃO DE VARIÁVEIS**

Após a análise completa e preliminar do banco de dados disponível, deu-se início ao estudo mais detalhado das variáveis, a fim de verificar quais deveriam entrar na análise e qual a possível relação dessas variáveis com o risco; ou seja, se uma determinada variável está mais relacionada com a probabilidade de haver queda de blocos, ou com as consequências dessa queda.

Como o objetivo desse trabalho é gerar um sistema de análise de risco de queda de blocos para taludes rodoviários que seja aplicável em taludes rochosos de qualquer lugar, foi necessário primeiramente verificar quais parâmetros do banco de dados são típicos do Colorado ou de regiões semelhantes ao Colorado. Dentre esses parâmetros, pode-se citar os ciclos de gelo e degelo e o aspecto do talude (ver subitem 2.4.1), pois são fatores climáticos relacionados a regiões que nevam, o que não é o caso do Brasil e de outros países tropicais, por exemplo. Portanto, esses dois parâmetros foram excluídos da análise.

O próximo passo foi analisar quais parâmetros estão mais relacionados à probabilidade de ocorrência de queda de blocos para taludes em rochas cristalinas. Nessa fase, foram observados quais os parâmetros podem ter mais influência na instabilização do talude, na ocorrência da queda de blocos em si, de acordo com o que é comentado pela literatura geotécnica clássica e adotado em outras classificações, sendo algumas citadas na Revisão Bibliográfica, além da experiência prática de campo.

Também foram verificados quais parâmetros estão ligados com as consequências dessas quedas para taludes rochosos. Em se tratando de taludes rodoviários, as consequências de uma queda de blocos podem ser desde financeiras ou estruturais, quando os blocos provocam danos nas estruturas rodoviárias e/ou carecem de ser removidos, como podem ser mais graves e representar danos à integridade física e à vida dos motoristas. Sendo assim, foram verificados tanto os parâmetros que estão relacionados com danos físicos aos motoristas, quanto os relacionados ao alcance dos blocos, pois se o bloco atinge a rodovia ou alguma estrutura importante, automaticamente já haverá gastos com a remoção desses blocos e até mesmo obras de recuperação, se for o caso.

Além disso, também foi analisado se no banco de dados existia uma ou mais variáveis que poderiam ser definidas como o *status* de alguma classificação. Por exemplo: existe alguma variável ali que pode indicar se a probabilidade de acontecer queda de blocos é alta, média ou baixa? Existe no banco de dados alguma variável que poderia ser a resposta final da análise de risco, ou seja, risco alto, médio ou baixo?

É importante destacar que isso foi feito porque a metodologia CRHRS não classifica os taludes quanto ao risco ou perigo e não havia qualquer menção quanto a isso no banco de dados. O CRHRS gera um índice dado pelo somatório das notas atribuídas às variáveis, hierarquizando os taludes mais problemáticos.

Ao observar as variáveis consideradas por Santi *et al.* (2009) (subitem 2.4.1), foi pensado em um primeiro momento que a frequência de queda poderia ser considerada como o *status* para a probabilidade, bem como o número de acidentes como *status* para o risco e também para as consequências. No entanto, os próprios autores que levantaram e utilizaram esse banco de

dados alertaram sobre as incertezas a cerca dessas variáveis (subitem 2.4.1), o que tornaria o seu uso tecnicamente inconsistente, mesmo se fossem feitas análises estatísticas e essas fossem bem sucedidas.

Com isso, essas duas variáveis também foram excluídas das análises. Outro motivo para exclusão é que nem todos os países, órgãos de trânsito ou concessionária de rodovias fazem um monitoramento regular sobre a frequência de ocorrência de quedas de blocos e acidentes causados em decorrência desse evento, o que tornaria um modelo de análise de risco contabilizando essas variáveis inviável de ser aplicado nessas circunstâncias.

Todas as variáveis do banco de dados original são qualitativas ordinais, cujas “notas” variam de 3 a 81, conforme explicado anteriormente. Para esse trabalho, as variáveis consideradas foram pontuadas de 1 a 4, seguindo a mesma ordem do CRHRS, conforme é mostrado na Tabela 3.1. Ou seja, nesse trabalho, foi considerado o posto das variáveis, o que é mais adequado para análises estatísticas em dados ordinais.

Tabela 3.1– Pesos atribuídos às variáveis no CRHRS e no trabalho proposto.

Nota adotada pelo CRHRS	Nota adotada nesse trabalho
3	1
9	2
27	3
81	4

A partir daí, foi então possível construir os bancos de dados para aplicação das técnicas de estatística multivariada e criação do modelo de probabilidade (*Likelihood*) do evento acontecer e do modelo de consequência do evento, para enfim, gerar a matriz de risco.

### **3.3. VIABILIDADE DE APLICAÇÃO DAS TÉCNICAS DE ESTATÍSTICA MULTIVARIADA**

Para verificar se os bancos de dados são adequados para aplicação das técnicas de estatística multivariada, ou seja, se as variáveis apresentam correlações significativas, foi feito o teste de esfericidade de Bartlett. Nessa fase também foi realizada a análise descritiva das variáveis através de gráficos de frequência, antes da aplicação das demais técnicas.

### 3.4. ANÁLISE DE AGRUPAMENTO

Conforme mencionado anteriormente, as variáveis utilizadas para a criação do modelo são ordinais e foram discretizadas em valores de 1 a 4. É comum no meio geotécnico esse tipo de discretização, visto que observações de campo são em grande parte qualitativas e aquelas que podem ser medidas, muitas vezes são feitas com certo grau de incerteza devido à variabilidade inerente às condições da natureza, o que torna preferível a adoção de pesos atribuídos a faixas de valores. É tão comum que em muitas classificações geomecânicas os índices obtidos através de somatórios desses pesos são utilizados para estimar parâmetros importantes de projetos de engenharia.

Sendo assim, foi possível assumir que o incremento entre as categorias dessas variáveis é igual a uma unidade, podendo trata-las como variáveis numéricas e aplicar técnicas de agrupamento mais apropriadas para esse tipo de variável. Segundo Nazareth (2020), as vantagens são que isso dá ao pesquisador maior flexibilidade na escolha de métodos de análises e preserva a informação de ordem. Além do mais, permite que as análises sejam feitas por métodos que são mais familiares ao público, e mais fáceis de entender.

A Análise de Agrupamento foi feita com o objetivo principal de agrupar os taludes em classes de probabilidade (*Likelihood*) de ocorrência de queda de blocos e também em classes de consequências, onde essas consequências foram abordadas sob dois aspectos: chance do bloco atingir a rodovia (Consequência A) e chance do bloco causar um acidente (Consequência B).

Isso foi feito porque no banco de dados não havia nenhuma variável que pudesse ser considerada como a variável dependente ou *status* para a criação do modelo de análise de risco, ou seja, não havia no banco de dados alguma variável que indicasse a condição de estabilidade do talude, nem a chance de haver acidentes devido a essas quedas de blocos.

Até existe no banco a variável “Número de Acidentes”, no entanto, os próprios responsáveis pelo levantamento mencionam as incertezas desse parâmetro. Outra questão relevante é que nem sempre há esse tipo de dado disponível em muitos locais, o que tornaria a aplicação do modelo inviável nessas situações.



Com isso, a Análise de Agrupamento foi aplicada primeiramente no banco de dados contendo as variáveis relacionadas à probabilidade e posteriormente no banco de dados relacionados à Consequência A. A intenção era agrupar os taludes de modo que se pudesse, nos dois casos, dividi-los em três classes: alta, média (zona de incerteza ou de transição) e baixa. Para a Consequência B foi adotada outra estratégia, o que é comentado com mais detalhes nos resultados.

Para isso, antes foi verificada para as duas análises se havia tendência de agrupamento; isso foi feito tanto pelo valor da Estatística Hopkins quanto visualmente pelo gráfico de calor. O número ótimo de *clusters* para os três casos foi obtido através da silhueta média, do método do cotovelo e também pela “regra da maioria”. Foi observado se esse número ótimo de grupos poderia gerar uma classificação coerente com os objetivos das análises.

Verificada a existência de tendência de agrupamento e tomadas as decisões quando ao número de grupos que deveria ser adotado para as análises de probabilidade (*Likelihood*) e Consequência A, partiu-se para a análise de agrupamento em si. Devido à natureza das variáveis (variáveis ordinais tratadas como numéricas), o método escolhido para os três casos foi o k-medóides (*PAM*), que se trata de um método não hierárquico em que os indivíduos mais representativos de um grupo são os próprios medóides desses grupos, não a média desses indivíduos (centróides), como no *k-means*.

Os grupos formados pela análise de agrupamento no *software* R (R Core Team, 2020) foram exportados para o Excel para verificar como se deu a formação dos grupos e se esses grupos eram interpretáveis e úteis tecnicamente, visto que estatisticamente sempre serão formados *clusters*. Nessa fase, foi feita uma análise das faixas dos somatórios das notas de cada grupo, com o objetivo de verificar se esses grupos formados poderiam ser interpretados como classes alta, média e baixa.

Um exemplo para ilustrar o que foi comentado acima quanto aos somatórios: supondo que para o modelo de probabilidade foram consideradas 8 variáveis que poderiam pontuar de 1 a 4. Logo, o somatório mínimo é 8 e o máximo é 32. Daí foi analisado nesses grupos formados pela análise de agrupamento, como estavam as faixas de somatório. Isto é, se é possível dizer

que no grupo 1 estão basicamente os taludes que pontuam entre 8 e X, no grupo 2 os que pontuam entre X+1 e Y e no grupo 3 os que pontuam entre Y+1 até 32.

Tanto para a probabilidade quanto para as consequências (quando a análise de agrupamento foi aplicada), essa divisão de acordo com faixas de somatório de notas pôde ser verificada, o que facilitou e tirou grande parte da subjetividade do processo de determinação das classes, pois após a verificação dessas faixas determinadas estatisticamente, foi feito também um estudo a fim de se verificar se era necessário algum ajuste nessas faixas. Com isso, é possível afirmar que a determinação de classes (alta, média e baixa) não foi feita apenas baseada em critérios subjetivos, mas sim em conjunto com um robusto embasamento estatístico.

Os grupos formados após a aplicação da análise de agrupamento e das modificações feitas nesses grupos, quando necessário, foram então considerados como as variáveis dependentes (*status*) para aplicação da Análise Discriminante.

### **3.5. ANÁLISE DE COMPONENTES PRINCIPAIS**

Para a criação dos modelos (equações e gráficos) de probabilidade e consequências, foi utilizada a Análise Discriminante, cujo pré-requisito é que as variáveis sejam quantificadas previamente. Sendo assim, a Análise de Componentes Principais (ACP) foi feita para quantificar as variáveis, de modo que os escores da ACP, e os grupos gerados com base nas Análises de Agrupamento formariam os novos bancos de dados para aplicação da Análise Discriminante. A extração dos Componentes Principais na ACP foi feita através da matriz de correlações das variáveis, para evitar qualquer viés relacionado a variáveis com maior variabilidade, por exemplo.

### **3.6. ANÁLISE DISCRIMINANTE E CRIAÇÃO DO MODELO**

As Análises Discriminantes foram aplicadas nos “novos” bancos de dados, onde as variáveis independentes eram os escores da ACP e as dependentes, os grupos gerados com base nas Análises de Agrupamento. Daí, foram criados os modelos, com equações e gráficos, onde é possível classificar novos taludes quanto à probabilidade de haver queda de blocos e quanto as consequências dessas quedas.

O teste de M Box (Box, 1949) foi realizado antes da Análise Discriminante a fim de verificar se os dados são homocedásticos. Esse resultado, juntamente com o resultado do teste de normalidade de Royston (1983) auxiliou na tomada de decisão entre Análise Discriminante Linear ou Quadrática. É importante salientar que a Análise Discriminante Quadrática pressupõe a normalidade dos dados e é bastante sensível a ela.

A validação dos modelos de classificação gerados pela Análise Discriminante foi realizada através do método da ressubstituição com divisão amostral, onde uma parte dos dados é utilizada apenas para a criação do modelo (amostra de treino) e a outra apenas para aplicação do modelo e estimação da probabilidade de erro (amostra de teste). No caso, 75% do banco de dados foi separado aleatoriamente como amostra de treino e 25% como amostra de teste, conforme é sugerido por Hair *et al.* (2009).

É importante ressaltar que as amostras de treino e de teste são separadas aleatoriamente e que diferentes amostras podem gerar diferentes resultados e, conseqüentemente, diferentes taxas de erro. Ou seja, a depender de quais indivíduos foram alocados na amostra de treino e de quais foram alocados na amostra de teste, o resultado da Análise Discriminante pode variar. No entanto, para gerar um modelo pronto para uso, com equações e gráficos de classificação, é necessário selecionar uma amostra específica de treino e uma de teste.

Sendo assim, antes de selecionar uma determinada amostra aleatoriamente para criar o modelo, foi feita uma avaliação considerando 31 diferentes amostras de treino e de teste a fim de se verificar a performance da técnica de Análise Discriminante, ou seja, a fim de se determinar o erro médio. Isso foi feito através de um *loop* no próprio algoritmo. Obtido o erro médio e verificada a boa performance da técnica estatística, uma amostra aleatória de treino e de teste, diferente dessas 31 amostras previamente avaliadas, foi selecionada para finalmente gerar o modelo. Se a taxa de erro do modelo gerado for próxima da média, pode-se considerar esse modelo também apresenta boa performance dentro do que era esperado para ele, aumentando ainda mais o nível de confiabilidade.

### 3.7. ELIPSES DE CONFIANÇA E FINALIZAÇÃO DO MODELO

A aplicação das elipses de confiança nos gráficos de probabilidade e de consequências gerados pelas Análises Discriminantes foi feita com objetivo de se determinar com mais exatidão os limites entre as classes. Foi escolhido um valor de  $\alpha$  igual a 0,05 e, portanto, geradas elipses com probabilidade de conter 95% dos pontos de um determinado grupo de dados. A metodologia para obtenção das elipses de confiança foi implementada por Pereira (2016) no *software* R (R Core Team, 2020) e foi adotada nesse trabalho de forma semelhante à apresentada por Santos (2019).

Todas as técnicas estatísticas utilizadas nesse trabalho foram desenvolvidas no *software* R (R Core Team, 2020), visto que o mesmo é um software livre, ou seja, o acesso ao código fonte é livre e possibilita a reprodução da metodologia por outros usuários. Os scripts utilizados para gerar os modelos estão disponíveis na plataforma GitHub, nos links abaixo:

- Para análise de cluster:  
[https://github.com/larissarcs/PAM\\_cluster\\_Likelihood](https://github.com/larissarcs/PAM_cluster_Likelihood)
- Para ACP, análise discriminante e elipses de confiança:  
<https://github.com/larissarcs/PCA-and-Discriminant-Analysis>
- Para avaliação da performance da análise discriminante:  
<https://github.com/larissarcs/Discriminant-Analysis-performance>

### 3.8. CONSTRUÇÃO DO SISTEMA DE ANÁLISE DE RISCO

A classificação de risco é dada através de uma matriz de risco, onde são cruzadas a classificação de probabilidade (*Likelihood*), obtida pelas equações e gráfico gerados pela Análise Discriminante, pela Consequência Final.

A Consequência Final, por sua vez, também é obtida através de uma matriz de consequências, onde são cruzadas as classificações da Consequência A, que é a chance do bloco atingir a rodovia, pela Consequência B, que é a chance do bloco causar um acidente. As equações e gráficos de classificação para a Consequência A foram gerados de forma similar ao modelo de probabilidade. Para a Consequência B, o procedimento foi um pouco diferente e mais simples. Os motivos e as discussões estão apresentados no capítulo de resultados.

# **CAPÍTULO 4**

---

## **RESULTADOS E DISCUSSÕES**

### **4.1. SELEÇÃO DE VARIÁVEIS E PREPARAÇÃO DOS BANCOS DE DADOS**

Ao selecionar as variáveis que deveriam constar nos bancos de dados utilizados nesse trabalho, tomou-se o cuidado de não excluir nenhum parâmetro importante quando se trata de quedas de blocos e suas possíveis consequências. Além do mais, conforme já foi explicado na metodologia, algumas variáveis são sazonais e típicas de regiões como o Colorado; outras dependem de monitoramento constante sobre ocorrência de quedas de blocos e acidentes causados por esse evento, sendo que em muitos locais, incluindo o Brasil, esse monitoramento nem sempre ocorre.

Sendo assim, as variáveis tomadas para as análises foram escolhidas de modo que tenham forte influência no risco de queda de blocos, a fim de gerar um sistema de análise robusto, representativo e seguro, mas que ao mesmo tempo sejam possíveis, fáceis e rápidas de se levantar em qualquer lugar, permitindo que a metodologia seja sim segura e eficiente, mas sem perder a praticidade e rapidez, que são aspectos requeridos em obras lineares como rodovias, onde muitas vezes se tem muitos quilômetros para se avaliar em pouco tempo e nem sempre com muitos recursos. Adiante seguem as variáveis consideradas nas análises com as devidas justificativas.

#### **4.1.1. Probabilidade**

Com relação à Probabilidade (*Likelihood*) de haver queda de blocos, a escolha dos parâmetros se deu de modo a selecionar todas as variáveis possíveis, dentro do banco original levantado e utilizado por Santi *et al.* (2009), que pudessem influenciar na instabilidade dos taludes em rochas cristalinas, que é o foco deste trabalho. Para isso, contou-se com a experiência prática em campo com relação à queda de blocos em taludes rochosos e também com embasamento literário, onde foram observados os parâmetros que tradicionalmente são avaliados em classificações geomecânicas voltadas para qualidade de maciços rochosos e estabilidade de

taludes, além de considerar as causas básicas de quedas de blocos pontuadas por Carvalho (1991), explicadas nos itens 2.2 e 2.3.

Com isso, as variáveis selecionadas foram: Presença de água (P1), Características do maciço (P2), Grau de descalçamento (P3), Tamanho do bloco/volume (P4), Persistência e Orientação das descontinuidades (P5), Abertura das descontinuidades (P6), Alteração/preenchimento das descontinuidades (P7) e Rugosidade das descontinuidades(P8). Observando a Tabela 2.8 (subitem 2.4.1) pode-se notar que foram selecionados praticamente todos os parâmetros relacionados às características geológicas e geotécnicas do maciço rochoso cristalino, com exceção dos parâmetros “alteração da rocha intacta” e “número de famílias”.

A alteração da rocha intacta foi excluída porque a queda de blocos pode acontecer mesmo que a rocha intacta esteja sã. Para o caso de taludes cristalinos, essa não é exatamente uma causa básica. Recorrendo novamente a Tabela 2.8, observa-se que as faixas pontuadas nesse parâmetro são: rocha fresca (3 pontos); descoloração superficial (9 pontos); alteração moderada (27 pontos); alterações internas (81 pontos). Na prática, não haveria uma diferença significativa entre uma rocha sã ou uma rocha com apenas leves descolorações superficiais, pois essas pequenas alterações na rocha intacta não são suficientes para desencadear um problema de queda de blocos, uma instabilização grave no talude.

Por outro lado, uma rocha fortemente alterada, já danificando a estrutura interna, talvez nem deva ser tratada no âmbito das quedas de blocos em taludes rochosos, pois os movimentos de massa nesse tipo de talude podem ser diferentes daqueles causados pelas feições estruturais do maciço e, inclusive, serem desencadeados por causas diferentes, como fortes precipitações e erosões.

Além do mais, em rochas que apresentam permeabilidade fissural, a condição de alteração de uma matriz rochosa decorre principalmente da expansão do processo de intemperismo a partir das descontinuidades, que são os caminhos preferenciais de percolação de água no maciço rochoso. Portanto, uma matriz rochosa alterada implica em condições de alteração severas nas descontinuidades. Somando-se a isso, destaca-se a tendência de que as descontinuidades condicionem o comportamento dos maciços rochosos, justificando a exclusão da variável Alteração da rocha para taludes em rochas cristalinas presente na Tabela 2.8.

Com isso, para queda de blocos em taludes de rocha cristalina, é mais importante considerar a alteração nas paredes das descontinuidades e a presença de preenchimento do que a alteração da rocha intacta, visto que se pode ter um evento catastrófico, com queda de grandes blocos de rocha sã, mas que perderam o atrito entre as paredes das descontinuidades devido à presença de preenchimento argiloso, por exemplo. Com isso, apenas a alteração na parede das juntas e a presença de preenchimento foi considerada.

Com relação ao número de famílias, todas as amostras de taludes cristalinos existentes no banco de dados original apresentavam número de famílias de descontinuidades maior que 2, o que indica que todos os taludes pontuavam 81 nesse parâmetro seguindo a metodologia do CRHRS, ou 4, seguindo a metodologia proposta nesse trabalho. Ou seja, não havia variabilidade, o que inviabiliza a inserção desse parâmetro.

No entanto, isso não é um grande problema nesse caso, visto que já estão incluídas todas as variáveis comumente levantadas para descontinuidades, o que tornaria a inclusão do número de famílias até mesmo redundante. É muito mais importante verificar, por exemplo, se as famílias presentes são persistentes e com orientação desfavorável, se há descalçamento, bem como o tamanho dos blocos (volume) formados, que está intimamente ligado ao espaçamento das descontinuidades. Essas variáveis já são uma indicação que há mais de uma família, ou há uma família crítica junto com descontinuidades ou trincas aleatórias, afinal, está sendo assumida a formação de blocos.

Quanto aos fatores climáticos, que poderiam ser considerados como gatilhos para as quedas de blocos, apenas a condição de água no talude foi mantida. Os parâmetros “ciclos de gelo/degelo anuais” e “aspecto do talude” são relacionados a regiões com nevascas sazonais. Como o objetivo desse trabalho é gerar um sistema de análise de risco que possa ser utilizado nas mais diversas situações e locais, esses dois parâmetros não foram contabilizados.

Já para a precipitação, embora possa ser um gatilho importante em qualquer lugar, as faixas consideradas no banco de dados original (CRHRS) são faixas típicas para o Colorado. Para considerar a precipitação em termos globais, seria necessário um estudo prévio relacionando a ocorrência de quedas de blocos com determinados milímetros de chuva em diversas partes do mundo. É comum encontrar esse tipo de análise para outros movimentos de massa, sobretudo

escorregamentos em maciços de solo e de detritos. Porém, para queda de blocos em taludes rochosos a literatura ainda é carente. Com isso, optou-se por considerar apenas a presença de água, que é um parâmetro que pode ser levantado facilmente em qualquer lugar do mundo e, obviamente, tem forte relação com a precipitação.

A Tabela 4.1 mostra as variáveis consideradas para a análise de probabilidade, com as pontuações atribuídas a cada uma das características que devem ser observadas em campo pelo responsável pelo levantamento. As notas atribuídas a essas características serão utilizadas para calcular os escores das componentes principais, que serão utilizados nas equações geradas pela análise discriminante. O valor será plotado no gráfico de probabilidade e a classificação será obtida. Esse procedimento e essas equações serão detalhados mais adiante. Nota-se que não houve quaisquer mudanças nas características e faixas já adotadas pelo CRHRS.

Tabela 4.1– Parâmetros e suas devidas pontuações para análise de probabilidade.

Parâmetro	Pontuação			
	1	2	3	4
Presença de água (P1)	Seco	Úmido	Gotejando	Com fluxo
Características do maciço (P2)	Homogêneo/ maciço	Pequenas falhas/veios fortes	Xistosidades e zonas de cisalhamento ( $< 15$ cm)	Pegmatitos/ micas/zonas de cisalhamento ( $>15$ cm)
Grau de descalçamento (P3)	0 a 0,3 m	0,3 a 0,6 m	0,6 a 1,2 m	$> 1,2$ m
Tamanho de bloco/volume (P4)	$<0,3$ m/ $<0,75$ m <sup>3</sup>	0,3 a 0,6m/ 0,75 a 2,3m <sup>3</sup>	0,6 a 1,5m/ 2,3 a 7,6m <sup>3</sup>	$> 1,5$ m/ $> 7,6$ m <sup>3</sup>
Persistência e orientação das descontinuidades (P5)	$< 3$ m mergulhando para dentro do talude	$> 3$ m cortando o talude	$< 3$ m mergulhando para fora do talude	$> 3$ m mergulhando para fora do talude
Abertura das descontinuidades (P6)	0	0,1 a 1 mm	1 a 5 mm	$> 5$ mm
Alteração/preenchimento das descontinuidades (P7)	Fresca	Descoloração superficial	Preench. granular	Preench. argiloso
Rugosidade das descontinuidades (P8)	Rugosa	Ondulada	Plana	Estriada



Cabe salientar que dos 224 taludes em rochas cristalinas contidos no banco de dados, após analisar em planilha Excel as variáveis relacionadas à probabilidade, 4 taludes foram excluídos por estarem com valores indefinidos, ou mesmo células vazias em alguma variável. Sendo assim, o banco de dados utilizado para gerar o modelo de probabilidade (*Likelihood*) possui 220 taludes.

#### **4.1.2. Consequências**

As quedas de blocos em taludes rodoviários podem causar diversos transtornos, que incluem interdição temporária da via e consequente congestionamento ou mudança de trajetos, gastos com remoção desses blocos e recuperação, quando necessário, de estruturas danificadas e, em um grau de severidade maior, acidentes com os que ali trafegam; o que significa risco de vida.

Sabendo que risco é dado pela probabilidade e pela consequência de um determinado evento, espera-se que um sistema completo de análise de risco para queda de blocos em taludes rodoviários meça tanto as consequências em um âmbito financeiro, quanto em relação à integridade física e a vida dos motoristas. Isso poderia ser feito através de levantamentos de custos de remoção de blocos, custos de obras de recuperação de estruturas danificadas e número de acidentes causados por blocos na via ou pela queda direta sobre os veículos.

Custos variam com o tempo e com o local onde a análise está sendo feita. Além do mais, o que representa um custo alto para um determinado estado, país ou província, não necessariamente é alto para outro local, o que dificultaria a classificação das consequências, num aspecto financeiro, como severa (alta), média, ou pequena (baixa), de modo que essa classificação seja aplicável para qualquer situação. Isso poderia ser feito em um sistema de análise de risco localizado, mas não global. Quanto ao número de acidentes, nem sempre esse dado está disponível e quando está, nem sempre é representativo.

Dada a importância dessas consequências e vistas as dificuldades no levantamento dos dados mencionados no parágrafo acima, foi necessário pensar em uma forma de contabilizar essas consequências através de parâmetros indiretos que independem de localização, tempo e monitoramento de problemas causados unicamente por quedas de blocos, visto que isso provavelmente ocorrerá apenas em locais em que quedas são frequentes, tornando o banco de

dados não representativo para locais com menos ocorrências, ou nenhuma. No entanto, para criar um modelo de consequências, onde elas podem ser pequenas ou severas, é necessário que o banco de dados seja variado e contenha parâmetros favoráveis e desfavoráveis.

Para que ocorra qualquer tipo de problema com quedas de blocos em rodovias, primeiro é necessário que o bloco atinja a rodovia. Se o bloco não alcança a rodovia, não haverá necessidade de remoção, muito menos acidentes causados por isso, considerando apenas acidentes com quem trafega no local e desconsiderando situações diversas, como uma pessoa andando no pé do talude. Esse não é o foco do trabalho. O simples fato do bloco atingir a via já pode ser considerado como uma consequência, visto que transtornos de ordem financeira já são esperados. Com isso, tem-se a primeira consequência (consequência A), que na verdade é um pré-requisito para qualquer outra, que é a chance do bloco atingir a via.

Embora a consequência A também tenha relação com a chance do bloco causar um acidente, devido à severidade, essa situação merece ser mais aprofundada, de modo que essa consequência seja tratada individualmente.

Os acidentes envolvendo quedas de blocos podem ser causados tanto pela colisão entre o veículo e os blocos, entre dois veículos quando um se desvia desses blocos, ou pela queda direta sobre o veículo. Com isso, tomando uma situação em que ocorre a queda, quanto maior o fluxo viário, maior a chance de acontecer acidente. Assim como a chance aumenta também a depender da distância de visibilidade para que o motorista tome alguma decisão em tempo hábil. A partir daí, tem-se a consequência B, que é então a chance do bloco causar um acidente.

As variáveis que devem ser observadas tanto no modelo para classificação da consequência A quanto para a consequência B são detalhadas adiante. A consequência final é baseada em ambas, através de uma matriz de consequências que será apresentada posteriormente.

Para o caso das consequências optou-se por utilizar tanto as amostras em taludes cristalinos quanto sedimentares, visto que essas variáveis são comuns aos dois tipos de taludes e o mecanismo de queda de blocos também é controlado por causas similares, visto que ambos são rochosos. Além do mais, o modelo de consequência já fica pronto para oportunidades

futuras, onde um modelo de probabilidade de queda de blocos para taludes sedimentares pode ser gerado.

Ainda assim, alguns taludes foram removidos do banco de dados por apresentarem valores inconsistentes ou células vazias. Com isso, o banco de dados reservado para o estudo relacionado com a chance dos blocos alcançarem a rodovia (consequência A) e com a chance dos blocos causarem acidentes (consequência B) possui 264 amostras, sendo 196 de taludes cristalinos e 68 sedimentares.

#### **4.1.3. Consequência A**

O alcance que um bloco de rocha toma durante o movimento de queda depende basicamente de sua energia cinética, da presença de irregularidades na superfície do talude e se existe alguma barreira ou estrutura responsável por captar esses blocos. Quanto maior a energia cinética do bloco, maior a distância que ele consegue alcançar após a queda, supondo a inexistência de barreiras. Essa energia está diretamente relacionada com a altura de queda, pois quanto maior a altura, maior a energia potencial e, conseqüentemente, cinética. Quando se tem um bloco rochoso caindo de um talude muito alto, mesmo após a colisão com a superfície, pode ser que ele ainda recupere boa parte da energia mecânica e continue em movimento, ricocheteando com a superfície, até perder toda a energia e parar.

As colisões, em teoria, podem ser perfeitamente elásticas, quando os materiais que colidem entre si conseguem recuperar toda a energia; ou perfeitamente inelásticas, quando toda essa energia do movimento é transformada em outra forma de energia. Para ilustrar essa situação, imagine um objeto caindo em queda livre de uma altura  $X$  e colidindo com o chão. Se esse objeto, após a colisão conseguisse voltar à altura  $X$ , a colisão seria perfeitamente elástica. Mas se esse objeto parasse imediatamente após a colisão, essa seria perfeitamente inelástica.

Essas situações descritas acima são teóricas e idealizadas. Na prática, o que ocorre após uma colisão entre um objeto e uma superfície qualquer, é uma perda gradual de energia, até que o objeto pare. Isso vai depender não só da magnitude da energia do movimento, como também dos coeficientes de restituição dos materiais que compõem o objeto e a superfície. Quanto mais próximos de 1 são esses coeficientes, maior a taxa de energia recuperada e, portanto,

maiores as chances do movimento continuar. Quanto mais próximos de 0, menor a taxa de energia recuperada após cada colisão.

Na literatura existem algumas tabelas de coeficientes de restituição que mostram que materiais como rochas mais rígidas e resistentes, asfalto e cimento tendem a ter valores mais próximos de 1 para esses coeficientes. Sendo assim, se a rocha que forma o talude e, conseqüentemente o bloco, for uma rocha mais resistente, ao colidir com a própria superfície do talude e com o asfalto, é provável que boa parte da energia seja recuperada após as colisões e que o movimento demore mais para parar.

Como esse trabalho se trata de uma análise de risco de quedas de blocos rochosos em taludes rodoviários, é assumida que a situação descrita acima, colisão entre blocos rochosos rígidos com talude de rocha rígida e asfalto, seja o “problema padrão”. Sendo assim, a altura do talude é um parâmetro crucial para se determinar as chances do bloco atingir a via, pois como é esperado que boa parte da energia potencial e cinética seja recuperada durante a colisão na maioria dos eventos estudados aqui, quanto maior a energia envolvida, mais crítica é a situação. Com isso, a altura do talude é a primeira variável que deve ser analisada para a classificação da consequência A.

Para uma maior compreensão a respeito desses coeficientes de restituição e análises de alcance de blocos, pode-se recorrer aos tutoriais de softwares específicos, como o RocFall da Rocscience® e também aos trabalhos de Stevens (1998), Heidenreich (2004) Asteriou *et al.* (2012), Tavares (2015) e de Silveira (2017), entre outras referências.

Outros parâmetros importantes quando se trata de alcance e trajetória de blocos é o ângulo de inclinação do talude e as condições da superfície, ou seja, presença de cicatrizes, bancos ou qualquer outra irregularidade. Pierson *et al.* (1993) estudaram a influência da inclinação dos taludes no tipo de movimento, como foi explicado anteriormente e ilustrado pela Figura 2.3. A depender do ângulo de inclinação do talude, o bloco pode cair em queda livre, saltar, ou rolar, sendo que essas duas últimas condições podem contribuir para que esse bloco alcance maiores distâncias.

Como já observou Silveira (2017), quando a geometria do talude é muito irregular, as trajetórias dos blocos apresentam maior variabilidade e dificuldade na estimação. De acordo com Pierson *et al.* (2001), taludes com superfícies mais irregulares podem aumentar o alcance de blocos, pois durante a queda, ao se chocar com essas feições estruturais, a velocidade do bloco adquire uma componente horizontal fazendo com que esse bloco se comporte como um projétil, alcançando maiores distâncias.

Diante disso, em uma análise onde se pretende obter as chances do bloco atingir a rodovia, a geometria do talude e suas condições superficiais devem ser levadas em consideração. Inclusive, esses são alguns dos parâmetros necessários para análise de alcance e trajetória de blocos em softwares como o RocFall (Rocscience Inc., 2001). Com isso, a altura do talude, o ângulo de inclinação e as feições de lançamento devem ser também observados e pontuados em campo.

No entanto, as chances do bloco atingir a via diminuem se houver alguma estrutura projetada para captar esses blocos. Com isso, a presença e eficácia da área de captação também devem ser observadas e pontuadas em campo. Conforme já foi explicado, a classificação do CRHRS proposta por Santi *et al.* (2009) pontua a eficácia da área de captação de duas maneiras, sendo adotada a pior situação: uma é de acordo com a relação entre as dimensões medidas em campo e as dimensões ideais, projetadas através do ábaco de Ritchie (1963); a outra maneira é através do ângulo que a parede externa da área de captação faz para dentro da mesma, onde quanto mais fechada é essa área de captação, mais segura ela é.

Contudo, a análise da eficácia da área de captação com base nas dimensões medidas em campo e com as dimensões propostas pelo ábaco de Ritchie (1963) levam em consideração a largura e a profundidade da área de captação, o que torna menos importante a análise do ângulo da parede externa da estrutura. Se uma área de captação é profunda e pouco larga, provavelmente o ângulo da parede externa será maior e a área de captação será mais fechada. Para que ela seja mais aberta, ou deve ser pouco profunda e pouco larga, ou profunda, porém mais larga. Com isso, basta analisar as dimensões.

É importante salientar que o ábaco de Ritchie (1963) projeta áreas de captação levando em consideração a altura e a inclinação do talude, o que reforça a importância dessas variáveis

para se determinar as chances dos blocos atingirem a rodovia. Além do mais, Santi *et al.* (2009) consideram que quando a pontuação atribuída às feições de lançamento (irregularidades no talude) é máxima, que as áreas de captação projetadas através desse ábaco podem não ser eficazes, pois o bloco pode tomar uma trajetória totalmente aleatória e atingir maiores alcances. Isso também será adotado nesse projeto.

Com isso, as variáveis selecionadas para avaliar as chances do bloco atingir a rodovia foram: altura e inclinação do talude, feições de lançamento e eficácia da área de captação. Para as feições de lançamento e também para a altura do talude, as características observadas e as devidas pontuações atribuídas se mantiveram conforme o que foi proposto para o CRHRS modificado de Santi *et al.* (2009). Para a eficácia da área de captação as faixas propostas e descritas na Tabela 2.8 também são mantidas, porém apenas para a eficácia quanto às dimensões. A análise feita para os ângulos da parede da área de captação não estão incluídas nesse trabalho.

Quanto à inclinação do talude, Santi *et al.* (2009) se basearam nos estudos de Maerz *et al.* (2005), que mostravam basicamente que para taludes com inclinações mais próximas de 30° as chances do bloco atingir a via são máximas e vai decrescendo até cerca de 70°, onde as chances são mínimas. Entretanto, para ângulos próximos de 85°, as chances de pequenos blocos atingirem a via também são máximas (ver Figura 2.9).

Essa é uma análise divergente do que sugere o estudo sobre a relação entre a inclinação do talude e a trajetória de blocos de Pierson *et al.* (1993) e também do que é visto na prática, onde blocos que caem de taludes verticais e sub-verticais tendem a alcançar menores distâncias. Além do mais, Santi *et al.* (2009) não detalham as dimensões dos blocos que podem tomar maiores alcances mesmo em taludes sub-verticais. Então, diante dessas divergências optou-se por modificar a pontuação atribuída aos ângulo de inclinação adotado pelo CRHRS modificado, enfatizando taludes sub-verticais.

Considerando as teorias apresentadas por Ritchie (1963) e por Pierson *et al.* (1993) (Figura 2.3) para trajetória e alcance de blocos, blocos que caem de taludes com cerca de 30° tendem a rolar. Blocos que caem de taludes cuja inclinação está entre 30° e 45° podem rolar e saltar, o que pode fazer com que eles alcancem maiores distâncias se não houver barreiras. Blocos que

caem de taludes com inclinação acima de 60° tendem a saltar, sendo que o alcance tende a ser menor para maiores inclinações, sobretudo para superfícies com menos irregularidades. Já em taludes verticais e subverticais, os blocos tendem a apresentar queda livre e, portanto, parar nas proximidades do pé do talude.

Na Tabela 4.2 é mostrado o que é adotado pelo CRHRS e o que será considerado nessa proposta, sendo que as mudanças sugeridas se baseiam nesses trabalhos supracitados e já difundidos mundialmente, além da experiência prática. Ainda na tabela 4.2 é possível perceber que as faixas adotadas por Santi *et al.* (2009) e as faixas sugeridas nesse trabalho são próximas; a mudança se concentra basicamente em taludes com ângulo de inclinação igual a cerca de 85°, que nessa proposta não são considerados como críticos.

Tabela 4.2 – Comparação entre o que é adotado no CRHRS modificado (Santi *et al.*, 2009) com o que é adotado nessa proposta.

Adotado pelo CRHRS modificado (Santi <i>et al.</i> , 2009)				
Escore	3	9	27	81
Ângulos (°)*	Cerca de 90	57 - 80	45 - 57	30 - 45; 85
Adotado nessa proposta				
Escore	1	2	3	4
Ângulos (°)	76 - 90	61 - 75	46 - 60	30 - 45

\*Faixas de ângulos variando de acordo com o que é adotado por Santi *et al.* (2009), apresentados na Tabela 2.8, com base no gráfico de Maerz et al. (2005) (Figura 2.9).

- **Variáveis e faixas consideradas para Consequência A**

Considerando tudo que foi explicado ao longo do subitem 4.1.3, para a consequência A, os parâmetros avaliados e os devidos escores são os que estão apresentados na Tabela 4.3.

Tabela 4.3 – Parâmetros e suas devidas pontuações para análise da consequência A.

Parâmetro	Faixas			
	1	2	3	4
Escores	1	2	3	4
Altura do talude (m) (CA1)	7,5 a 15	15 a 23	23 a 30	> 30
Inclinação do talude (°) (CA2)	76 - 90	61 - 75	46 - 60	30 - 45
Feições de lançamento (CA3)	Nenhuma (talude suave)	Menores	Maiores	Extremas
Área de Captação (CA4)	95 a 100%	65 a 94%	30 a 64%	< 30%

#### 4.1.4. Consequência B

Sabendo que existe um talude com uma considerável probabilidade de gerar queda de blocos e que esses blocos podem atingir a rodovia, as chances desses blocos causarem acidentes aumentam quando se tem os seguintes fatores relacionados à exposição, sobretudo se eles são combinados: maior fluxo diário, pouca visibilidade da rodovia a maiores distâncias e também a extensão do talude.

Sendo assim, variáveis como Risco Médio ao Veículo (RMV) e Percentual de Distância de Visibilidade (PDV), já consideradas no CRHRS modificado (Santi et al., 2009), contemplam as chances de um bloco causar acidentes, tendo em vista que consideram os parâmetros relacionados com a exposição dos usuários da via à quedas de blocos. Além do mais, esses parâmetros são dependentes de medições e/ou de outros parâmetros de fácil obtenção.

O PDV (CB1) é uma variável que mede basicamente se há condições e tempo suficiente para que um motorista consiga realizar alguma manobra caso precise se desviar de um bloco na via, por exemplo. Como já foi explicado no subitem 2.4.1, é dado pela relação entre a distância de visibilidade medida, pela distância ideal para tomada de decisão. A distância de visibilidade medida para tomada de decisão é definida como a distância em que um objeto de 15 cm colocado na beira da estrada é visível para um motorista (Pierson e Van Vickle, 1993).

O RMV (CB2), como mostra a equação 2.20, depende da extensão do talude e do fluxo médio diário, de modo que quanto maior o fluxo e quanto maior a extensão, maior a exposição dos veículos a uma condição de risco. De fato, se um talude suscetível à queda de blocos é extenso, maior será a distância de exposição e se o fluxo é maior, será maior também a chance de algum veículo ser atingido, ou colidir com algum bloco que já esteja na rodovia. Na Tabela 4.4 estão as variáveis (CB1 e CB2) suas respectivas características e escores que devem ser observados para obtenção do índice e classificação da consequência B.

Tabela 4.4 – Parâmetros e suas devidas pontuações para análise da consequência B.

Parâmetro	Faixas			
	1	2	3	4
Escores				
PDV (%) (CB1)	> 80	60 - 80	40 - 60	< 40
RMV (%) (CB2)	0 - 24	25 - 49	50 - 74	> 75



Como se pode perceber, a consequência B se trata na verdade de um caso bivariado. Além do mais, é importante salientar uma particularidade dessa análise relacionada à chance dos blocos causarem acidentes: apesar de ambas as variáveis serem função de parâmetros medidos, há certa subjetividade inerente a eles, pois a chance de ocorrer um acidente pode aumentar de acordo com o comportamento do motorista, como por exemplo, com a não observação dos limites de velocidade, que inclusive entram no cálculo das variáveis (ver subitem 2.4.1). O comportamento dos motoristas, entre outras condições adversas, não são consideradas matematicamente nessa análise.

## 4.2. TESTES ESTATÍSTICOS PRELIMINARES

Para verificar se os bancos de dados são adequados para aplicação das técnicas de estatística multivariada, ou seja, se os bancos de dados apresentam correlações significativas, foi feito o teste de esfericidade de Bartlett. Devido à natureza ordinal das variáveis, foi utilizada a correlação por postos de Spearman. A matriz de correlação das 8 variáveis relacionadas com a probabilidade de queda de blocos é mostrada na Figura 4.1. Nas Figuras 4.2 e 4.3 estão, respectivamente, as matrizes de correlação das 4 variáveis observadas para o cálculo da consequência A e das 2 avaliadas para o cálculo da consequência B. Na Tabela 4.5 estão os resultados dos testes estatísticos.

	P1	P2	P3	P4	P5	P6	P7	P8
P1	1.00000000	0.12765310	0.42426928	0.305455297	0.183845156	0.038575614	0.17369628	0.111890914
P2	0.12765310	1.00000000	0.25045378	0.137530869	0.094991547	0.125831125	0.21418776	0.055583069
P3	0.42426928	0.25045378	1.00000000	0.567468074	0.278354413	0.240716630	-0.07479194	0.143048698
P4	0.30545530	0.13753087	0.56746807	1.00000000	0.338609172	0.367237410	-0.16802045	-0.003333632
P5	0.18384516	0.09499155	0.27835441	0.338609172	1.00000000	0.007100835	0.07388036	0.068674355
P6	0.03857561	0.12583112	0.24071663	0.367237410	0.007100835	1.00000000	0.04156554	-0.054397914
P7	0.17369628	0.21418776	-0.07479194	-0.168020451	0.073880364	0.041565539	1.00000000	0.021433470
P8	0.11189091	0.05558307	0.14304870	-0.003333632	0.068674355	-0.054397914	0.02143347	1.00000000

Figura 4.1– Matriz de correlações das variáveis relacionadas à probabilidade.

	CA1	CA2	CA3	CA 4
CA1	1.00000000	0.08162587	0.4991526	0.1267242
CA2	0.08162587	1.00000000	-0.1173813	-0.2059240
CA3	0.49915263	-0.11738127	1.00000000	0.1864239
CA4	0.12672416	-0.20592405	0.1864239	1.00000000

Figura 4.2– Matriz de correlações das variáveis relacionadas à consequência A.

$$\begin{array}{cc}
 & \begin{array}{cc} \text{CB1} & \text{CB2} \end{array} \\
 \begin{array}{c} \text{CB1} \\ \text{CB2} \end{array} & \begin{bmatrix} 1.0000000 & 0.1036195 \\ 0.1036195 & 1.0000000 \end{bmatrix}
 \end{array}$$

Figura 4.3– Matriz de correlações das variáveis relacionadas à consequência B.

**Tabela 4.5 – Resultados dos testes de esfericidade de Bartlett.**

Banco de Dados	Chi-quadrado ( $X^2$ )	Graus de Liberdade (df)	p-valor
Probabilidade	261,42	28	$9,93 \times 10^{-40}$
Consequência A	105,43	6	$1,84 \times 10^{-20}$
Consequência B	2,82	1	0,09

Um p-valor menor que 0,05, ou seja, próximo de zero, significa que existe mais de 95% de confiança de que existam correlações significativas no banco de dados. Apenas o banco de dados utilizado para avaliar a consequência B apresentou um p-valor maior que 0,05, o que indica que esse banco de dados relacionado às chances dos blocos causarem acidentes pode não ser adequado para aplicação das técnicas de estatística multivariada. Com isso, para a consequência B foi adotada uma outra estratégia, baseada na análise direta dos parâmetros, sem utilizar as técnicas de estatística multivariada.

### 4.3. MODELO DE PROBABILIDADE DE QUEDA DE BLOCOS

#### 4.3.1. Análise de Agrupamento

A Análise de Agrupamento foi feita a fim de definir as classes de probabilidade (*Likelihood*), ou *status*, para aplicação da Análise Discriminante, que foi a técnica utilizada para a criação do modelo. O método utilizado foi o k-medóides, através do algoritmo *PAM*, visto que as variáveis são de natureza ordinal, porém consideradas como numéricas. Na Figura 4.4 é mostrado o gráfico de frequência com o número de indivíduos por escore para cada variável.

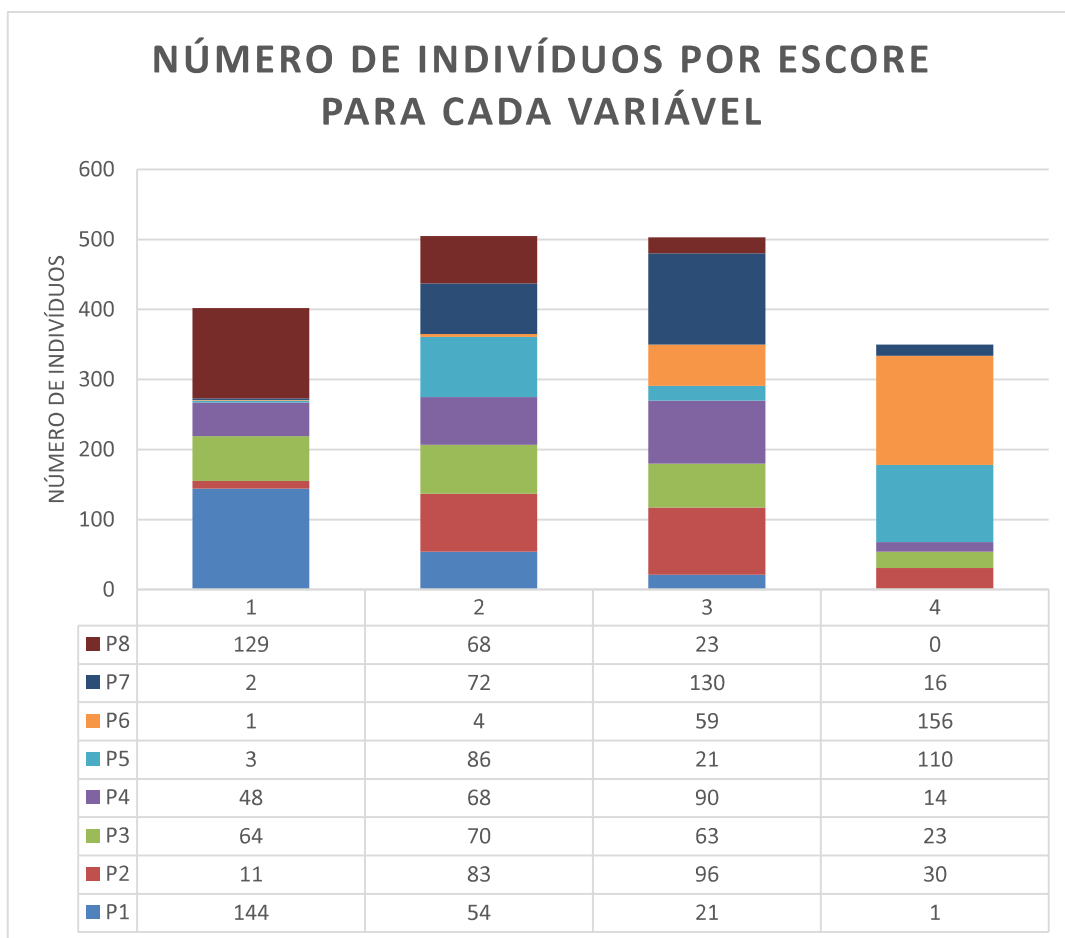


Figura 4.4 – Gráfico de frequência mostrando a variabilidade das variáveis relacionadas à probabilidade.

Observando o gráfico de frequência para os 220 indivíduos que compõem o banco de dados para a probabilidade (Fig. 4.4), é possível notar que o banco de dados apresenta variabilidade; quase todas as variáveis possuem indivíduos pontuando de 1 a 4. Apenas a rugosidade (P8) não apresenta indivíduos pontuando 4 (descontinuidades estriadas) (Tabela 4.1). No entanto, quedas de blocos podem ocorrer, sejam as descontinuidades lisas ou rugosas, especialmente se o grau de descalçamento e as condições de persistência / orientação forem desfavoráveis.

Também é importante observar que algumas variáveis possuem poucos indivíduos em um determinado escore. No entanto, isso reflete as condições frequentemente presentes em taludes rochosos de rodovias. Por exemplo, existem poucos indivíduos com pontuação 4 para a condição da água. No entanto, é difícil encontrar taludes de rocha em rodovias com água escorrendo em excesso em dias não chuvosos. Além do mais, esse parâmetro pode variar

fortemente com as condições climáticas na época em que foram feitos os levantamentos de campo.

Uma situação semelhante ocorre com a abertura, pois é mais comum encontrar discontinuidades abertas do que totalmente fechadas em levantamentos de campo envolvendo quedas de blocos e, normalmente, em análises de campo há uma tendência em se tomar as situações mais críticas, o que pode explicar a maior quantidade de taludes com juntas abertas no banco de dados. Ainda assim, a maioria das variáveis possui indivíduos com pontuação de 1 a 4, o que torna o banco de dados adequado para o prosseguimento das análises.

A tendência de agrupamento foi avaliada através da Estatística Hopkins, que apontou um valor de H igual a 0,37, mostrando que há tendência de formação de agrupamentos significativos. Na Figura 4.5 é mostrado o mapa de calor, que para esse caso não forneceu uma informação precisa a respeito do número de grupos formados.

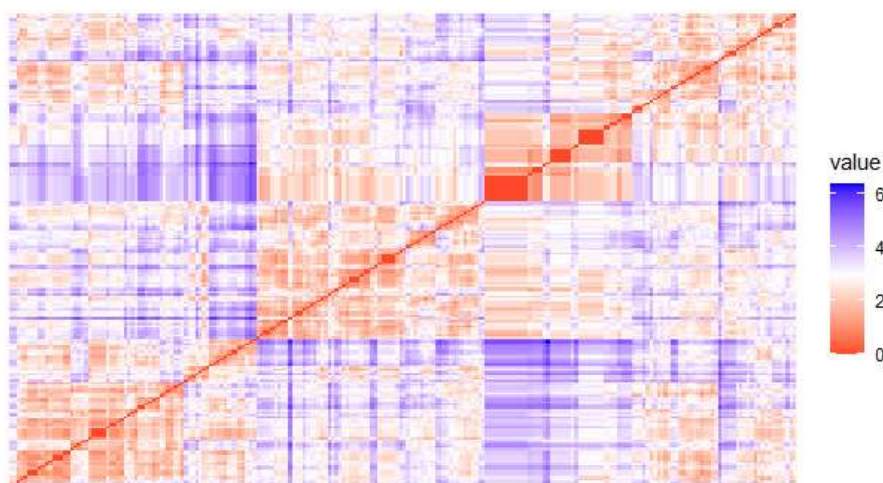


Figura 4.5 - Mapa de Calor das dissimilaridades para as variáveis relacionadas à probabilidade.

Com isso, número ótimo de grupos foi analisado através do método do Cotovelo (Figura 4.6), da Silhueta Média (Figura 4.7) e pela “regra da maioria” (Figura 4.8), que com excessão do método do cotovelo, apontaram que 2 grupos é o ideal estatisticamente.

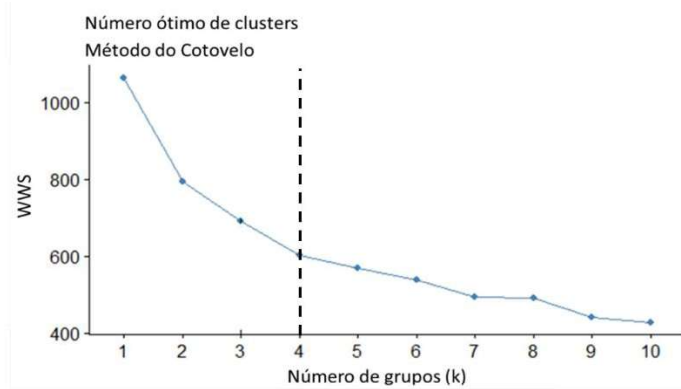


Figura 4.6– Número ótimo de clusters pelo Método do Cotovelo para as variáveis relacionadas à probabilidade.

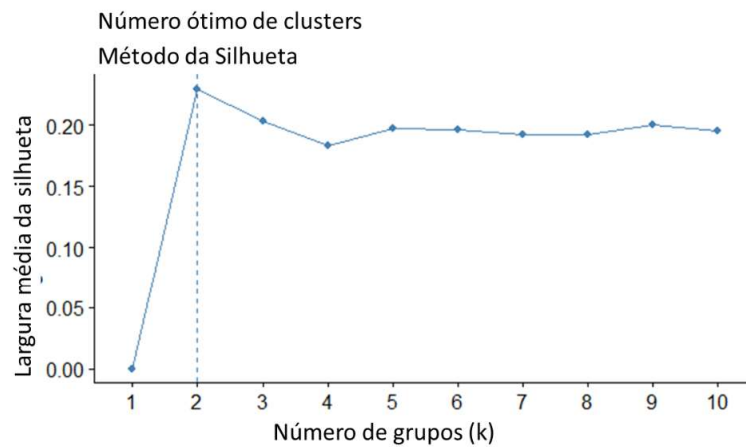


Figura 4.7– Número ótimo de clusters pelo Método da Silhueta para as variáveis relacionadas à probabilidade.

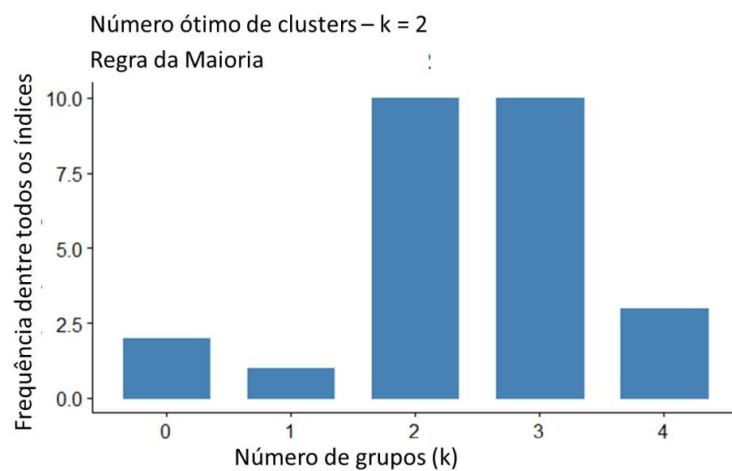


Figura 4.8– Número ótimo de clusters pela Regra da Maioria para as variáveis relacionadas à probabilidade.

Como o objetivo da Análise de Agrupamento nesse caso é definir as classes de probabilidade (*likelihood*) de ocorrência de quedas de blocos como alta, média (zona de transição ou incerteza) e baixa, o ideal tecnicamente seria aplicar a técnica definindo a princípio 3 grupos. No entanto, 2 grupos foi o apontado como ideal em termos estatísticos. Com isso, decidiu-se prosseguir com a análise tanto para 2 grupos quanto para 3 e verificar qual apresentaria o melhor padrão. Para o caso de 2 grupos foi observado se havia alguma zona de interseção entre esses dois grupos e se essa zona poderia ser caracterizada como o grupo médio, ou zona de transição. A análise com 3 grupos também foi feita, visto que era o esperado e, pela Regra da Maioria, também aparece com grande frequência (Figura 4.8).

O agrupamento gerado pelo k-medóides é mostrado na Figura 4.9, sendo que a Figura 4.9(a) mostra o agrupamento para  $k = 2$ , enquanto que o agrupamento para  $k = 3$  é mostrado na Figura 4.9(b). O gráfico de agrupamento é representado pelas duas primeiras componentes principais, mostradas como “Dim 1” e “Dim 2” nos eixos do gráfico. A métrica de distância utilizada foi a Euclidiana.

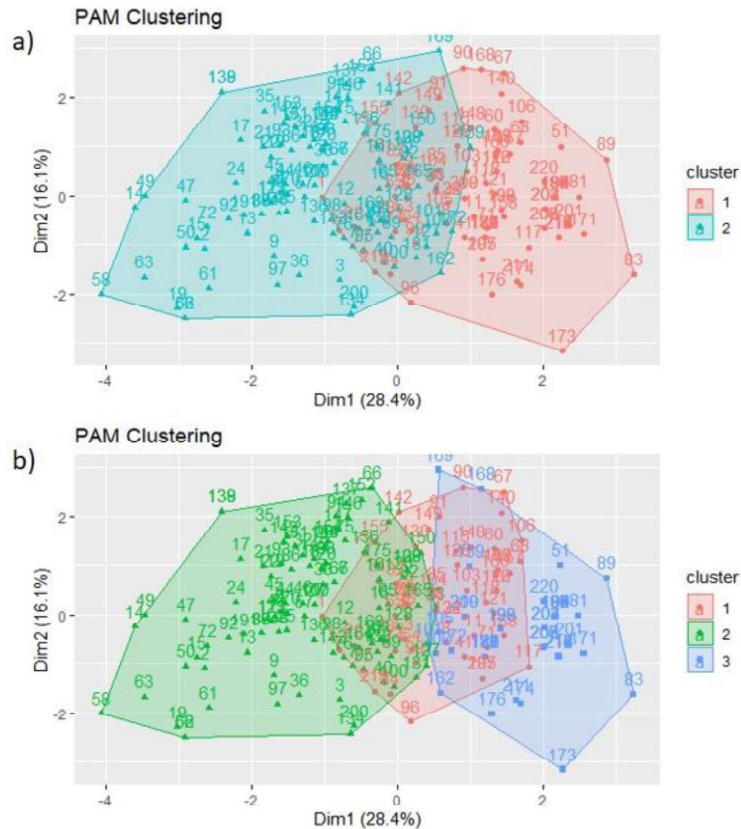


Figura 4.9 – a) Agrupamento gerado pelo R considerando dois grupos; b) Agrupamento gerado pelo R considerando três grupos.

Na Figura 4.9(a) pode ser verificado que existe uma zona de interseção entre o grupo 1 e 2. Na Figura 4.9(b) é possível perceber que o grupo 1 é na verdade um grupo que se sobrepõe aos grupo 2 e 3 e que os limites do grupo 1 no agrupamento com  $k = 3$  é praticamente igual à zona de interseção no agrupamento feito considerando  $k = 2$ . Com isso, a análise prosseguiu primeiramente considerando o agrupamento com  $k = 2$ .

Na Tabela 4.6 consta o medóide de cada grupo, com as devidas pontuações recebidas em cada parâmetro, onde se pode notar que o medóide do grupo 1 recebe notas mais baixas (1 e 2) na maioria dos parâmetros, enquanto que o medóide do grupo 2 recebe 3 e 4 na maioria das variáveis. Deve-se ressaltar que os medóides mostrados na Tabela 4.6 são os medóides finais do agrupamento com  $k = 2$ , ou seja, o elemento que “representa” o grupo.

Tanto o grupo 1 quanto o grupo 2 receberam 110 taludes, sendo que cada um desses grupos, bem como a zona de interseção, foram estudados a fim de tentar identificar os critérios desse agrupamento e se esses grupos poderiam representar classes de probabilidade de queda. Com isso o banco de dados constando as variáveis juntamente com os grupos em que cada talude foi inserido foi analisado em planilha Excel.

Tabela 4.6 – Taludes medóides de cada grupo referente à probabilidade.

		Escore ou “nota” atribuída, de 1 a 4.							
		Variável							
Grupo	Talude medóide	P1	P2	P3	P4	P5	P6	P7	P8
1	Talude 132	1	3	2	2	2	4	3	1
2	Talude 37	2	3	3	3	4	4	3	2

O resultado do agrupamento gerado pelo R (R Core Team, 2020) foi levado para uma planilha Excel onde cada grupo foi cuidadosamente analisado. Com isso, foi possível concluir que o critério de agrupamento pode ser dado pela soma final dos escores atribuídos às variáveis, como se pode observar na Tabela 4.7.

Como para se definir a probabilidade de haver queda de blocos são considerados 8 parâmetros, que podem receber pontuações entre 1 e 4, representando respectivamente uma característica favorável a estabilidade e uma desfavorável, a soma mínima possível é 8 e a

máxima é 32. Ou seja, um talude com somatório igual a 8 tende a ter probabilidade baixa de apresentar queda de blocos, ao contrário de um com somatório igual a 32. É importante informar que no banco de dados não existiam taludes pontuando entre 8 e 12 e nem entre 30 e 32.

Tabela 4.7– Resultado da Análise de Agrupamento para a probabilidade gerado pelo algoritmo R.

Grupo	Faixa de somatório	Classe de probabilidade Inferida
1	13 - 22	Baixa
2	17 - 29	Alta
Zona de Interseção	17 - 22	Média/Zona de Transição

Observando os resultados do agrupamento mostrados na Tabela 4.7, nota-se que no grupo 1 se concentram os taludes com menores somatórios e no grupo 2 os taludes com somatórios maiores, sendo que existe uma zona de interseção referente a esses somatórios, que está na faixa de 17 a 22, ou seja, taludes que pontuam entre 17 a 22 podem estar em qualquer um dos grupos. Com isso, foi feita uma análise referente a essa zona de interseção, a fim de saber, para cada escore entre 17 e 22, quantos taludes ficaram no grupo 1 e quantos ficaram no grupo 2. O resultado é mostrado na Tabela 4.8.

Tabela 4.8– Quantidade de taludes por grupo para taludes que somam entre 17 e 22.

Escore da zona de interseção	Número de taludes que se alocaram no grupo 1	Número de taludes que se alocaram no grupo 2
17	21	2
18	14	0
19	13	11
20	9	15
21	6	15
22	1	27

A situação descrita acima (Tabela 4.8) mostra três padrões. O primeiro é que taludes que pontuam 21 e 22 tendem a se alocar no grupo 2, que é o grupo de taludes com maiores somatórios. O segundo é que há de fato uma incerteza a respeito dos taludes que pontuam 19 e 20, ou seja, não se sabe ao certo em qual grupo esses taludes devem entrar. O terceiro é que



taludes que pontuam 17 e 18 deveriam pertencer ao grupo 1, que é o grupo com menores somatórios.

Observando primeiramente os taludes que pontuam 22, foi possível perceber que a maior parte deles recebe escores 3 e 4 em 5 a 6 variáveis, dentre 8 variáveis. Isso significa que a maioria das variáveis apresenta uma condição desfavorável à estabilidade. Situação praticamente idêntica acontece com os taludes que pontuam 21. Com isso, considerar que esses taludes são zonas de incerteza em vez de considerá-los como taludes de alta probabilidade de ocorrência de quedas de blocos seria admitir uma condição perigosa em taludes permanentes de obras civis, o que não corresponde ao que pregam as boas práticas de segurança na engenharia geotécnica. Portanto, taludes que pontuam 21 e 22 foram considerados como de alta probabilidade.

Com relação aos taludes que pontuam 17 e 18, que estatisticamente tendem a pertencer ao grupo 1, que é o grupo de baixa probabilidade de queda, foi feita a mesma análise que foi feita para aqueles que pontuam 21 e 22. Os taludes que pontuam 18 chegaram a apresentar até 4 variáveis com escores 3 e 4, sendo as demais pontuando 1 e 2. Taludes que pontuam 17 apresentaram até 3 variáveis com escores 3 e 4.

A depender de quais variáveis pontuam 3 e 4, a situação pode ser crítica quanto à estabilidade mesmo que apenas 3 dessas variáveis apresentem esses escores. Para ilustrar, basta imaginar um talude que tenha escores 3 ou 4 em variáveis como persistência e orientação das descontinuidades, grau de descalçamento e abertura. Nessas condições, pode-se inferir que há formação de blocos, que eles estão descalçados e desprendidos, retirando qualquer efeito do atrito. Nesse caso, mesmo que as demais variáveis pontuem 1 ou 2, não se pode afirmar que esse talude é seguro.

O contrário também é válido. Pode ocorrer alguma situação em que as paredes das descontinuidades sejam lisas, que haja certa percolação de água e que haja preenchimento argiloso, por exemplo. No entanto, essas descontinuidades podem não ter orientação favorável à formação de blocos e não serem persistentes; nesse caso é provável que nem mesmo haja descalçamento. Nessas condições, afirmar que esse talude é inseguro seria uma visão excessivamente conservadora que poderia acarretar gastos desnecessários. Diante desse

impasse, é mais prudente considerar os taludes que pontuam 17 e 18 como zona de transição, juntamente com os que pontuam 19 e 20.

Com isso, e sabendo da não existência no banco de dados de taludes que pontuam entre 8 e 12 e entre 30 e 32, a divisão em classes de probabilidade (*Likelihood*) que deu origem ao *status* que foi considerado para a geração do modelo de probabilidade pela Análise Discriminante é dado conforme a Tabela 4.9. Ou seja, a variável dependente é a classe alta, média ou baixa, de cada talude do banco de dados, dada de acordo com a faixa de somatórios dos escores atribuídos a cada variável.

Deve-se destacar que apesar de não existir, nesse banco de dados utilizado para gerar o modelo, taludes que pontuam entre 8 e 12 e entre 30 e 32, essas situações podem ocorrer na prática e, portanto, elas devem ser contabilizadas, conforme as tendências de agrupamento. Afinal, se todos os taludes que pontuaram 13, por exemplo, se alocaram no grupo 1, é provável que taludes que pontuam menos que isso também ficariam no grupo 1 caso fosse feita uma análise de agrupamento com um banco de dados mais completo. Do mesmo modo, taludes que pontuam mais de 29 provavelmente estariam no grupo 2.

Tabela 4.9 – Classes de probabilidade obtidas após análise final do agrupamento gerado pelo algoritmo R.

Grupo	Faixa de somatório	Classe de Probabilidade Adotada
1	8 - 16	Baixa
2	21 - 32	Alta
Zona de Interseção	17 - 20	Média/Zona de Transição

Mesmo que os resultados mostrados acima para o agrupamento com 2 grupos tenham sido satisfatórios, os resultados do agrupamento com 3 grupos também foram analisado em planilha Excel. Na Tabela 4.10 está o resultado do agrupamento gerado pelo R (R Core Team, 2020).

Tabela 4.10– Resultado da Análise de Agrupamento para a probabilidade para  $k = 3$ .

Grupo	Faixa de somatório	Classe de Likelihood Inferida
1 (interseção)	15 - 21	Média/Zona de Transição
2	19 - 29	Alta
3	13 - 20	Baixa

Observando os resultados do agrupamento mostrados na Tabela 4.10, nota-se que o grupo 1 é praticamente a mesma zona de interseção formada no agrupamento para  $k = 2$  (Tabela 4.7) e que todos os taludes alocados no grupo 1 podem pertencer também a outro grupo, no caso, ou ao grupo 1 e 2, ou ao grupo 1 e 3, ou mesmo ao grupo 1, 2 e 3. No grupo 2 estão os taludes com somatórios maiores; no grupo 3 estão os taludes com somatórios menores. Com isso, foi feita também uma análise referente a essa zona de interseção, a fim de saber, para cada escore entre 15 e 21, quantos taludes ficaram no grupo 1, quantos ficaram no grupo 2 e quantos ficaram no grupo 3. O resultado é mostrado na Tabela 4.11.

Tabela 4.11– Quantidade de taludes por grupo para taludes que somam entre 15 e 21 para  $k = 3$ .

Escore da zona de interseção	Número de taludes que se alocaram no grupo 1	Número de taludes que se alocaram no grupo 2	Número de taludes que se alocaram no grupo 3
15	4	0	15
16	10	0	14
17	9	0	14
18	11	0	3
19	12	9	3
20	9	13	2
21	6	15	0

Os resultados do agrupamento para  $k = 3$  reforçam as discussões feitas anteriormente, ainda para o agrupamento com  $k = 2$ , pois ao observar a Tabela 4.11, nota-se que de fato há certa indefinição entre os taludes que pontuam 19 e 20, que poderiam estar alocados tanto no grupo 1 (intermediário), quanto no grupo 2 (maiores somatórios) e até mesmo no grupo 3 (menores somatórios). Isso significa que é mais prudente manter esses taludes em uma zona de transição. Esses resultados também mostram que taludes que pontuam 21 e 22 devem ser considerados como de alta probabilidade de ocorrência de queda de blocos, tanto que taludes que pontuam 22 já nem entraram em um grupo intermediário considerando  $k = 3$  (Tabela 4.10).

Com relação aos taludes que pontuam 17 e 18, as discussões feitas anteriormente continuam válidas. Já os taludes que pontuam entre 15 e 16 devem ser mantidos no grupo com baixa probabilidade de queda, que é inclusive o que mostram as tendências de agrupamento quando se compara os agrupamentos feitos para 2 e 3 grupos. Seria muito conservador coloca-los em

uma zona de transição. Portanto, o resultado final das análises de agrupamento e a definição dos *status* para Análise Discriminante permanecem conforme a Tabela 4.9.

Vale ressaltar que o resultado da Análise de Agrupamento, por si só, já é bastante relevante, pois conseguiu determinar as classes de probabilidade (*Likelihood*) que um talude tem à queda de blocos através de 8 variáveis, sendo essas classes devidamente justificadas. Mas até esse momento, todas as variáveis ainda possuem o mesmo peso. Por exemplo, persistência e orientação ainda tem o mesmo grau de importância que a característica da rocha que compõe o maciço. Daí uma das justificativas em se aplicar as demais técnicas cujos resultados são mostrados logo adiante. Através dessas técnicas é que se tem um modelo de classificação quantitativo.

#### 4.3.2. Análise de Componentes Principais

Apesar de nesse trabalho todas as variáveis serem ordinais podendo ter apenas valores atribuídos iguais a 1, 2, 3 e 4, optou-se por utilizar a matriz de correlações para extrair as componentes, a fim de assegurar que de fato nenhuma variável com maior variabilidade impacte mais nos resultados das análises. Além do mais, como as variáveis são ordinais, o teste de Esfericidade de Bartlett foi feito a partir da matriz de correlação por postos de Spearman. Ao utilizar a matriz de correlação para extrair as componentes, essa matriz também é a matriz de correlação de Spearman. Sendo assim, as análises ficam mais consistentes umas com as outras.

A Análise de Componentes Principais gera novas variáveis não correlacionadas a partir da combinação linear dos autovetores normalizados da matriz de correlações. Como para o modelo de probabilidade são utilizadas 8 variáveis, foram geradas 8 componentes. A Tabela 4.12 apresenta a proporção das variâncias explicadas por cada uma das componentes geradas e nas Equações 4.1 a 4.8 são dadas as componentes principais ( $Y_i$ ).

Tabela 4.12– Variância explicada pelas componentes principais geradas na análise para probabilidade.

Componente Principal	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Proporção da Variância Explicada	0,284	0,161	0,137	0,115	0,105	0,096	0,054	0,048
Proporção acumulada	0,284	0,445	0,582	0,697	0,802	0,898	0,952	1,000

$$Y1 = 0,405 Z1 + 0,254 Z2 + 0,546 Z3 + 0,518 Z4 + 0,334 Z5 + 0,280 Z6 + 0,078 Z7 + 0,098 Z8 \quad (4.1)$$

$$Y2 = 0,238 Z1 + 0,468 Z2 - 0,096 Z3 - 0,337 Z4 + 0,057 Z5 - 0,278 Z6 + 0,702 Z7 + 0,165 Z8 \quad (4.2)$$

$$Y3 = 0,071 Z1 - 0,180 Z2 + 0,121 Z3 - 0,034 Z4 + 0,242 Z5 - 0,552 Z6 - 0,306 Z7 + 0,700 Z8 \quad (4.3)$$

$$Y4 = 0,073 Z1 - 0,211 Z2 - 0,102 Z3 + 0,056 Z4 + 0,655 Z5 - 0,416 Z6 + 0,046 Z7 - 0,574 Z8 \quad (4.4)$$

$$Y5 = 0,713 Z1 - 0,511 Z2 + 0,088 Z3 - 0,053 Z4 - 0,422 Z5 - 0,121 Z6 + 0,102 Z7 - 0,127 Z8 \quad (4.5)$$

$$Y6 = 0,127 Z1 + 0,575 Z2 + 0,153 Z3 + 0,047 Z4 - 0,342 Z5 - 0,445 Z6 - 0,457 Z7 - 0,321 Z8 \quad (4.6)$$

$$Y7 = 0,494 Z1 + 0,214 Z2 - 0,596 Z3 - 0,162 Z4 + 0,255 Z5 + 0,324 Z6 - 0,385 Z7 + 0,099 Z8 \quad (4.7)$$

$$Y8 = 0,025 Z1 - 0,016 Z2 + 0,530 Z3 - 0,763 Z4 + 0,193 Z5 + 0,220 Z6 - 0,194 Z7 - 0,113 Z8 \quad (4.8)$$

Onde:

Y<sub>i</sub>: componente principal i, sendo i = 1, 2, 3, ..., 8.

Z<sub>i</sub>: valor padronizado dos escores de 1 a 4 de cada variável relacionada à probabilidade (Equação 4.9).

$$Z_i = \frac{P_i - \mu_i}{\sigma_i} \quad (4.9)$$

Onde

P<sub>i</sub>: escores de 1 a 4 (Tabela 4.1) de cada variável, i = 1, 2, 3, ..., 8.

μ<sub>i</sub>: média da variável P<sub>i</sub>, i = 1, 2, 3, ..., 8 (Tabela 4.13).

σ<sub>i</sub>: desvio padrão da variável P<sub>i</sub>, i = 1, 2, 3, ..., 8 (Tabela 4.13).

Tabela 4.13 – Valores de média e desvio padrão para cada variável relacionada à probabilidade.

Variável	P1	P2	P3	P4	P5	P6	P7	P8
μ <sub>i</sub>	1,450	2,659	2,205	2,318	3,082	3,682	2,727	1,518
σ <sub>i</sub>	0,684	0,775	0,979	0,886	0,971	0,531	0,603	0,679

Então, para calcular os escores das componentes principais basta aferir em campo a nota de 1 a 4 (P<sub>i</sub>) dada a cada uma das 8 variáveis de acordo com a Tabela 4.1. Posteriormente, essas

notas devem ser padronizadas de acordo com a Equação 4.9 e com os dados da Tabela 4.13. Por fim, esses valores padronizados ( $Z_i$ ) são inseridos nas Equações 4.1 a 4.8.

Isso foi feito para cada um dos 220 taludes presentes no banco de dados utilizado para a criação do modelo de probabilidade (*Likelihood*). Na Tabela 4.14 é dado o exemplo de um dos taludes do banco de dados, onde é mostrada a nota atribuída de acordo com as características descritas na Tabela 4.1, os valores padronizados dessas notas e os escores das componentes principais calculados.

Tabela 4.14 – Exemplo de cálculo dos escores das componentes principais para um talude do banco de dados referente à probabilidade.

Variável ( $P_i$ )	P1	P2	P3	P4	P5	P6	P7	P8
Nota atribuída	3	4	3	3	4	4	3	1
Nota padronizada	2,266	1,731	0,813	0,770	0,946	0,599	0,452	-0,763
Componente Principal ( $Y_i$ )	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Valor de $Y_i$	2,646	1,091	-0,855	0,592	0,432	0,882	1,065	0,185

### 4.3.3. Análise Discriminante

Com as variáveis devidamente quantificadas através da Análise de Componentes Principais para todos os taludes do banco de dados e obtidas as classes de probabilidade de acontecer queda de blocos para cada um desses taludes através da Análise de Agrupamento, foi possível finalmente aplicar a Análise Discriminante, que é a técnica estatística utilizada de fato para gerar a regra de classificação, que poderá ser aplicada posteriormente em qualquer talude cuja condição de estabilidade é desconhecida a priori. Ou seja, a Análise Discriminante foi aplicada no “novo” banco de dados, cujas variáveis independentes são os escores das componentes principais e a variável dependente (*status*) é a classe de probabilidade de ocorrência de queda de blocos (alta, média e baixa).

Para aplicar a Análise Discriminante utilizando as funções canônicas de Fisher, é requerido que os dados sejam homocedásticos; para o caso em que isso não ocorre, é recomendado que

se avalie a aplicabilidade de funções quadráticas. No entanto, as funções quadráticas são sensíveis à normalidade multivariada dos dados e, se os dados não forem normais multivariados, é preferível utilizar as funções canônicas de Fisher mesmo que a homocedasticidade não seja verificada. Para verificar se os dados são homocedásticos foi utilizado o teste M Box (Box, 1949) e para verificar se há normalidade multivariada, foi aplicado o teste de normalidade de Royston. Os resultados são apresentados na Tabela 4.15.

Tabela 4.15 – Teste M Box e Royston para o banco de dados referente ao modelo de probabilidade.

Teste	p-valor
M Box	$2,2 \times 10^{-16}$
Royston	$5,27 \times 10^{-10}$

Como o tanto o p-valor do teste M Box quanto do teste de Royston são menores que 0,05, é possível afirmar que os dados não são nem homocedásticos e nem normais multivariados. Com isso a análise prosseguiu utilizando as funções canônicas de Fisher.

Conforme explicado no subitem 3.6 da Metodologia, antes de aplicar a Análise Discriminante para gerar a regra de classificação definitiva proposta nesse trabalho, composta por equações discriminantes e gráfico discriminante, foi verificada a performance da técnica como um todo, quando aplicada em 31 diferentes amostras de treino e de teste, visto que os resultados são diferentes quando aplicados em amostras diferentes. Isso foi feito através de ressubstituição com divisão amostral. Ou seja, 75% do banco de dados (165 indivíduos) foi separado apenas para treinar o modelo e 25% (55 indivíduos) apenas para testar o modelo.

Isso significa que para avaliar a performance da técnica de uma forma geral, o banco de dados foi separado na proporção de 75% e 25% de 31 maneiras diferentes (31 subdivisões diferentes). Logo, foram geradas 31 regras de classificação diferentes e o erro para cada uma delas foi obtido a fim de se conhecer o erro médio. A Figura 4.10 mostra o erro obtido para cada uma das 31 amostras de teste.

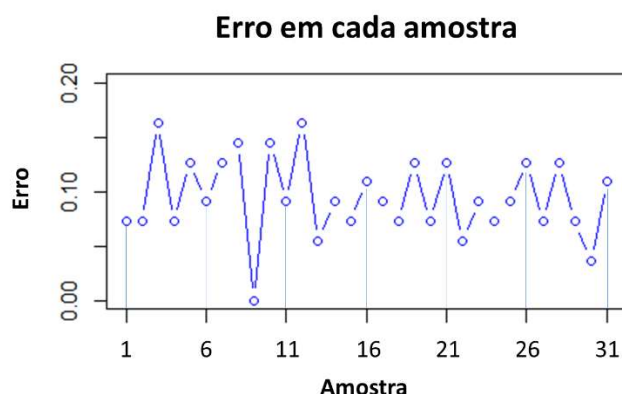


Figura 4.10 – Erro para as 31 diferentes divisões amostrais.

Dentre as amostras testadas, o maior erro obtido foi de 0,164 (16,4%) para as amostras 3 e 12; o menor erro obtido foi 0 (0%), para nona amostra. O erro médio foi de 0,095 (9,5%). Levando em consideração todas as incertezas inerentes aos levantamentos de campo e a complexidade do movimento de queda de blocos, pode-se considerar que essa taxa de erro médio é baixa e que a Análise Discriminante apresenta boa performance e é adequada para gerar um modelo de classificação confiável para prever probabilidade de queda de blocos. Santos (2019) utilizou Análise Discriminante para criar sistema de classificação de perigo para taludes de mina, considerando algumas variáveis semelhantes às consideradas nesse trabalho. O sistema proposto por Santos (2019) apresentou uma taxa de erro de 11,36%.

Com isso, foi fixada uma amostra específica de treino (75% do banco de dados) e teste (25% do banco de dados), diferente das subdivisões avaliadas acima, para aplicação da Análise Discriminante com a finalidade de gerar a regra de classificação da probabilidade de queda de blocos proposta nesse trabalho e validá-la, respectivamente. A regra de classificação gerada é dada pelas equações 4.10 e 4.11. Na Figura 4.11 é mostrado o gráfico de classificação referente aos escores discriminantes gerados pela Análise Discriminante, ainda sem os limites predefinidos. Esses limites serão dados pela aplicação das Elipses de Confiança no subitem seguinte.

$$LD1 = 1,505 Y1 + 0,317 Y2 + 0,187 Y3 + 0,031 Y4 - 0,318 Y5 - 0,113 Y6 + 0,013 Y7 + 0,106 Y8 \quad (4.10)$$

$$LD2 = 0,029 Y1 + 0,136 Y2 - 0,078 Y3 - 0,325 Y4 + 0,657 Y5 - 0,820 Y6 + 0,258 Y7 + 0,313 Y8 \quad (4.11)$$

Onde:



$Y_i$ : escores das componentes principais  $i$ , sendo  $i = 1, 2, 3, \dots, 8$ .

$L_{di}$ : escores das funções discriminantes, sendo  $i = 1$  e  $2$ .

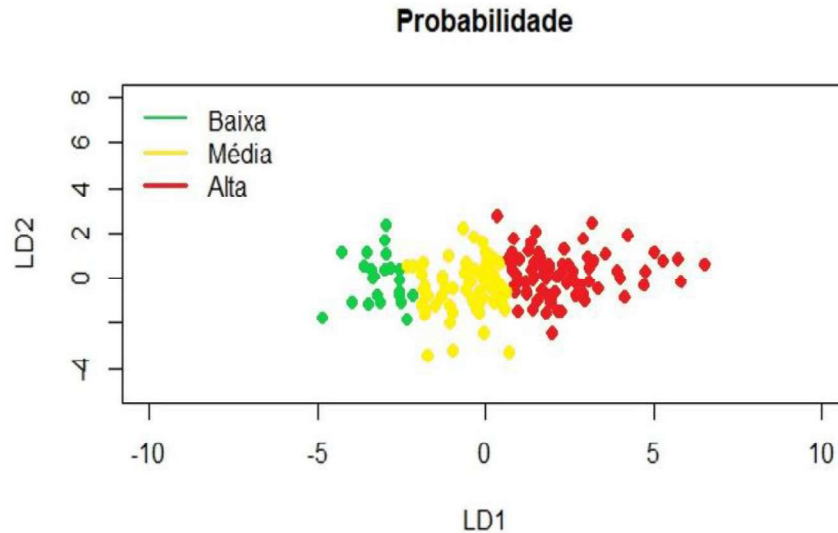


Figura 4.11– Gráfico de probabilidade gerado pela Análise Discriminante, ainda sem os limites entre classes.

Observando o gráfico gerado pela Análise Discriminante, é possível perceber que basicamente quanto maior o valor de LD1, maiores são as chances de um talude apresentar alta probabilidade de ocorrência de quedas de blocos. Ao se analisar LD1 (Equação 4.10), nota-se que a componente que possui o maior coeficiente positivo é  $Y_1$ , ou seja, quanto maior  $Y_1$ , maior tende a ser LD1.

$Y_1$  pode ser considerada como um índice global de queda de blocos, visto que quanto maiores os valores de  $Z_i$ , maior também será  $Y_1$ , para  $i = 1, 2, 3, \dots, 8$  (Equação 4.1), pois todas as variáveis padronizadas  $Z_i$  são acompanhadas de coeficientes positivos. Ainda observando a Equação 4.1, percebe-se que apesar de todas as variáveis contribuírem positivamente para a ocorrência da queda de blocos, as variáveis  $Z_1, Z_3, Z_4$  e  $Z_5$  são as que possuem maior peso, pois são acompanhadas dos maiores coeficientes. Isso é um indicativo de que as variáveis mais influentes no modelo são  $P_1$  (presença de água),  $P_3$  (grau de descalçamento),  $P_4$  (tamanho do bloco/volume/espacamento) e  $P_5$  (persistência e orientação), pois se essas variáveis recebem nota 3 ou 4, os valores padronizados  $Z_1, Z_3, Z_4$  e  $Z_5$  também são maiores, o que aumenta  $Y_1$  e, conseqüentemente, LD1.

Com isso, pode-se inferir que a condição estrutural do maciço rochoso, aliado à presença de água e ao descalçamento são os fatores que mais pesam para determinar se um talude possui alta, média ou baixa probabilidade de ter problemas com quedas de bloco. Lembrando que a componente Y1 é sempre a que possui maior porcentagem da variância explicada e que a função discriminante LD1 também é a que possui maior poder de discriminação, validando ainda mais os pesos das variáveis.

Isso está de acordo com o que foi dito por Carvalho (1991) (Figura 2.2), que as principais causas para quedas de blocos em taludes de rochas cristalinas são justamente as questões estruturais e o descalçamento de blocos. Sendo assim, o modelo proposto é capaz de atribuir pesos adequados conforme a importância de cada parâmetro no resultado final, sem deixar de considerar nenhum deles.

Como consequência da discussão anterior, a regra de classificação gerada é adequada para taludes em rochas cristalinas e minimiza parte da subjetividade da classificação quando atribui pesos às variáveis, em vez de considerá-las como de mesma importância para o processo. No entanto, para comprovar a eficiência da técnica e saber de fato o quanto ela é capaz de prever a probabilidade de ocorrência de quedas, é necessário validar o modelo e verificar a taxa de erro para esse modelo específico. Na Tabela 4.16 está o resultado da validação ao testar a regra de classificação na amostra de teste.

Tabela 4.16 – Resultado da validação da regra de classificação para a probabilidade.

	Classificação dada pela regra de classificação gerada pela Análise Discriminante				
	Classe	Baixa	Média	Alta	Total
Classificação real	Baixa	9	4	0	13
	Média	1	17	0	18
	Alta	0	0	24	24
	Total	10	21	24	55

Pela Tabela 4.16 é possível perceber que dos 55 taludes separados para testar a regra de classificação (25% do banco de dados), 5 foram classificados incorretamente, o que significa que a regra tem cerca de 9,1% de chance de errar a classificação de um talude qualquer

quanto a probabilidade de ocorrência de queda de blocos, que é muito próxima da taxa de erro médio de 9,5% quando diferentes amostras foram testadas.

Ainda observando a Tabela 4.16, percebe-se que esses erros tendem a envolver de alguma forma a região de incerteza (probabilidade média), visto que dos 13 taludes que possuíam probabilidade baixa, 4 foram classificados como de probabilidade média de ocorrência de queda; dos 17 que possuem probabilidade média, 1 foi classificado como de probabilidade baixa, o que pode ser um erro mais grave. Com isso, e diante do baixo índice de erro, pouco menor do que a média, pode-se concluir que a regra de classificação gerada é eficaz, visto que os erros se concentram justamente na região de incerteza, justificando inclusive a sua existência nesse tipo de regra de classificação.

#### 4.3.4. Elipses de confiança e gráfico final

O gráfico final para a classificação de probabilidade (*Likelihood*) de haver queda de blocos foi gerado utilizando elipses de confiança. Através dessas elipses que os limites entre uma classe e outra puderam ser evidenciados matematicamente. O nível de significância utilizado foi igual a 0,05, ou seja, cada elipse gerada para cada classe continha 95% de probabilidade de conter os taludes da respectiva classificação. As Equações 4.12, 4.13 e 4.14 representam as equações das elipses de confiança, para as classes baixa, média e alta, respectivamente. LD1 e LD2 são os escores das funções discriminantes.

$$3,633LD1^2 + 1,005LD2^2 + 22,387LD1 + 2,542LD2 + 1,012LD1LD2 + 29,095 = 0 \quad (4.12)$$

$$1,437LD1^2 + 0,903LD2^2 + 2,134LD1 + 0,665LD2 + 0,090LD1LD2 - 4,200 = 0 \quad (4.13)$$

$$0,596LD1^2 + 1,117LD2^2 - 2,735LD1 - 0,008LD2 - 0,148LD1LD2 - 3,246 = 0 \quad (4.14)$$

Na Figura 4.12 são mostradas as elipses e as retas ligando os pontos de interseção, que é o que determina os limites entre uma classe e outra. Os pontos de interseção entre as elipses das classes baixa e média são (-2,145; 1,309) e (-2,194; -1,756). A reta que liga esses pontos de interseção é, portanto, o limite entre a classe baixa e entre a classe média; a equação dessa reta de interseção é dada pela equação (4.15), que é também mostrada na Figura 4.12, sendo que o y da Figura 4.12 corresponde a LD2 e x corresponde a LD1.

Os pontos de interseção entre as elipses das classes média e alta são (0,0808; 1,770) e (0,646; -2,019). A reta que liga esses pontos de interseção é o limite entre a classe alta e entre a classe média; a equação dessa reta é dada pela equação (4.16), que é também mostrada na Figura 4.12, sendo que y corresponde a LD2 e x corresponde a LD1.

$$LD2 = 62,271LD1 + 134,876 \quad (4.15)$$

$$LD2 = -6,705LD1 + 2,312 \quad (4.16)$$

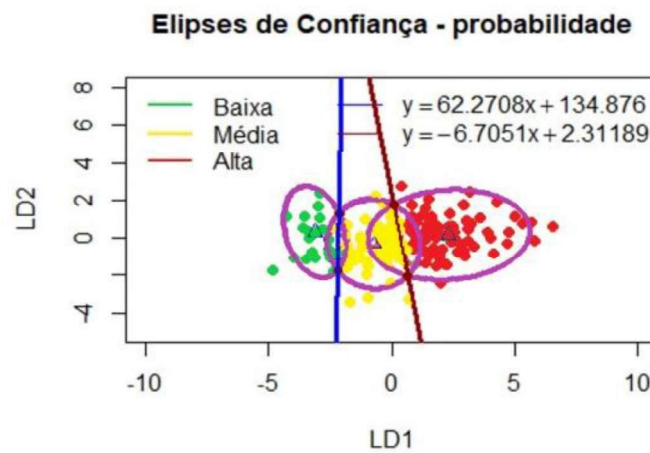


Figura 4.12 – Elipses de confiança e limites entre as classes de probabilidade.

Determinados os limites entre as classes de probabilidade através das Elipses de Confiança, foi possível então gerar o gráfico de classificação quanto a probabilidade (Likelihood) de ocorrência de quedas de blocos (Figura 4.13).

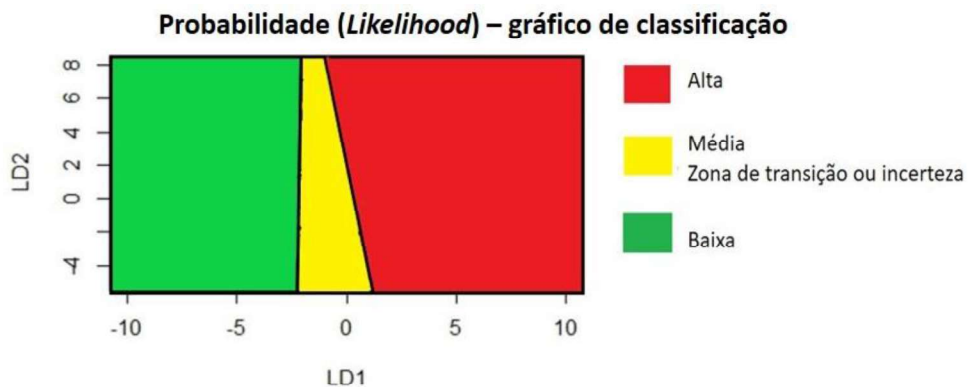


Figura 4.13– Gráfico de probabilidade.

Finalmente, para determinar a probabilidade (*Likelihood*) de ocorrência de quedas de blocos em um talude desconhecido qualquer, basta que o usuário dessa metodologia atribua notas de 1 a 4 para as variáveis P1 a P8 de acordo com as características mencionadas na Tabela 4.1. Essas notas devem ser padronizadas através da Equação 4.9, utilizando os dados da Tabela 4.13, obtendo as variáveis padronizadas Z1 a Z8. Os valores Z1 a Z8 são inseridos nas equações 4.1 a 4.8 a fim de se obter os escores das componentes principais (Yi). Esses por sua vez, devem ser inseridos nas equações geradas pela Análise Discriminante (4.10 e 4.11) para obter os valores de LD1 e LD2. Esses valores são, por fim, plotados no gráfico de probabilidade (Figura 4.13), obtendo a classificação. Na Tabela 4.17 está um exemplo de aplicação com um talude do banco de dados. O manual de instruções de uso do sistema de análise de risco completo está no Apêndice A; no Apêndice B é apresentado o cálculo detalhado desse exemplo mostrado na Tabela 4.17.

Tabela 4.17 – Exemplo de classificação quanto à probabilidade para um talude qualquer.

Variável (Pi)	P1	P2	P3	P4	P5	P6	P7	P8
Nota atribuída	3	3	2	2	4	3	3	2
Nota padronizada	2,266	0,440	-0,209	-0,359	0,946	-1,284	0,452	0,710
Componente Principal (Yi)	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Valor de Yi	0,791	1,732	1,363	0,840	1,104	0,297	1,118	-0,055
Valor de LD1	1,644							
Valor de LD2	0,632							
Classificação	Alta probabilidade ( <i>Likelihood</i> )							

## 4.4. MODELO PARA CONSEQUÊNCIA A

### 4.4.1. Análise de Agrupamento

Assim como para a definição de classes para a probabilidade, a Análise de Agrupamento nesse caso foi feita com o intuito de definir se a chance do bloco atingir a rodovia é alta, média ou baixa, sendo essas classes os *status* para aplicação da Análise Discriminante. O método utilizado também foi o k-medóides, através do algoritmo PAM. Na Figura 4.14 é mostrado o gráfico de frequência com o número de indivíduos por nota (escore) para cada variável, considerando os 264 indivíduos. Lembrando que para gerar o modelo de consequência foram considerados os taludes cristalinos e sedimentares do banco de dados,

visto que são os mesmo parâmetros avaliados, totalizando 264 taludes, conforme explicado no subitem 4.1.2.

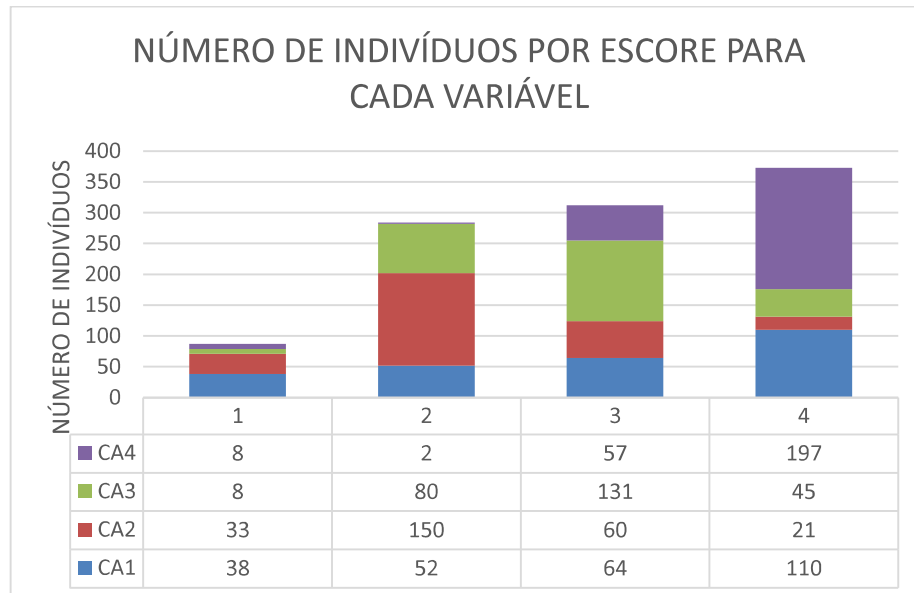


Figura 4.14 – Número de indivíduos por escore para cada variável relacionada à consequência A.

A tendência de agrupamento foi avaliada através da Estatística Hopkins, que apontou um valor de H igual a 0,52. Por ser muito próximo de 0,5, não há a princípio tendência de formação de agrupamentos significativos. No entanto, ao visualizar o mapa de calor (Figura 4.15), se observa que há tendência de agrupamento devido à concentração de retângulos vermelhos na diagonal, embora não esteja claro na figura o número inferido de grupos. Diante disso a Análise de Agrupamento prosseguiu, pois mesmo que estatisticamente o número de Hopkins não aponte formação de grupos significativos, tecnicamente os agrupamentos formados podem levar a discussões e conclusões relevantes.

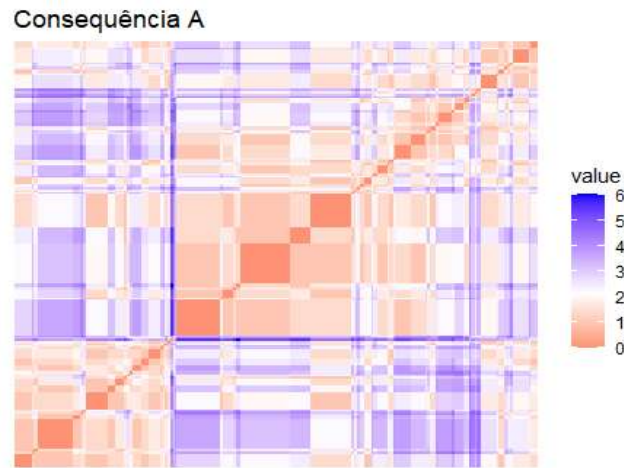


Figura 4.15 - Mapa de Calor para as variáveis relacionadas à consequência A.

O número ótimo de grupos foi analisado através do método do cotovelo (Figura 4.16), da silhueta média (Figura 4.17) e pela “regra da maioria” (Figura 4.18), sendo que o primeiro apontou que 3 é o número ótimo de grupos. O método da silhueta apontou que 2 ou 4 é o número ótimo; já a regra da maioria apontou que 2 grupos é o ideal estatisticamente.

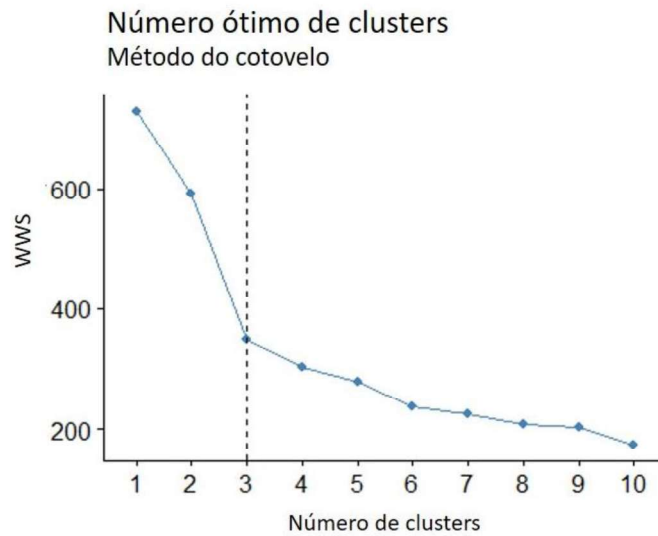


Figura 4.16– Número ótimo de clusters pelo método do cotovelo para as variáveis relacionadas à consequência A.

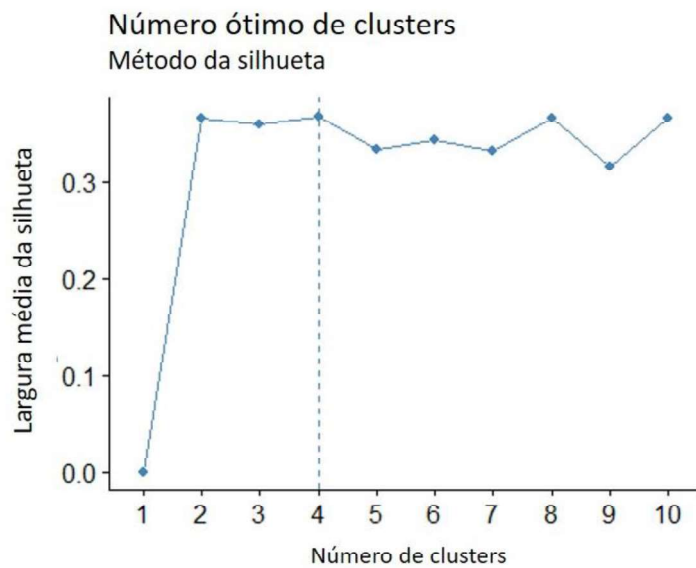


Figura 4.17 – Número ótimo de clusters pelo método da silhueta para as variáveis relacionadas à consequência A.

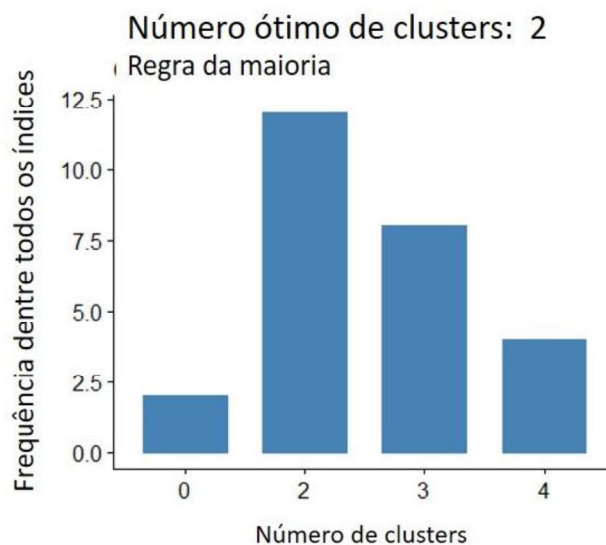


Figura 4.18– Número ótimo de clusters pela regra da maioria para as variáveis relacionadas à consequência A.

Como o objetivo da Análise de Agrupamento nesse caso é definir se a chance dos blocos atingirem a rodovia é alta, média (zona de incerteza) ou baixa, o ideal tecnicamente também seria aplicar a técnica definindo a princípio 3 grupos. No entanto, como o número ótimo de grupos sugeridos pela maioria das técnicas apresentadas acima é dois, a análise prosseguiu primeiramente para dois grupos, visto que poderia haver alguma zona de interseção entre eles



que poderia representar uma classe intermediária, tal qual ocorreu para a análise de probabilidade de ocorrência de quedas de blocos.

O agrupamento gerado pelo k-medóides com  $k = 2$  é mostrado na Figura 4.19. A distância utilizada foi a Euclidiana. Na Tabela 4.18 consta o medóide final de cada grupo, com as devidas pontuações recebidas em cada parâmetro, onde é possível perceber um medóide com pontuações mais baixas nas variáveis (grupo 1) e outro com pontuações mais altas (grupo 2). É importante destacar que a variável CA4 (eficácia da área de captação) recebe pontuações altas, ou seja, 3 e 4, na maior parte dos taludes do banco de dados e como o medóide é um indivíduo representativo, "médio", dos grupos, é esperado que as diferenças de pontuações se concentre mais nas demais variáveis, mas que ainda assim essa diferença não seja tão extrema.

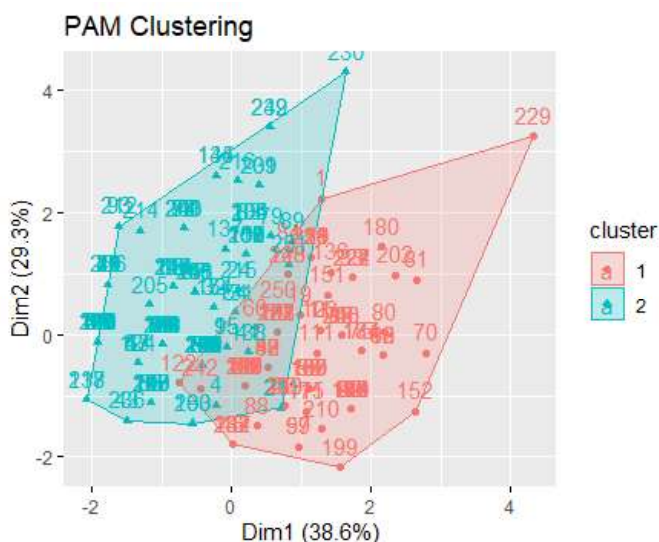


Figura 4.19 – Agrupamento gerado pelo R para as variáveis relacionadas à consequência A.

Tabela 4.18 – Medóides de cada grupo referente à consequência A.

		Escore ou nota atribuída			
		Variável			
Grupo	Talude medóide	CA1	CA2	CA3	CA4
1	Talude 3	2	2	2	4
2	Talude 5	4	2	3	4

O grupo 1 recebeu 108 taludes, já o grupo 2 recebeu 156 taludes. Como se pode notar pela Figura 4.19, de fato há uma zona de interseção entre os grupos 1 e 2 que pode ser uma

possível classe intermediária entre a classe alta e baixa. Cada um desses grupos e a zona de interseção foram estudados a fim de tentar identificar os critérios desse agrupamento e se esses grupos poderiam representar classes para a consequência A. Com isso o banco de dados constando as variáveis juntamente com os grupos em que cada talude foi inserido foi analisado em planilha Excel.

Com as análises feitas em cada talude de cada grupo em planilha Excel, foi possível concluir que o critério de agrupamento, assim como para probabilidade (*Likelihood*), pode ser dado pela soma final dos escores atribuídos às variáveis. Como para se definir a classe para a Consequência A são considerados 4 parâmetros, que podem receber pontuações entre 1 e 4, a soma mínima possível é 4 e a máxima é 16. No entanto, no banco de dados não existiam taludes pontuando entre 4 e 6. Na Tabela 4.19 é mostrado o resultado do agrupamento gerado pelo R. Na Tabela 4.20 é mostrado o resultado final determinado após as análises em planilha Excel.

Tabela 4.19– Resultado da Análise de Agrupamento para a consequência A gerado pelo R.

Grupo	Faixa de somatório	Classe inferida para Consequência A
1	7 - 13	Baixa chance
2	10 - 16	Alta chance
Zona de interseção	10 - 13	Média (Zona de incerteza)

Tabela 4.20– Classes para consequência A obtidas após análise final do agrupamento gerado pelo R.

Grupo	Faixa de somatório	Classe adotada para Consequência A
1	4 - 9	Baixa chance
2	10 - 12	Média (Zona de incerteza)
3	13 - 16	Alta Chance

Observando os resultados do agrupamento mostrados na Tabela 4.19, nota-se que no grupo 1 se concentram os taludes com menores somatórios, no grupo 2 os taludes com somatórios maiores, sendo que há uma zona de interseção, que são os taludes que possuem somatório final de notas entre 10 e 13; esses, podem estar tanto no grupo 1 quanto no grupo 2. Com isso,

foi feita uma análise referente a essas zonas de interseção a fim de saber quantos taludes ficaram no grupo 1 e quantos ficaram no grupo 2. O resultado é mostrado na Tabela 4.21.

Tabela 4.21 – Quantidade de taludes por grupo para taludes que somam 10, 11, 12 e 13.

Escore da zona de interseção	Número de taludes que se alocaram no grupo 1	Número de taludes que se alocaram no grupo 2
10	35	6
11	29	8
12	50	5
13	1	39

A situação descrita acima (Tabela 4.21) mostra claramente que taludes que pontuam 13 devem ser alocados no grupo 2, que são os taludes com maiores somatórios. De fato, para somar 13, pode acontecer de todas as variáveis pontuarem 3 ou 4, como é o caso de um dos taludes do banco que pontua 3 em três variáveis e 4 em uma variável, o que é uma situação crítica. Bem como, pode acontecer de duas variáveis pontuarem 4, uma pontuar 3 e uma pontuar 2, que também é uma situação crítica. Sendo assim, taludes que somam 13 foram considerados como pertencentes ao grupo 2 (Tabela 4.20).

Já para os taludes que somam 12, pode ocorrer de duas variáveis pontuarem 4 e duas pontuarem 2, o que o coloca em uma situação intermediária a depender de quais parâmetros recebem 4 e quais recebem 2. No entanto, também pode acontecer de duas variáveis receberem 4 pontos, uma receber 3 e uma 1, o que já é uma situação mais crítica. Sendo assim, o somatório igual a 12 para a Consequência A pode ser considerado o limite superior do grupo intermediário e, portanto, o resultado gerado pela análise de cluster pode ser respeitado.

Apesar de taludes com nota 9 não estarem em nenhuma região de interseção de acordo com os resultados da análise de agrupamento, é interessante analisar se esses taludes se encaixam bem como limite superior do grupo 1 (baixas chances do bloco atingir a via). Para somar 9 pontos devem acontecer as seguintes situações: três variáveis pontuarem 2 e uma pontuar 3; ou deve acontecer de uma variável pontuar 4, duas pontuarem 2 e uma pontuar 1; ou, em uma situação menos favorável à segurança, duas variáveis pontuarem 3 ou 4 e duas pontuarem 1 ou 2.

Isso significa que taludes que pontuam 9 possuem mais características que podem dificultar que o bloco chegue a rodovia, do que características que facilitem esse processo. Porém, como tem uma situação que pode ser um pouco mais crítica, mas ainda assim intermediária, que é o caso do talude que possui duas variáveis pontuando 3 ou 4 e duas pontuando 1 ou 2, pode-se considerar que esses taludes representam bem o limite superior do grupo com baixas chances do bloco atingir a via.

Sendo assim, a divisão em classes de chance do bloco atingir a via, que deu origem ao *status* que foi considerado para a geração do modelo de consequência A pela Análise Discriminante, é dado conforme a Tabela 4.20. Ou seja, a variável dependente é a classe alta, média ou baixa, de cada talude do banco de dados, dada de acordo com a faixa de somatórios dos escores atribuídos a cada variável.

Houve também a tentativa de fazer a análise de agrupamento via k-medoides considerando três grupos ( $k = 3$ ). No entanto os resultados não foram satisfatórios e não geraram discussões relevantes o suficiente para serem inseridas nesse trabalho.

#### 4.4.2. Análise de Componentes Principais

Assim como foi feito para o modelo de probabilidade, a Análise de Componentes Principais nesse caso também foi realizada utilizando as matrizes de correlações e com o intuito de quantificar as variáveis antes da aplicação da Análise Discriminante.

Como para o modelo de consequência A, ou seja, chance de um bloco atingir a rodovia, são utilizadas 4 variáveis, foram geradas 4 componentes principais. A Tabela 4.22 apresenta a proporção das variâncias explicadas por cada uma das componentes geradas e nas Equações 4.17 a 4.20 são dadas as componentes principais ( $Y_{cai}$ ).

Tabela 4.22– Variância explicada pelas componentes principais geradas na análise para consequência A.

Componente Principal	Yca1	Yca2	Yca3	Yca4
Proporção de variância explicada	0,386	0,293	0,207	0,114
Proporção acumulada	0,386	0,679	0,886	1,000

$$Y_{ca1} = 0,641 ZA1 - 0,121 ZA2 + 0,696 ZA3 + 0,299 ZA4$$

(4.17)

$$Y_{ca2} = 0,367 ZA1 + 0,740 ZA2 + 0,032 ZA3 - 0,562 ZA4 \quad (4.18)$$

$$Y_{ca3} = 0,019 ZA1 - 0,603 ZA2 + 0,208 ZA3 - 0,770 ZA4 \quad (4.19)$$

$$Y_{ca4} = 0,673 ZA1 - 0,272 ZA2 - 0,686 ZA3 + 0,044 ZA4 \quad (4.20)$$

Onde:

$Y_{cai}$ : componente principal  $i$ , sendo  $i = 1, 2, \dots, 4$ .

$ZA_i$ : valor padronizado dos escores de 1 a 4 de cada variável relacionada à consequência A (Equação 4.21).

$$ZA_i = \frac{CA_i - \mu_{cai}}{\sigma_{cai}} \quad (4.21)$$

Onde:

$CA_i$ : escores de 1 a 4 (Tabela 4.3) de cada variável,  $i = 1, 2, \dots, 4$ .

$\mu_{cai}$ : média da variável  $CA_i$ ,  $i = 1, 2, \dots, 4$  (Tabela 4.23).

$\sigma_{cai}$ : desvio padrão da variável  $CA_i$ ,  $i = 1, 2, \dots, 4$  (Tabela 4.23).

Tabela 4.23 – Valores de média e desvio padrão para cada variável relacionada à consequência A.

Variável	CA1	CA2	CA3	CA4
$\mu_{cai}$	2,932	2,261	2,807	3,678
$\sigma_{cai}$	1,091	0,777	0,748	0,646

Então, para calcular os escores das componentes principais basta obter a nota de 1 a 4 ( $CA_i$ ) dada a cada uma das 4 variáveis de acordo com a Tabela 4.3. Posteriormente, essas notas devem ser padronizadas de acordo com a Equação 4.21, utilizando os dados da Tabela 4.23. Por fim, esses valores padronizados ( $ZA_i$ ) são inseridos nas Equações 4.17 a 4.20.

Isso foi feito para cada um dos 264 taludes presentes no banco de dados utilizado para a criação do modelo de consequência A. Na Tabela 4.24 é dado o exemplo de um dos taludes do banco de dados, onde é mostrada a nota atribuída de acordo com as características descritas na Tabela 4.3, os valores padronizados dessas notas e os escores das componentes principais calculados.

Tabela 4.24 – Exemplo de cálculo dos escores das componentes principais para um talude do banco de dados referente à consequência A.

Variável (CAi)	CA1	CA2	CA3	CA4
Nota atribuída	4	2	3	4
Nota padronizada	0,980	-0,336	0,258	0,499
Componente Principal (Ycai)	Yca1	Yca2	Yca3	Yca4
Valor de Ycai	0,998	-0,161	-0,109	0,595

#### 4.4.3. Análise Discriminante

Com as variáveis devidamente quantificadas através da Análise de Componentes Principais para todos os taludes do banco de dados e obtidas as classes referentes à chance dos blocos atingirem a rodovia através da Análise de Agrupamento, foi possível finalmente aplicar a Análise Discriminante, cujas variáveis independentes são os escores das componentes principais e a variável dependente (*status*) é a classe determinada para a Consequência A (alta, média ou baixa).

Para verificar se os dados são homocedásticos foi utilizado o teste M Box (Box, 1949) e para verificar se há normalidade multivariada, foi aplicado o teste de normalidade de Royston. Os resultados são apresentados na Tabela 4.25.

Tabela 4.25– Teste M Box e Royston para o banco de dados referente ao modelo de consequência A.

Teste	p-valor
M Box	$2,2 \times 10^{-16}$
Royston	$7,30 \times 10^{-23}$

Como o tanto o p-valor do teste M Box quanto do teste de Royston são menores que 0,05, é possível afirmar que os dados não são nem homocedásticos e nem normais multivariados. Com isso a análise prosseguiu utilizando as funções canônicas de Fisher.

Assim como para o modelo de probabilidade de queda de blocos, antes de aplicar a Análise Discriminante para gerar a regra de classificação para a Consequência A, foi verificada a performance da técnica quando aplicada em 31 diferentes amostras de treino e de teste. Sendo

assim, para avaliar a performance geral da técnica, o banco de dados foi separado na proporção de 75% (treino) e 25% (teste) de 31 maneiras diferentes (31 amostras diferentes) e o erro para cada um dos modelos foi obtido a fim de se conhecer o erro médio. A Figura 4.20 mostra o erro obtido para cada uma das 31 amostras.

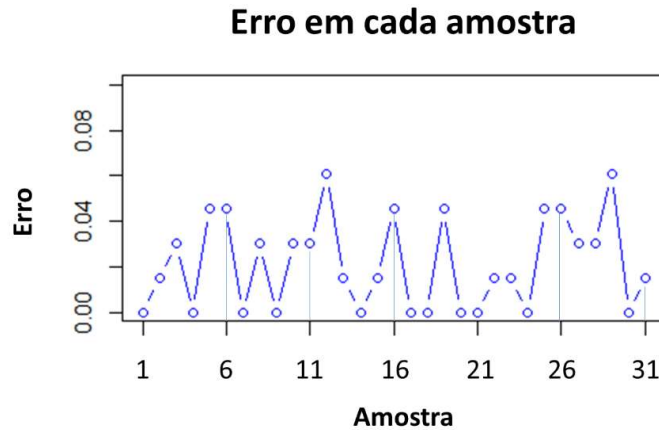


Figura 4.20 – Erro para os 31 diferentes modelos gerados para cada uma das 31 divisões amostrais para a Consequência A.

Dentre as amostras testadas, o maior erro obtido foi de 0,061 (6,1%) para as amostras 12 e 29; o menor erro obtido foi 0 (0%) para as amostras 1, 4, 7, 9, 14, 17, 18, 20, 21, 24 e 30. O erro médio foi de 0,022 (2,2%), que é uma taxa baixa, sobretudo quando se considera as incertezas quanto às chances do bloco atingir de fato a rodovia e a trajetória que um bloco pode tomar em queda. Logo, a Análise Discriminante apresenta boa performance e é adequada para gerar um modelo de classificação confiável para predizer a chance do bloco, ou dos blocos, atingir a rodovia.

Feito isso, a regra de classificação para a Consequência A foi criada considerando uma subdivisão amostral específica de treino e de teste, diferente das 31 avaliadas acima. A regra gerada é dada pelas equações 4.22 e 4.23. Na Figura 4.21 é mostrado o gráfico de classificação dos escores discriminantes.

$$LDca1 = 1,764 Yca1 + 0,808 Yca2 - 0,876 Yca3 + 0,254 Yca4 \quad (4.22)$$

$$LD2ca = 0,079 Yca1 + 0,348 Yca2 + 1,028 Yca3 + 0,688 Yca4 \quad (4.23)$$

Onde:

$Y_{cai}$ : escores das componentes principais  $i$ , sendo  $i = 1, 2, \dots, 4$ .

$LD_{cai}$ : escores das funções discriminantes, sendo  $i = 1$  e  $2$ .

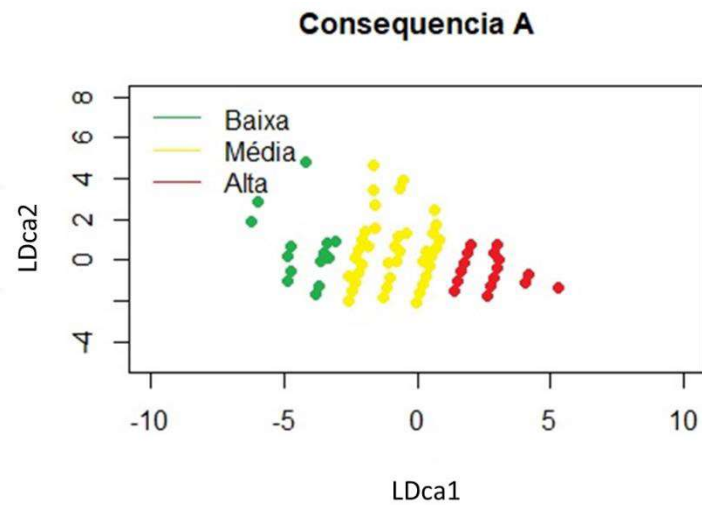


Figura 4.21– Gráfico de chance do bloco atingir a rodovia (Conequência A) gerado pela Análise Discriminante, ainda sem os limites entre classes.

De acordo com o gráfico mostrado na Figura 4.21, quanto maior o valor de  $LD_{ca1}$ , maiores as chances de um ou mais blocos atingirem a rodovia. Observando a função  $LD_{ca1}$  (Equação 4.22), nota-se que quanto maior  $Y_{ca1}$ , maior  $LD_{ca1}$ .  $Y_{ca1}$  (Equação 4.17) é maior principalmente quando as variáveis  $ZA1$  (altura do talude padronizada),  $ZA3$  (feições de lançamento padronizada) e  $ZA4$  (área de captação padronizada) são maiores, com ênfase em  $ZA1$  e  $ZA3$ .

De fato, se um talude é alto, os blocos atingem altos valores de energia potencial (Hoek, 2007) e isso, aliado às irregularidades da superfície do talude mais uma área de captação ineficiente ou inexistente, faz com que as chances de um bloco em queda atingir a rodovia aumentem. Barrett e White (1991) sugerem que inclinações severamente irregulares muitas vezes tornam os projetos de área de captação baseados no Ábaco de Ritchie inaplicáveis, o que reforça a importância de atribuir pesos maiores às feições de lançamento.

De todo modo, mesmo que as variáveis altura de talude ( $CA1$ ) e feições de lançamento ( $CA3$ ) se destaquem, as demais também impactam nas chances do bloco atingir a rodovia. Basta observar que na Equação 4.22,  $LD_{ca1}$  é maior também se for atribuído à componente  $Y_{ca3}$  um valor negativo. Ao observar a Equação 4.19 ( $Y_{ca3}$ ), nota-se que quanto maior o valor das



componentes padronizadas ZA2 (ângulo de inclinação do talude) e ZA4 (eficácia da área de captação), menor (mais negativo) é o valor de  $Y_{ca3}$ , o que mostra que o modelo é capaz de avaliar adequadamente a chance do bloco atingir a rodovia considerando pesos diferentes para as variáveis.

A validação da regra de classificação gerada pela Análise Discriminante também foi feita através de ressubstituição com divisão amostral, sendo 75% das amostras reservadas aleatoriamente apenas para gerar o modelo e 25% apenas para testar o modelo. Na Tabela 4.26 está o resultado do teste.

Tabela 4.26 – Resultado da validação da regra de classificação para a consequência A.

Classificação	Classificação dada pela regra de classificação gerada pela Análise Discriminante				Total
	Classe	Baixa	Média	Alta	
real	Baixa	7	1	0	8
	Média	0	34	0	34
	Alta	0	0	24	24
	Total	7	35	24	66

Pela Tabela 4.26 nota-se que dos 66 taludes separados para testar a regra de classificação (25% dos 264 taludes), 1 foi classificado incorretamente, o que significa que a regra tem cerca de 1,5% de chance de errar a classificação de um talude qualquer quanto a chance dos blocos atingirem a rodovia. Nota-se que essa taxa de erro é próxima da taxa média, sendo ainda menor.

No caso, assim como para a análise de probabilidade, esse erro envolveu a região de incerteza (classe média), visto que dos 8 taludes que inseridos na classe baixa, 1 foi classificado incorretamente como classe média pela análise discriminante, o que não é um erro grave levando em consideração a segurança dos motoristas. Com isso, e diante do baixo índice de erro, pode-se concluir que a regra de classificação gerada é eficaz, visto que os erros se concentram justamente na região de incerteza.

#### 4.4.4. Elipses de confiança e gráfico final

O gráfico final para a classificação referente à consequência A foi também gerado utilizando elipses de confiança. O nível de significância utilizado foi igual a 0,05, ou seja, cada elipse

gerada para cada classe representa 95% de probabilidade dos taludes pertencerem à respectiva classificação. As Equações 4.24, 4.25 e 4.26 representam as equações das elipses de confiança, para as classes baixa, média e alta, respectivamente.

$$2,311LDca1^2 + 18,694LDca1 + 0,668LDca2^2 + 1,046LD1LDca2 + 4,491LDca2 + 37,837 = 6,859 \quad (4.24)$$

$$0,781LDca1^2 + 1,182LDca1 + 0,641LDca2^2 - 0,300LDca1LDca2 - 0,487LDca2 + 0,475 = 5,417 \quad (4.25)$$

$$1,518LDca1^2 - 7,818LDca1 + 3,762LDca2^2 + 2,222LDca1LDca2 - 4,507LDca2 + 10,192 = 6,388 \quad (4.26)$$

Na Figura 4.22 são mostradas as elipses e as retas que delimitam as classes. Os pontos de interseção entre as elipses das classes baixa e média são (-2,865; 1,463) e (-2,171; -2,576). A reta que delimita as classes baixa e média é, portanto, a reta que liga esses dois pontos, cuja equação é dada pela equação (4.27) e também mostrada na Figura 4.22.

Os pontos de interseção entre as elipses das classes média e alta são (1,951; 1,251) e (1,489; -0,943). A reta que delimita as classes média e alta é a reta que liga esses dois pontos de interseção. A equação dessa reta é dada pela equação (4.28) e é também mostrada na Figura 4.22. Lembrando que essas equações das retas mostradas na Figura 4.22 estão em termos de y e x, onde y representa LDca2 e x representa LDca1. Na Figura 4.23, está o gráfico final para determinar a chance dos blocos atingirem a rodovia (consequência A).

$$LDca2 = -5,812LDca1 - 15,190 \quad (4.27)$$

$$LDca2 = 4,748LDca1 - 8,014 \quad (4.28)$$

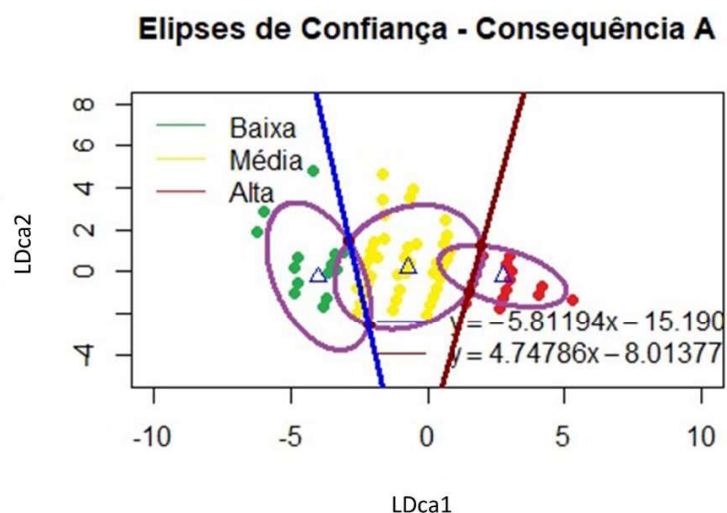


Figura 4.22– Elipses de confiança e limites entre as classes de consequência A.



Figura 4.23– Gráfico para Consequência A.

Por fim, para determinar a chance de um bloco atingir a rodovia em caso de queda de um talude qualquer, basta que o usuário dessa metodologia atribua notas de 1 a 4 para as variáveis CA1 a CA4 de acordo com as características mencionadas na Tabela 4.3. Essas notas devem ser padronizadas através da Equação 4.21, utilizando os dados da Tabela 4.23, obtendo as variáveis padronizadas ZA1 a ZA4. Os valores ZA1 a ZA4 são inseridos nas equações 4.17 a 4.20 a fim de se obter os escores das componentes principais ( $Y_{cai}$ ). Esses por sua vez, devem ser inseridos nas equações geradas pela Análise Discriminante (4.22 e 4.23) para obter os valores de LDca1 e LDca2. Esses valores são, por fim, plotados no gráfico de consequência A

(Figura 4.23), obtendo a classificação. Na Tabela 4.27 está um exemplo de aplicação com um talude do banco de dados. No Apêndice A consta o manual de instruções do sistema. No Apêndice B os cálculos dos exemplos mostrados nesse trabalho são detalhados.

Tabela 4.27 – Exemplo de classificação quanto à consequência A para um talude qualquer.

Variável (CAi)	CA1	CA2	CA3	CA4
Nota atribuída	4	2	3	4
Nota padronizada	0,980	-0,336	0,258	0,499
Componente Principal (Ycai)	Yca1	Yca2	Yca3	Yca4
Valor de Ycai	0,998	-0,161	-0,109	0,595
Valor de LDca1				1,876
Valor de LDca2				0,320
Classificação	Alta chance de atingir a rodovia (classe alta)			

#### 4.5. MODELO PARA CONSEQUÊNCIA B

Para avaliar a probabilidade (*Likelihood*) de ocorrência de quedas de blocos, 8 parâmetros relacionados às condições do maciço rochoso devem ser observados e medidos em campo. Para avaliar a chance dos blocos atingirem a rodovia (consequência A), 4 variáveis relacionadas com a geometria do talude e as dimensões da área de captação também são observadas em campo.

Todos esses 12 parâmetros são medidos ou observados diretamente e nenhum deles depende de situações que tornam essas observações/medições ainda mais subjetivas, como o comportamento humano. Portanto, foi possível realizar uma análise estatística a fim de verificar a influência que esses parâmetros exercem, juntos, no movimento de queda de blocos e, por fim, obter uma regra de classificação com equações que atribuem adequadamente os pesos às variáveis de acordo com o impacto em relação ao evento queda de blocos. Se tratam de casos multivariados.

Para avaliar a chance dos blocos causarem acidentes (consequência B), apenas duas variáveis são avaliadas, o Percentual de Distância de Visibilidade (CB1) e o Risco Médio do Veículo,

RMV, (CB2), variáveis essas detalhadas no subitem 2.4.1. A consequência B é um caso bivariado.

CB1 e CB2 medem o quanto os motoristas estão expostos a acidentes devido à queda de blocos de rocha por meio de parâmetros que podem ser medidos diretamente, como fluxo médio diário, extensão do talude, distância de visibilidade e limite de velocidade. Porém, com exceção da extensão do talude, a medição desses parâmetros pode ser muito mais complexa e subjetiva quando comparadas com os 12 parâmetros relacionados com a probabilidade e consequência A, pois são fortemente influenciados pelo comportamento do motorista, entre outras situações que são difíceis de abordar em simples medições.

O limite de velocidade da rodovia nem sempre é respeitado pelos motoristas e isso pode alterar a condição de visibilidade e, conseqüentemente, o tempo para o motorista perceber a existência de bloqueios na via e agir. Além disso, o tráfego médio diário pode variar consideravelmente em determinadas épocas do ano, como nos feriados. Sendo assim, os valores atribuídos a CB1 e CB2 podem não representar exatamente a real condição do local. Mas ainda são as variáveis que podem representar melhor o nível de exposição dos motoristas e dar uma dimensão das conseqüências e severidade de um possível acidente, visto que contabilizar o número exato de acidentes devido às quedas de blocos é algo que tem se mostrado inviável e não praticado.

Por toda essa subjetividade, seria difícil dizer, mesmo após aplicações de técnicas de estatística multivariada, qual desses dois parâmetros, CB1 ou CB2, tem o maior impacto na chance de acidentes por queda de blocos.

Além do mais, foi comprovado através do teste de esfericidade de Bartlett (Tabela 4.5) que o banco de dados reservado para gerar o modelo de consequência B não é adequado para a aplicação de técnicas de estatística multivariada. Aliado a isso, cabe salientar que tanto CB1 quanto CB2 são variáveis que dependem de outras variáveis (distância de visibilidade, extensão do talude, fluxo médio diário, entre outras) e que isso, por si só, já contra indica a aplicação das técnicas de estatística multivariada nessas variáveis.

De todo modo, uma vez que apenas duas variáveis pontuando entre 1 e 4 em escala ordinal são avaliadas para a Consequência B, não há muitas combinações possíveis para obter um determinado somatório final, que estará entre 2 e 8, considerando apenas números inteiros. Portanto, não é necessário reproduzir os mesmos passos adotados para gerar os modelos de probabilidade e consequência A, onde mais variáveis são consideradas. Para o caso da consequência B, esse processo pode se dar de forma mais intuitiva.

Sendo assim, a proposta para a consequência B é diferente do que foi adotado para gerar o modelo de probabilidade e de consequência A. Para a consequência B, as variáveis CB1 e CB2 terão o mesmo peso e a regra de classificação gerada é baseada na análise direta dos parâmetros. Para isso, primeiramente foi verificada e confirmada a variabilidade do banco de dados através de gráfico de frequência (Figura 4.24), onde se pode ver que para as duas variáveis, há indivíduos pontuando de 1 a 4.

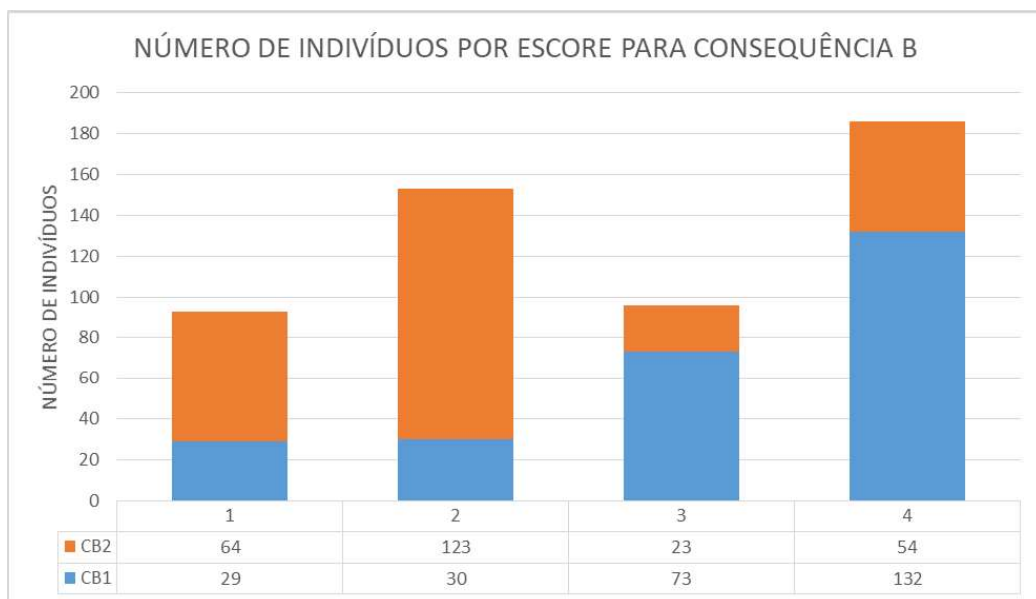


Figura 4.24 - Variabilidade das variáveis relacionadas à consequência B.

Como para se definir a classe para a consequência B são considerados 2 parâmetros, que podem receber pontuações entre 1 e 4, a soma mínima possível é 2 e a máxima é 8. Para um talude obter nota final igual a 2, é necessário que ambas as variáveis pontuem 1, o que é, indiscutivelmente, uma situação favorável à segurança em relação às chances de acontecer algum acidente devido a queda de blocos. O mesmo pode ser dito para taludes que possuem nota final igual a 3, onde a única situação possível é uma variável receber 1 ponto e a outra 2.

Com isso, pode-se dizer que para um talude que possui somatório final igual a 2 ou 3, as chances de acontecer algum acidente são baixas.

Em contrapartida, para que um talude obtenha nota final igual a 8, é necessário que as duas variáveis pontuem 4, o que é claramente uma situação desfavorável a segurança. Situação que não é muito diferente para taludes que possuem escore final igual a 7, visto que essa situação só ocorre se uma variável pontuar 3 e a outra 4. Com isso, taludes que possuem somatório final de escores igual a 7 e 8 devem ser considerados como de altas chances de causar acidentes.

Observando os taludes que somam 4, pode ocorrer de uma variável pontuar 3 e outra 1, que pode ser uma situação incerta quanto à segurança, visto que uma variável recebe uma nota alta e a outra, uma nota baixa. Ou pode ocorrer de ambas pontuarem 2, o que pode ser tomado como uma situação ainda favorável a segurança, porém no limite se for levado em consideração que notas 1 e 2 são baixas (favorável à segurança) e notas 3 e 4 são altas (desfavorável). Isso indica que o somatório igual a 4 é adequado como um limite inferior para uma classe média (zona de incerteza ou transição) quanto às chances dos blocos causarem acidentes.

Para os taludes com somatório igual a 6 podem acontecer também duas situações: as duas variáveis pontuarem 3, que é uma situação mais crítica; ou uma variável pontuar 4 e a outra 2, que é uma situação intermediária. Sendo assim, pode-se considerar que o somatório igual a 6 é adequado como um limite superior para a classe média (zona de incerteza) para a consequência B.

Com isso, a classificação quanto à consequência B é dada conforme a Tabela 4.28, ou mesmo quanto à matriz de Consequência B, mostrada na Figura 4.25, que representam de forma resumida e gráfica a discussão feita logo acima.

Sendo assim, para conhecer a classificação quanto às chances dos blocos causarem acidentes para um talude qualquer, o usuário da metodologia proposta nesse trabalho deve pontuar as variáveis características CB1 e CB2, de acordo com a Tabela 4.4 e somar essas pontuações. A classificação é dada conforme a Tabela 4.28 ou, se preferir, o pesquisador pode plotar as

pontuações de CB1 e CB2 diretamente na matriz de consequência B apresentada na Figura 4.23.

Tabela 4.28– Classes para consequência B obtidas após análise dos parâmetros.

Grupo	Faixa de somatório	Classe adotada para Consequência B
1	7 - 8	Alta chance
2	4 - 6	Média (Zona de incerteza)
3	2 - 3	Baixa Chance

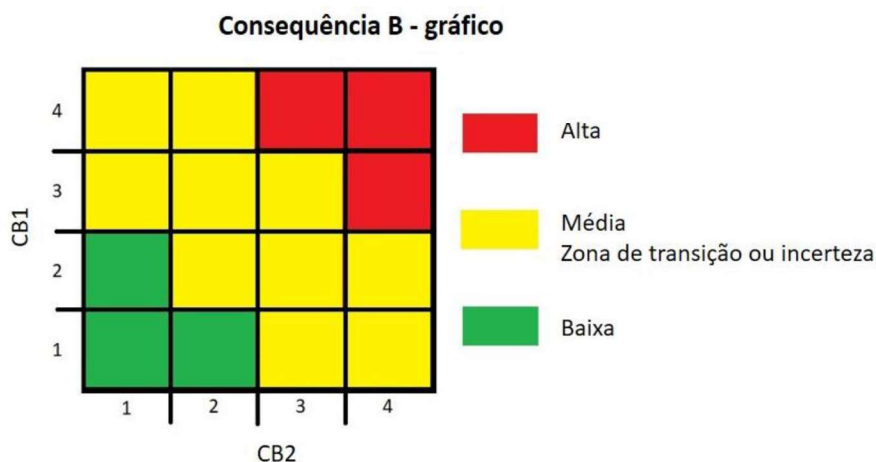


Figura 4.25 – Matriz de classificação para consequência B.

#### 4.6. SISTEMA DE ANÁLISE DE RISCO POR MATRIZ DE RISCO

Conforme explicado anteriormente, para obter a probabilidade (*Likelihood*) de ocorrência de queda de blocos, a chance desses blocos atingirem a rodovia (consequência A) e causar acidentes (consequência B), primeiramente o pesquisador deve atribuir valores de 1 a 4 para as variáveis, de acordo com suas características apresentadas nas Tabelas 4.1, 4.3 e 4.4, respectivamente.

Para determinar a probabilidade e a consequência A, esses valores são padronizados e depois inseridos nas equações das componentes principais, a fim de se obter os valores quantificados, denominados escores dos componentes principais ( $Y_i$  para probabilidade e  $Y_{cai}$  para consequência A). Por fim, esses escores são inseridos nas equações geradas pela Análise Discriminante e os resultados são plotados nos respectivos gráficos de classificação. Para a consequência B, basta que os valores atribuídos para as respectivas variáveis sejam plotados na matriz de classificação para a consequência B (Figura 4.25).



Sabendo que o risco é definido matematicamente como uma medida da probabilidade de um evento ocorrer multiplicado por uma medida de suas consequências, para obter a classificação de risco de queda de blocos de um talude qualquer, basta cruzar o resultado da classificação de probabilidade (*Likelihood*), obtido através do gráfico mostrado na Figura 4.13, com a consequência global dessas quedas de blocos. Essa consequência global é obtida cruzando os resultados das classificações de consequências A (Figura 4.23) e consequência B (Figura 4.25), pela matriz de consequências (Figura 4.26). A matriz de risco é, finalmente, apresentada na Figura 4.27.

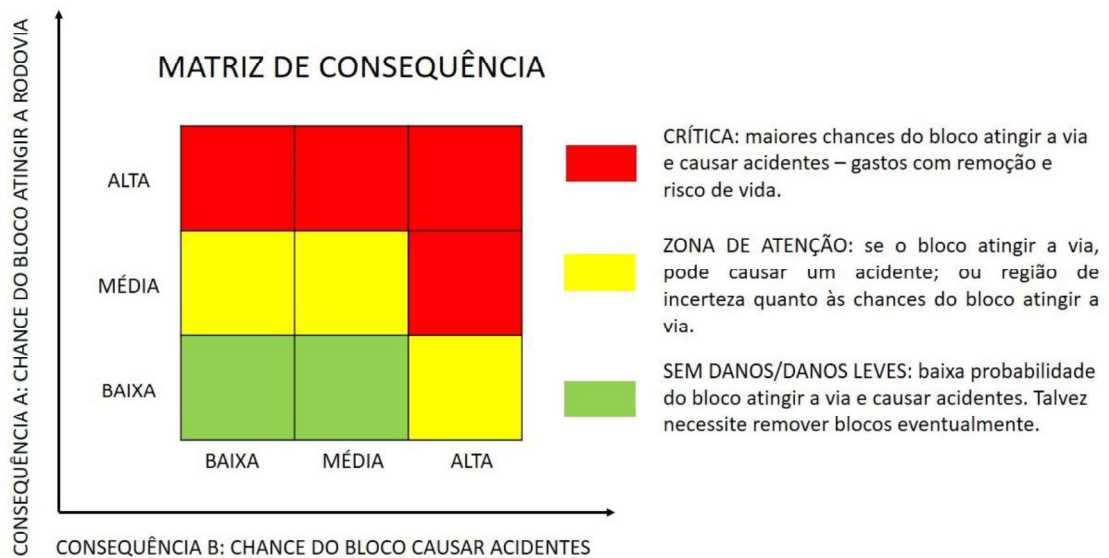


Figura 4.26 – Matriz de Consequências.

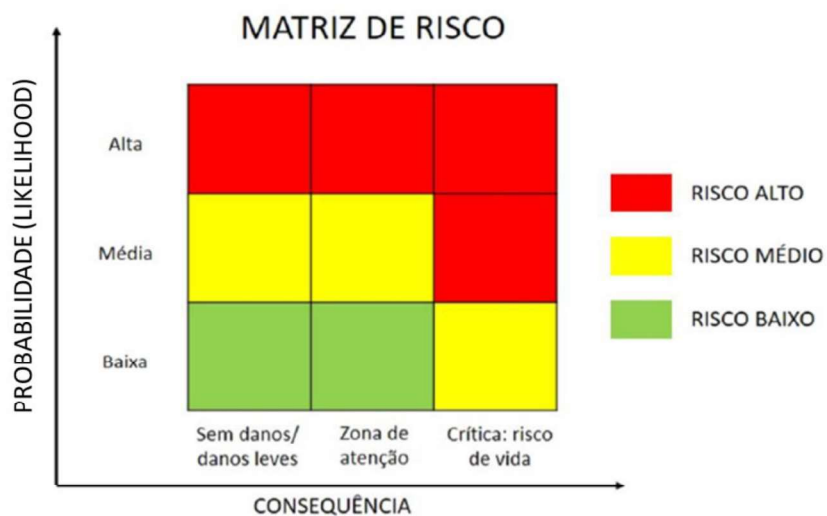


Figura 4.27 – Matriz de Risco.

Devido à subjetividade e às incertezas relacionadas à chance dos blocos causarem acidentes (consequência B), se as chances dos blocos atingirem a rodovia (consequência A) são altas, é razoável considerar que a consequência global é crítica (alta). Se os blocos chegarem à rodovia, mesmo que as condições de tráfego sejam favoráveis à segurança, não se pode afirmar com certeza que não haverá acidentes. Por outro lado, se é pouco provável que os blocos cheguem à rodovia, as chances de um acidente só aumentam se as condições de tráfego forem muito desfavoráveis. Nesta situação, a rodovia pode ser inserida em uma zona de atenção, conforme é mostrado na Figura 4.26.

Se a probabilidade à ocorrência de queda de blocos em um talude for baixa, seria conservador considerar que o risco de queda de blocos é alto mesmo se o nível de dano for alto, uma vez que é mais provável que o evento não aconteça. Porém, esta metodologia proposta, mesmo com robustez estatística, não acerta 100% das vezes, principalmente em eventos controlados por aspectos naturais. Portanto, considerar que o risco é baixo seria admitir que o talude não precisa de nenhum monitoramento, o que não é razoável quando se tem um alto nível de dano em caso de ocorrência do evento. Nesta situação, o risco é médio (Figura 4.27). Por outro lado, devido às incertezas inerentes às classificações das consequências, se a probabilidade for alta, o risco de ocorrência de quedas de blocos também deve ser alto, afinal, só existem consequências quando o evento ocorre.

Tomando como exemplo para aplicação do método o mesmo talude utilizado nos subitens anteriores para o cálculo dos escores das funções discriminantes (LD1 e LD2; LDca1 e LDca2) e classificação quanto à probabilidade e quanto a consequência A. Esse talude foi classificado como de probabilidade (*Likelihood*) alta para ocorrência de quedas de blocos (Tabela 4.17). Além do mais, apresentou chance alta de que esses blocos atinjam a rodovia (consequência A) (Tabela 4.27). Se esse talude pontua, para a Consequência B, 3 na variável CB1 e 4 na variável CB2, a chance de ocorrência de acidentes é alta, caso o bloco atinja a rodovia (plotar na Figura 4.23).

Pela Matriz de Consequências (Figura 4.26), tem-se que esse talude se enquadra como crítico, visto que para ambas as consequências (A e B) a classificação é alta. Sabendo que a

probabilidade à queda de blocos é alta e que as consequências são críticas, pela Matriz de Risco (Figura 4.27) esse talude é considerado como de alto risco.

No Apêndice A consta um manual de uso do sistema proposto, onde as tabelas, equações e gráficos são colocados novamente, a fim de sumarizar a metodologia para o usuário, da forma que ela deve ser usada. No Apêndice B consta o cálculo do exemplo de aplicação evidenciado no Capítulo 4 dessa tese, mostrando o passo a passo que o usuário do sistema deve reproduzir.

# **CAPÍTULO 5**

---

## **CONCLUSÕES**

O sistema proposto neste trabalho é uma ferramenta poderosa para prever o risco de queda de blocos em taludes localizados em rodovias, pois quantifica e classifica a probabilidade (*Likelihood*) de ocorrência dessas quedas e suas consequências, atribuindo pesos adequados às variáveis de acordo com sua importância no processo, conforme mostrado ao longo do desenvolvimento do trabalho. Além do mais, a taxa de erro do sistema proposto é evidenciada e detalhada. Ou seja, o usuário desse sistema sabe o quanto ele pode errar e como ele pode errar. Isso é fundamental para diminuir a subjetividade inerente a esse tipo de análise geotécnica que envolve observações de campo e, muitas vezes, carrega as percepções do engenheiro ou geólogo.

A robustez estatística deste trabalho é a responsável pela ponderação adequada para as variáveis, de acordo com sua importância no processo e isso pode explicar as baixas taxas de erro ao testar este sistema em taludes cuja classificação quanto à probabilidade, chance dos blocos atingirem a rodovia e chance deles causarem acidentes já era conhecida. Além disso, este sistema considera todas as variáveis mais importantes relacionadas à queda de blocos em rodovias e pode ser utilizado em qualquer região, pois não são consideradas variáveis tipicamente regionais e sazonais, como ciclos de gelo e degelo.

Todas as variáveis consideradas no método apresentado são fáceis de observar e avaliar em campo. Com isso, basta o pesquisador ir a campo atribuir notas de 1 a 4 para cada variável de acordo com suas características. Com essas pontuações, é necessário utilizar as equações, gráficos e matrizes apresentadas para obter o risco de queda. Nos Apêndices A e B constam o manual de uso e um exemplo detalhado de cálculo, respectivamente. O usuário deste sistema proposto não precisa fazer novas análises estatísticas utilizando programação de computador. Isso já foi feito para gerar o sistema de análise de risco. O usuário pode, inclusive, já deixar essas equações propostas nesse trabalho implementadas em uma planilha eletrônica, como Excel.

No entanto, deve-se destacar que esse sistema de análise de risco proposto nesse trabalho se trata de uma proposta inicial e, portanto, é encorajado o aperfeiçoamento da metodologia como forma de dar continuidade à mesma. Isso pode ser feito através da inserção de novos taludes ao banco de dados e também através de testes com outras técnicas de estatística multivariada e aprendizado de máquina.

De todo modo, pode-se afirmar que este sistema de análise de risco de queda proposto é simples e prático, o que contribui para que o trabalho do engenheiro seja viável tanto em termos de tempo de conclusão quanto financeiramente, o que é um pré-requisito para análises em obras viárias.

É importante destacar que essa proposta não tem o objetivo de substituir análises de estabilidade de taludes pelos métodos de equilíbrio-limite ou por métodos numéricos. São abordagens diferentes e complementares, pois se aplicam a momentos e condições específicas durante uma análise de estabilidade de taludes sujeito a queda de blocos. Por exemplo, se o objetivo é avaliar o mecanismo de tensão e deformação em um talude, obviamente essa metodologia de análise de risco não é indicada, mas sim uma análise utilizando métodos numéricos. No entanto, se a finalidade é avaliar em campo a condição de risco de vários taludes, com um tempo restrito e recursos escassos, essa ferramenta é adequada, dada a facilidade de uso, o fato de ser gratuita e não exigir ensaios de laboratório.

Por fim, a metodologia proposta nesse trabalho é focada na análise de risco para taludes rodoviários. No entanto, o modelo de probabilidade (*Likelihood*) pode ser utilizado para qualquer talude em rocha cristalina (ígneas e metamórficas), seja ele em rodovias, urbano ou de mineração. O modelo de consequência A também pode ser adaptado para outras situações, no entanto deve-se pensar em uma alternativa para tratar o parâmetro “área de captação” e deixá-lo de acordo com a situação em análise. Já o modelo de Consequência B proposto nesse trabalho é apropriado para uso apenas em rodovias. De modo geral, todo o trabalho pode ser adaptado ou servir de base para outros modelos de análise de risco geotécnico para outras situações, assim como outros trabalhos foram referências importantes para o desenvolvimento deste. Mas para isso é necessário que seja utilizado um banco de dados adequado para o contexto e para aplicação das técnicas de estatística multivariada.

## **REFERÊNCIAS**

---

Alboukadel Kassambara and Fabian Mundt (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. <https://CRAN.R-project.org/package=factoextra>

Anderson, S.A.; King, T.J.; Schloner M.; 1999. Urban Rockfall Hazard And Mitigation Cost Rating, In: Amandei, B.; Kranz, R.L.; Scott, G.A.; Smeallie, P.H.; Proceedings Of The 37th U.S. Rock Mechanics Symposium, Volume 2, Rock Mechanics For Industry, Pp. 1229-1234.

Andrew, R.D., 1994. The Colorado Rockfall Hazard Rating System, Colorado Dept. Of Transportation, Report # CTI-CDOT-2-94.

Arnold, J.G.; Allen, P.M.; Ramanarayanan, T.S.; Srinivasan, R.; Muttiah, R.S.; 1996. The Geographical Distribution Of Freeze/Thaw And Wet/Dry Cycles In The United States: Environmental & Engineering Geoscience, Vol. II, No. 4, Pp. 596-603.

Asteriou, P. [et al.] (2012) - Geotechnical and kinematic parameters affecting the coefficients of restitution for rock fall analysis. International Journal of Rock Mechanics and Mining Sciences [Em linha]. 54:0. 103-113. Disponível em WWW: <URL: <http://www.sciencedirect.com/science/article/pii/S1365160912001098>>. ISSN 1365-1609.

Barrett, R.K. & White, J.L. 1991. Rockfall Prediction And Control: National Symposium On Highway And Railroad Slope Maintenance, 34th Annual Meeting Of The Association Of Engineering Geologists, Pp. 23-40.

Bartlett, M. S. 1951. The Effect Of Standardization On A Chi Square Approximation In Factor Analysis. Biometrika; 38:337-44.

Bar, N.; Barton, N. 2017. The Q-Slope Method For Rock Slope Engineering. 16p

Barton, N., Lien, R.; Lunde, J. 1974. Engineering Classification Of Rock Masses For The Design Of Rock Support. *Rock Mechanics*, V. 6, P. 189-236.

Bateman, V., 2003. The Development Of A Database To Manage Rockfall Hazard: The Tennessee Rockfall Hazard Database: 2003 Transportation Research Board Annual Meeting, 13p.

Bauer, M. & Neumann, P. 2011. A Guide To Processing Rock-Fall Hazard From Field Data.- In: Vogt, N., Schuppener, B., Straub, D. & Bräu, G., *Geotechnical Safety And Risk, Proceedings Of The 3rd International Symposium On Geotechnical Safety And Risk*, June 2011: 149-156

Bezdek, J.C & Hathaway, R. J. 2002. “VAT: A Tool for Visual Assessment of (Cluster) Tendency,” *Proceedings of the 2002 International Joint Conference on Neural Networks*, Honolulu, 12-17 May 2002, pp. 2225-2230.

Bieniawski, Z.T. 1989 *Engineering Rock Mass Classifications: A Complete Manual For Engineers And Geologists In Mining, Civil, And Petroleum Engineering*. New York : John Wiley & Sons, 251 P.

Box G. 1949. A General Distribution Theory For A Class Of Likelihood Criteria. *Biometrika*; 36:317–46.

Breiman, I.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and regression trees*. Wadsworth International: California, USA, 1984.

Brown, E. T. 1981. *Rock Characterization Testing & Monitoring-Isrm Suggested Methods*. Pergamon Press, Oxford, 211p.

Bunchoft, G; Kellner, S R.O. 1999 *Estatística Sem Mistérios*. 2.Ed. Petrópolis: Vozes. V.2, 303p.

Callegari-Jacques, S. M. 2003. *Bioestatística: Princípios E Aplicações*. Porto Alegre: Artemed. 255p.

Campos, S. L. S. 2019. Busca não supervisionada de padrões por técnicas de agrupamento clássica e nebulosa. Trabalho de Conclusão de Curso (Graduação). Universidade Federal de Juiz de Fora, 149 P

Carvalho, P. A. S. 1991. Taludes De Rodovias: Orientação Para Diagnóstico E Soluções De Seus Problemas. São Paulo. Ipt. 410 P.

Cascini, L.; Bonnard, C.; Corominas, J.; Jibson, R.; Montero-Olarte, J. 2005. Landslide Hazard And Risk Zoning For Urban Planning And Development. In: Hungr, O., Fell, R., Couture, R., Eberhardt, E. (Eds.). Landslide Risk Management. Londres: Taylor And Francis. P. 199–235.

Cattell, R. B. 1966. The Screen Test For The Number Of Factors. Multivariate Behavioral Research, 1, P. 140-161.

Charrad M.; Ghazzali N.; Boiteau V.; Niknafs A. 2015. Nbclust: Determining The Best Number Of Clusters In A Data Set. R Package, Version 3.0. In: <https://CRAN.Rproject.org/package=Nbclust>.

Data Novia. Cluster Analysis In R: Practical Guide. <<https://www.datanovia.com/en/blog/cluster-analysis-in-r-practical-guide>>. Acesso Em 23/02/2021.

Efron, B. 1992. Bootstrap Methods: Another Look At The Jackknife. In Breakthroughs In Statistics. New York: Springer. P. 569-593.

Elliasen, T.D. & Ingraham, P.C., 2000. Review, Rockfall Hazard Rating And Repair Of Aging Interstate Rock Cuts In Northeastern Vermont: 51st Annual Highway Geology Symposium Proceedings, Seattle Washington, Pp. 233-243.

Everitt, B. S.; Landau, S.; Leese, M.; Stahl, D. 2011. Cluster Analysis. 5<sup>a</sup> Ed. Chichester: Wiley. 330 P. ISBN: 978-0-470-74991-3.



Fell R., Corominas J., Bonnard C., Cascini L., Leroi E., Savage W.Z. 2008. Guidelines For Landslide Susceptibility, Hazard And Risk Zoning For Land Use Planning. *Engineering Geology* 102 (2008) 85–98. [www.elsevier.com/locate/enggeo](http://www.elsevier.com/locate/enggeo).

Ferreira, D. F. 2018. *Estatística Multivariada*. 3ª Ed. Lavras: Editora UFLA. 624 P. ISBN: 978-85-8127-063- 0.

Fish, M., Lane, R. 2001. GIS And The New Hampshire Rock Cut Inventory: 81<sup>st</sup> Transportation Research Board Annual Meeting, National Research Council, Washington D.C.

Fisher R. 1936. The Use Of Multiple Measurements In Taxonomic Problems. *Ann Eugen*;7:179–88.

Giani, G. P. 1992. *Rock Slope Stability Analysis*. A.A. Balkema Publishers, Rotterdam, Netherlands.

Gower, J. C.; Legendre, P. 1986. Metric And Euclidean Properties Of Dissimilarity Coefficients. *Journal Of Classification*, V.3, N.1, P.5–48. <https://doi.org/10.1007/bf01896809>.

Hair, J.F.; Black, W. C.; Babin, B. J.; Anderson, R. E.; Tatham, R. L. 2009. *Análise Multivariada De Dados*. Porto Alegre: Bookman, 688 P.

Heidenreich, B. (2004) - Small-and half-scale experimental studies of rockfall impacts on sandy slopes. Citeseer.

Hoek, E. 2000. *Practical Rock Engineering*. Edition by Dr. E. Hoek, 328p.

Hoek, E. 2007. *Practical Rock Engineering, Analysis of Rockfall Hazards*. Rocscience, University Of Toronto, 341p. . (Available from the publisher at <http://www.rocscience.com/hoek/PracticalRockEngineering.asp>).

Hopkins, b.; Skellam, j. G. A new method for determining the type of distribution of plant individuals. *Annals of Botany*, v.18, n.2, p. 213 - 227, 1954.

Jain, A. K. 2010 Data Clustering: 50 Years Beyond K-Means. *Pattern Recognition Letters*, N. 31, P. 651-666. <https://doi.org/10.1016/j.patrec.2009.09.011>.

Hotelling H. 1933. Analysis of A Complex Of Statistical Variables Into Principal Components. *J Educ Psychol*;24:417–41.

Kaiser, H. F. A. 1970. Second Generation Little Jiffy. *Psychometrika*, V. 35, N. 4, P. 401-415.

Kassambara, A. 2017 *Practical Guide to Cluster Analysis In R Unsupervised Machine Learning*. London: STHDA.

Kaufman, L.; Rousseeuw P. J. 1990. *Finding Groups In Data: An Introduction To Cluster Analysis*. Wiley Series In Probability And Statistics. ISBN: 0-47 1-73578-7. <https://doi.org/10.1002/9780470316801>.

Lawson, Richard G.; Peter C. Jurs. 1990. “New Index For Clustering Tendency And Its Application To Chemical Problems.” *Journal Of Chemical Information And Computer Sciences* 30 (1): 36–41. <http://pubs.acs.org/doi/abs/10.1021/ci00065a010>.

Lienhart, D.A., 1988. The Geographic Distribution Of Intensity And Frequency Of Freezethaw Cycles: *Bulletin Of The Association Of Engineering Geologists*, Vol. XXV, No. 5, Pp. 465-469.

LIRA, S. A. 2004. *Análise De Correlação: Abordagem Teórica E De Construção Dos Coeficientes Com Aplicações*. Dissertação De Mestrado. Universidade Federal Do Paraná, 196 P.

Maerz, N.H.; Youssef A., Fennessey, T.W. 2005. New Risk-Consequence Rockfall Hazard Rating System For Missouri Highways Using Digital Image Analysis. *Environmental And Engineering Geoscience* 11, 229–249 2005.

McMillan, P.; Matheson, G. D. 1998. Rock Slope Hazard Assessment: A New Approach. London: Engineering Geology Special Publications N. 15, P. 177-183.

MICROSOFT. 2013. Microsoft® Excel® 2013. <https://www.microsoft.com/pt-br>.

Miller, S.M., 2003. Development And Implementation Of The Idaho Highway Slope Instability And Management System (Hisims). Idaho Transportation Department, Report # N03-07.

Mingoti, S. A. 2013. Análise De Dados Através De Métodos De Estatística Multivariada: Uma Abordagem Aplicada. Belo Horizonte: Editora UFMG, 295 P.

Nazareth, A. F. D. V. 2020. Análise De Agrupamentos Aplicada À Setorização Geotécnica De Minas A Céu Aberto. Dissertação (Mestrado). Programa De Pós-Graduação Em Engenharia Mineral, Universidade Federal De Ouro Preto, Ouro Preto-MG, 170 P.

New York State Dept. Of Transportation, 1996. Rock Slope Rating Procedure. Geotechnical Engineering Bureau, Report # GEM-15.

Pack, R.T.; Boie K., 2002. Utah Rockfall Hazards Inventory, Phase I: Research Division Of Utah Department Of Transportation, Report # UT-03.01.

PAPAIVANOU, A; SUORINENI, F.T. Development of a generalised dilution – based stability graph for open stope design. Jornal of Research Projects Review, Austrália, 2015, v.4, 36 – 43p.

Pierson, L. A.; Davis, S. A. Van Vickle, R. 1990. Rockfall Hazard Rating System Implementation Manual Federal Highway Administration (FHWA), Report FHWA-OR- EG-90-01, FHWA, U.S. Dep. Of Transp.

Pierson, L. A., Van Vickle R. 1993. Rockfall Hazard Rating System. Transportation Research Record N° 1343, National Research Board, Washington, D.C., Pp 6-19.

Pierson, L. A., Gullixon, G. F., Chassie, R. G. 2001. Rockfall Catchment Area Design Guide Final Reportspr-3 (032). Technical Report Form DOT F 1700.7. Oregon, U. S.

Pereira, T, M. 2015. Estatística Multivariada I. Uma Abordagem Aplicada Utilizando O Software R. Apostila. Departamento De Estatística. Universidade Federal De Ouro Preto (UFOP). Ouro Preto. 254p.

PEREIRA,T.M. Algoritmo Bootstrap aplicado às distâncias de Mahalanobis para obtenção de elipses de confiança para variáveis não normais. Ouro Preto, 2016.

R Core Team (2020). R: A Language And Environment For Statistical Computing. R Foundation For Statistical Computing, Vienna, Austria. URL <https://www.r-project.org/>.

Ritchie, A. M. 1963 Evaluation Of Rockfall And Its Control. U. S. Department Of Commerce, Bureau Of Public Roads, And The Washington State Highway Commission.

Rocscience Inc. (2001). Rocfall Version 6.0 - Statistical Analysis Of Rockfalls. Www.Rocscience.Com, Toronto, Ontario, Canada.

Romana, M. 1985. New Adjustment Ratings For Application Of Bieniawski Classification To Slopes. Int. Symp. On The Role Of Rock Mechanics ISRM. Zacatecas, P. 49-53.

Romana, M., 1993. SMR classification: Romana, M Proc 7th ISRM International Congress on Rock Mechanics, Aachen, 16-20 September 1991V2, P955-960. Publ Rotterdam: A A Balkema, 1991. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts 30, A231-A231.

Royston, J. B. 1983. Some Techniques For Assessing Multivariate Based On The Shapiro-Wilk W. Applied Statistics, London, V. 32, N. 2, P. 121-133.

Russel, C. P.; Santi, P; Higgins, J. D. 2008. Modification and Statistical Analysis of The Colorado Rockfall Hazard Rating System. Colorado School of Mines. Golden – Colorado. 137p

Santi, P. M.; Russel, C. P.; Higgins, J. D.; Spriet, J. I. 2009. Modification and Statistical Analysis of The Colorado Rockfall Hazard Rating System. Engineering Geology 104 (2009) 55–65.

Santos, T. B. 2019. Análise De Risco Geotécnico Em Taludes Rochosos De Mina Com Uso De Técnicas Estatísticas Multivariadas E De Aprendizado De Máquina. Tese (Doutorado). Programa De Pós-Graduação Em Engenharia Mineral, Universidade Federal De Ouro Preto, Ouro Preto- MG, 106 P.

Senior, S.A., 1999. Rockfall Hazard Remediation Along Ontario Highways: 50th Annual Highway Geology Symposium, Proceedings And Field Trip Guide, Pp. 276-286.

Shakoor, A., 2005. Development Of A Rockfall Hazard Rating Matrix For The State Of Ohio: The Ohio Department Of Transportation Office Of Research & Development Executive Summary Report.

Silva, L. C. 2016. Aplicação Do Sistema De Avaliação De Perigo “Rockfall Hazard Rating System” (RHRS) Aos Taludes Da BR-356 / Rodovia Dos Inconfidentes, MG. Dissertação (Mestrado). Programa De Pós-Graduação Em Geotecnia, Universidade Federal De Ouro Preto, Ouro Preto- MG, 140 P.

Silveira, L. R C. 2017. Avaliação Do Perigo De Queda De Blocos Em Taludes Urbanos E Ferroviários E Simulação De Sua Trajetória. Dissertação (Mestrado). Programa De Pós-Graduação Em Engenharia Mineral, Universidade Federal De Ouro Preto, Ouro Preto- MG, 161 P.

Standards Australia. AS/NZS 4360: Risk Management. Standards Australia, Sydney, 2004.

Stevens, W. D. RocFall, a tool for probabilistic analysis, design of remedial measures and prediction of rockfalls. 1998.

Stover, B.K., 1992. Highway Rockfall Research Report: Colorado Geological Survey, Special Publication 37, 27 P.

SUORINENI, F.T; KAISER, P.K; TANNANT, D.D. Likelihood statistic for interpretation of the stability graph for open stope design. International Journal of Rock Mechanics & Mining Sciences, 2001, p.735 – 744.

Szwilski, A.B., 2002. Rockfall Rating Evaluation And Data Management Systems For Highway And Railway Rock Slopes: Appalachian Transportation Institute And West Virginia Department Of Transportation.

Tavares, L. M. M. S. C. (2015). Análise Paramétrica No Estudo Da Queda De Blocos Em Encostas Rochosas. Mestrado Integrado em Engenharia Civil - 2014/2015 - Departamento de Engenharia Civil, Faculdade de Engenharia da Universidade do Porto, Porto, Portugal, 2015. 110 p.

Tinsley, H. E. A.; Steven, D. B. 2000 Handbook Of Applied Multivariate Statistics And Mathematical Modeling. 1ª Ed. San Diego: Academic Press, 2000. 721 P. ISBN: 0-12-691360-9.

Vandewater, C. J.; Dunne, W. M.; Mauldon, M.; Drumm, E. C.; Batemann, V.; 2005. Classifying And Assessing The Geologic Contribution To Rockfall Hazard: Environmental & Engineering Geoscience, Vol XI, No. 2, Pp. 141-154.

Vassalo, T. F. S. 2017. Metodologia não paramétrica para delimitação dos estados de estabilidade no gráfico de estabilidade. Dissertação (Mestrado). Programa De Pós-Graduação Em Engenharia Mineral, Universidade Federal De Ouro Preto, Ouro Preto-MG, 66 P.

Vela, P. A.; Kingravi, H. A.; Celebi, M. E. A. 2013 Comparative Study Of Efficient Initialization Methods For The K-Means Clustering Algorithm. Expert Systems With Application, N. 40, P. 200-210. <https://doi.org/10.1016/j.eswa.2012.07.021>.

Witteman, T. A.; Lamaskin, T. A.; Watts, C. F.; 1992. Examples Of The Application Of The Rockfall Hazard Rating System To Virginia Highways: Virginia Journal Of Science, Vol. 43, No. 2, P. 266.

Wyllie, D.C., 1987. Rock Slope Inventory: Proceedings, Federal Highway Administration Rockfall Mitigation Seminar, FHWA Region 10, Portland Oregon, 25 P.

## APÊNDICE

### **APÊNDICE A: MANUAL DE USO DO SISTEMA DE ANÁLISE DE RISCO DE QUEDA DE BLOCOS PARA TALUDES RODOVIÁRIOS**

Para determinar a classificação de risco de queda de blocos em um talude rodoviário em rocha cristalina, o primeiro passo é determinar a probabilidade de ocorrência do evento. Para isso, os parâmetros P1 a P8, relacionados com as características do maciço rochoso que compõe o talude devem ser pontuadas de acordo com a Tabela A.1. A pontuação atribuída pode ser 1, 2, 3 ou 4, sendo 1 a nota atribuída quando o parâmetro é muito favorável à estabilidade e 4 quando o parâmetro for muito adverso.

**Tabela A.1**– Parâmetros e suas devidas pontuações para análise de probabilidade.

Parâmetro	Pontuação			
	1	2	3	4
Presença de água (P1)	Seco	Úmido	Gotejando	Fluxo
Características do maciço (P2)	Homogêneo/ maciço	Pequenas falhas/veios fortes	Xistosidades e zonas de cisalhamento ( $< 15$ cm)	Pegmatitos/ micas/zonas de cisalhamento ( $>15$ cm)
Grau de descalçamento (P3)	0 a 0,3 m	0,3 a 0,6 m	0,6 a 1,2 m	$> 1,2$ m
Tamanho de bloco/volume (P4)	$<0,3$ m/ $<0,75$ m <sup>3</sup>	0,3 a 0,6m/ 0,75 a 2,3m <sup>3</sup>	0,6 a 1,5m/ 2,3 a 7,6m <sup>3</sup>	$> 1,5$ m/ $> 7,6$ m <sup>3</sup>
Persistência e orientação das descontinuidades (P5)	$< 3$ m mergulhando para dentro do talude	$> 3$ m cortando o talude	$< 3$ m mergulhando para fora do talude	$> 3$ m mergulhando para fora do talude
Abertura das descontinuidades (P6)	0	0,1 a 1 mm	1 a 5 mm	$> 5$ mm
Alteração/preenchimento das descontinuidades (P7)	Fresca	Descoloração superficial	Preench. granular	Preench. argiloso
Rugosidade das descontinuidades (P8)	Rugosa	Ondulada	Plana	Estriada

Após atribuir as notas de 1 a 4 para cada parâmetro, ou variável, P1 a P8 (Tabela A.1), essas notas devem ser padronizadas de acordo com a equação A.1, utilizando os valores das médias ( $\mu_i$ ) e desvios padrões ( $\sigma_i$ ) de cada variável  $P_i$  (Tabela A.2).



$$Z_i = \frac{P_i - \mu_i}{\sigma_i} \quad (\text{A.1})$$

Onde

$Z_i$ : valor padronizado dos escores de 1 a 4 de cada variável relacionada à probabilidade.

$P_i$ : escores de 1 a 4 (Tabela A.1) de cada variável,  $i = 1, 2, 3, \dots, 8$ .

$\mu_i$ : média da variável  $P_i$ ,  $i = 1, 2, 3, \dots, 8$  (Tabela A.2).

$\sigma_i$ : desvio padrão da variável  $P_i$ ,  $i = 1, 2, 3, \dots, 8$  (Tabela A.2).

**Tabela A.2**– Valores de média e desvio padrão para cada variável relacionada à probabilidade.

Variável	P1	P2	P3	P4	P5	P6	P7	P8
$\mu_i$	1,450	2,659	2,205	2,318	3,082	3,682	2,727	1,518
$\sigma_i$	0,684	0,775	0,979	0,886	0,971	0,531	0,603	0,679

Os valores padronizados  $Z_i$ , por sua vez, devem ser inseridos nas Equações A.2 a A.9, a fim de se obter os valores das componentes principais  $Y_i$ .

$$Y_1 = 0,405 Z_1 + 0,254 Z_2 + 0,546 Z_3 + 0,518 Z_4 + 0,334 Z_5 + 0,280 Z_6 + 0,078 Z_7 + 0,098 Z_8 \quad (\text{A.2})$$

$$Y_2 = 0,238 Z_1 + 0,468 Z_2 - 0,096 Z_3 - 0,337 Z_4 + 0,057 Z_5 - 0,278 Z_6 + 0,702 Z_7 + 0,165 Z_8 \quad (\text{A.3})$$

$$Y_3 = 0,071 Z_1 - 0,180 Z_2 + 0,121 Z_3 - 0,034 Z_4 + 0,242 Z_5 - 0,552 Z_6 - 0,306 Z_7 + 0,700 Z_8 \quad (\text{A.4})$$

$$Y_4 = 0,073 Z_1 - 0,211 Z_2 - 0,102 Z_3 + 0,056 Z_4 + 0,655 Z_5 - 0,416 Z_6 + 0,046 Z_7 - 0,574 Z_8 \quad (\text{A.5})$$

$$Y_5 = 0,713 Z_1 - 0,511 Z_2 + 0,088 Z_3 - 0,053 Z_4 - 0,422 Z_5 - 0,121 Z_6 + 0,102 Z_7 - 0,127 Z_8 \quad (\text{A.6})$$

$$Y_6 = 0,127 Z_1 + 0,575 Z_2 + 0,153 Z_3 + 0,047 Z_4 - 0,342 Z_5 - 0,445 Z_6 - 0,457 Z_7 - 0,321 Z_8 \quad (\text{A.7})$$

$$Y_7 = 0,494 Z_1 + 0,214 Z_2 - 0,596 Z_3 - 0,162 Z_4 + 0,255 Z_5 + 0,324 Z_6 - 0,385 Z_7 + 0,099 Z_8 \quad (\text{A.8})$$

$$Y_8 = 0,025 Z_1 - 0,016 Z_2 + 0,530 Z_3 - 0,763 Z_4 + 0,193 Z_5 + 0,220 Z_6 - 0,194 Z_7 - 0,113 Z_8 \quad (\text{A.9})$$

Por fim, os valores obtidos para  $Y_i$  (Equações A.2 a A.9) são inseridos nas equações discriminantes LD1 e LD2 (Equações A.10 e A.11) e o resultado plotado no gráfico de probabilidade de queda de blocos (Figura A.1), onde o talude pode ser classificado

como de alta probabilidade de queda, baixa probabilidade, ou média probabilidade (zona de transição ou de incerteza).

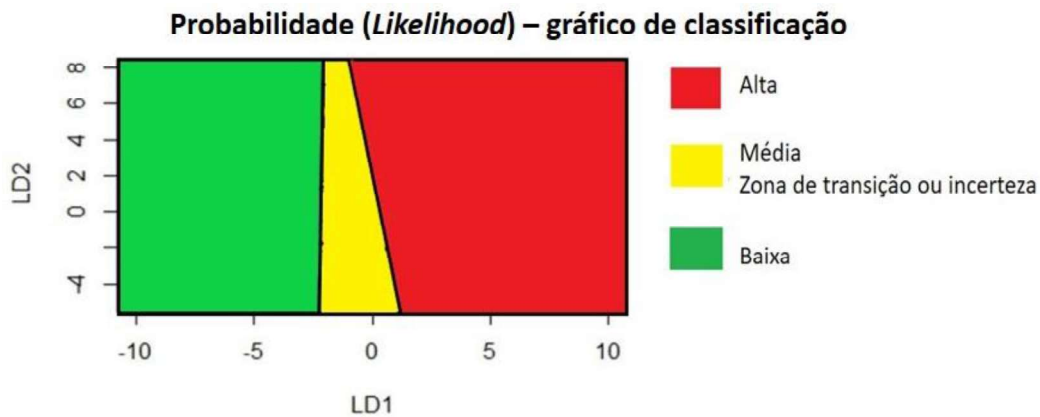
$$LD1 = 1,505 Y1 + 0,317 Y2 + 0,187 Y3 + 0,031 Y4 - 0,318 Y5 - 0,113 Y6 + 0,013 Y7 + 0,106 Y8 \quad (A.10)$$

$$LD2 = 0,029 Y1 + 0,136 Y2 - 0,078 Y3 - 0,325 Y4 + 0,657 Y5 - 0,820 Y6 + 0,258 Y7 + 0,313 Y8 \quad (A.11)$$

Onde:

Y<sub>i</sub>: escores das componentes principais i, sendo i = 1, 2, 3, ..., 8.

L<sub>d</sub>i: escores das funções discriminantes, sendo i = 1 e 2.



**Figura A.1** – Gráfico de classificação quanto à probabilidade de queda de blocos.

Obtida a classificação quanto à probabilidade, o próximo passo é determinar as chances do bloco (ou dos blocos) atingir a rodovia, sendo essa a Consequência A. Para isso, deve-se, avaliar os parâmetros altura do talude (CA1), inclinação do talude (CA2), feições de lançamento (CA3) e eficácia da área de captação (CA4). A essas variáveis são atribuídas as notas de 1 a 4 de acordo com as características observadas, conforme a Tabela A.3.

**Tabela A.3** – Parâmetros e suas devidas pontuações para análise da consequência A.

Parâmetro	Faixas			
	1	2	3	4
Altura do talude (m) (CA1)	7,5 a 15	15 a 23	23 a 30	> 30
Inclinação do talude (°) (CA2)	76 - 90	61 - 75	46 - 60	30 - 45
Feições de lançamento (CA3)	Nenhuma (talude suave)	Menores	Maiores	Extremas
Área de Captação (CA4)	95 a 100%	65 a 94%	30 a 64%	< 30%

Após atribuir as notas de 1 a 4 para cada variável, CA1 a CA4 (Tabela A.3), essas notas devem ser padronizadas de acordo com a equação A.12, utilizando os valores das médias ( $\mu_{cai}$ ) e desvios padrões ( $\sigma_{cai}$ ) de cada variável CA<sub>i</sub> (Tabela A.4).

$$ZA_i = \frac{CA_i - \mu_{cai}}{\sigma_{cai}} \quad (A.12)$$

Onde:

CA<sub>i</sub>: escores de 1 a 4 (Tabela A.3) de cada variável,  $i = 1, 2, \dots, 4$ .

$\mu_{cai}$ : média da variável CA<sub>i</sub>,  $i = 1, 2, \dots, 4$  (Tabela A.4).

$\sigma_{cai}$ : desvio padrão da variável CA<sub>i</sub>,  $i = 1, 2, \dots, 4$  (Tabela A.4).

**Tabela A.4** – Valores de média e desvio padrão para cada variável relacionada à consequência A.

Variável	CA1	CA2	CA3	CA4
$\mu_{cai}$	2,932	2,261	2,807	3,678
$\sigma_{cai}$	1,091	0,777	0,748	0,646

Os valores padronizados ZA<sub>i</sub>, por sua vez, devem ser inseridos nas Equações A.13 a A.16, a fim de se obter os valores das componentes principais Yca<sub>i</sub>.

$$Yca1 = 0,641 ZA1 - 0,121 ZA2 + 0,696 ZA3 + 0,299 ZA4 \quad (A.13)$$

$$Yca2 = 0,367 ZA1 + 0,740 ZA2 + 0,032 ZA3 - 0,562 ZA4 \quad (A.14)$$

$$Yca3 = 0,019 ZA1 - 0,603 ZA2 + 0,208 ZA3 - 0,770 ZA4 \quad (A.15)$$

$$Yca4 = 0,673 ZA1 - 0,272 ZA2 - 0,686 ZA3 + 0,044 ZA4 \quad (A.16)$$

Onde:

$Y_{cai}$ : componente principal  $i$ , sendo  $i = 1, 2, \dots, 4$ .

$Z_{Ai}$ : valor padronizado dos escores de 1 a 4 de cada variável relacionada à consequência A.

Por fim, os valores obtidos para  $Y_{cai}$  (Equações A.13 a A.16) são inseridos nas equações discriminantes  $LD_{ca1}$  e  $LD_{ca2}$  (Equações A.17 e A.18) e o resultado plotado no gráfico de classificação da Consequência A (Figura A.2), onde o talude pode ser classificado como de alta chance do bloco atingir a rodovia, baixa chance, ou média chance (zona de transição ou de incerteza).

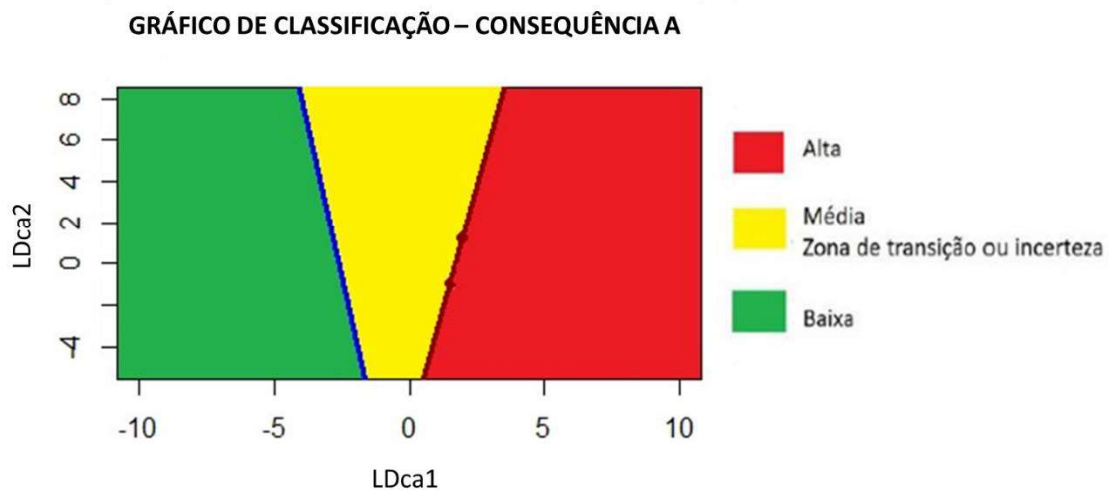
$$LD_{ca1} = 1,764 Y_{ca1} + 0,808 Y_{ca2} - 0,876 Y_{ca3} + 0,254 Y_{ca4} \quad (A.17)$$

$$LD_{ca2} = 0,079 Y_{ca1} + 0,348 Y_{ca2} + 1,028 Y_{ca3} + 0,688 Y_{ca4} \quad (A.18)$$

Onde:

$Y_{cai}$ : escores das componentes principais  $i$ , sendo  $i = 1, 2, \dots, 4$ .

$LD_{cai}$ : escores das funções discriminantes, sendo  $i = 1$  e  $2$ .



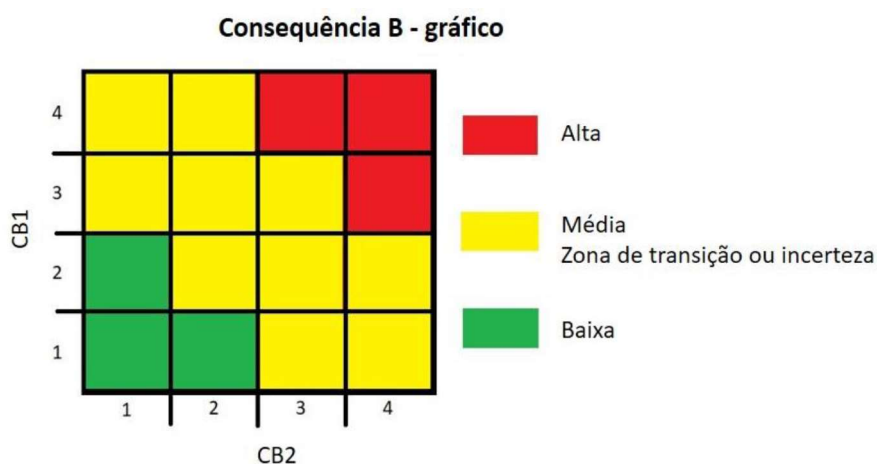
**Figura A.2** – Gráfico para Consequência A.

Para determinar a consequência B, ou chance de acontecer algum acidente devido à queda de blocos, o procedimento é diferente e mais simples do que os adotados para determinar a probabilidade e a consequência A. No caso da consequência B, basta que o usuário da metodologia atribua as notas de 1 a 4 de acordo com as características

apontadas na Tabela A.5 para os parâmetros Percentual de Distância de Visibilidade (PDV - CB1) e Risco Médio ao Veículo (RMV – CB2) e plote essas notas diretamente na matriz de consequência B (Figura A.3) para obter a classificação quanto à consequência B, que pode ser alta chance de ocorrência de acidentes, baixa chance, ou média chance (zona de transição ou de incerteza).

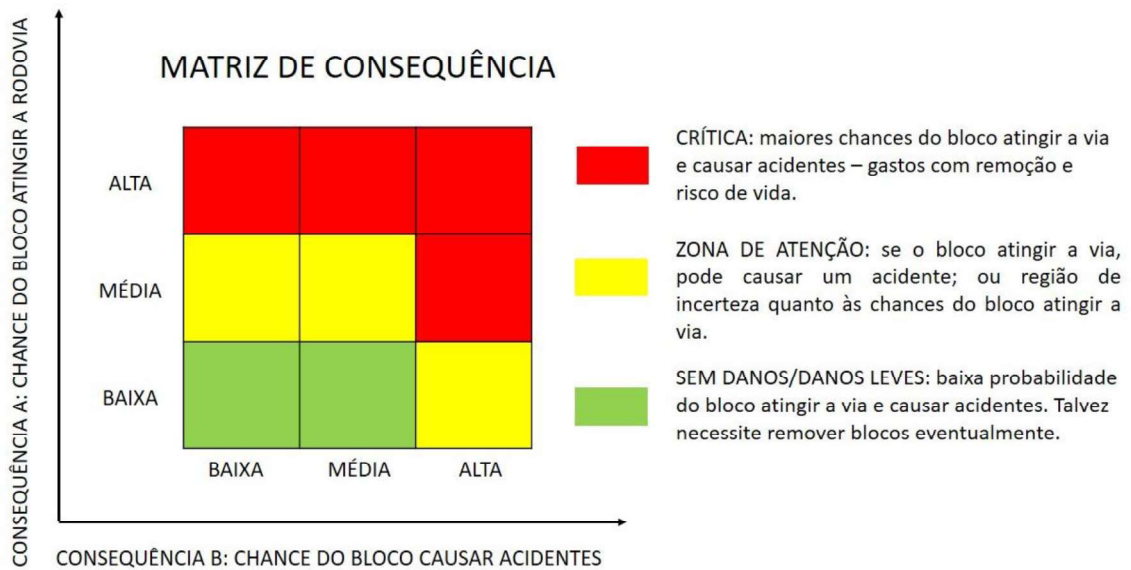
**Tabela A.5** – Parâmetros e suas devidas pontuações para análise da consequência B.

Parâmetro	Faixas			
Escores	1	2	3	4
PDV (%) (CB1)	> 80	60 - 80	40 - 60	< 40
RMV (%) (CB2)	0 - 24	25 - 49	50 - 74	> 75

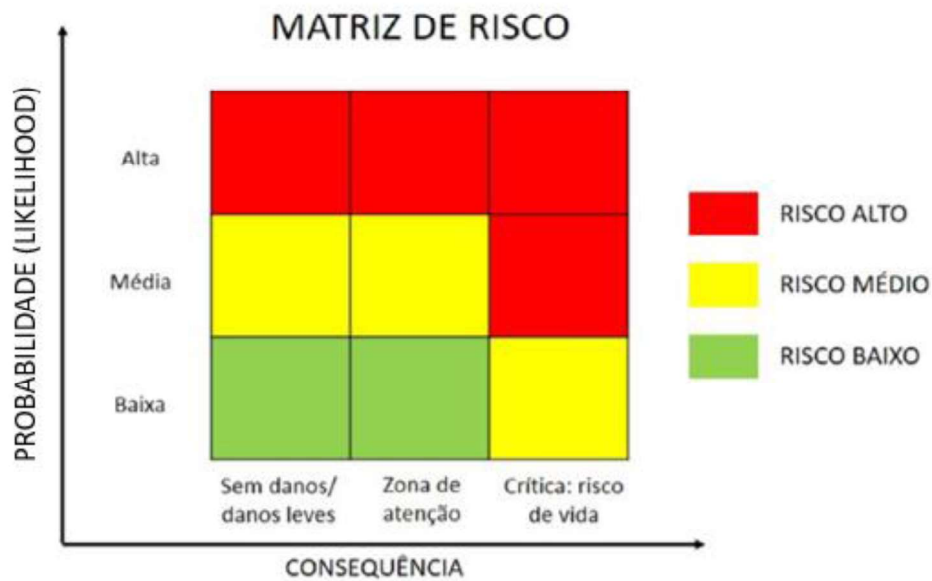


**Figura A.3** – Matriz de classificação para consequência B.

Conhecidas as classificações quanto à probabilidade de queda de blocos, à chance desses blocos atingirem a rodovia (consequência A) e à chance desses blocos causarem acidentes (consequência B), é possível finalmente obter o risco. Para isso, primeiramente se obtém a consequência global através da matriz de consequências (Figura A.4), onde são cruzadas as classificações da consequência A e da consequência B. A classificação de risco é então determinada através da matriz de risco (Figura A.6), onde a classificação da probabilidade é cruzada com a consequência global.



**Figura A.5 – Matriz de Consequências.**



**Figura A.6 – Matriz de Risco.**

Caso haja dúvida sobre como aferir um determinado parâmetro, bem como as equações, tabelas e ábacos para determinar parâmetros como Eficácia da Área de Captação (CA4), Percentual de Distância de Visibilidade (CB1) e Risco Médio ao Veículo (CB2), deve-se consultar o subitem 2.4.1, do Capítulo 2 dessa tese, onde todas as variáveis são devidamente explicadas. No Apêndice B consta um exemplo de aplicação.

## APÊNDICE B: EXEMPLO DE APLICAÇÃO E CÁLCULOS

O exemplo utilizado para demonstração dos cálculos é o mesmo cujos resultados foram mostrados no Capítulo 4. Trata-se de um dos taludes cristalinos do banco de dados utilizado para desenvolver esse trabalho. Lembrando que esse banco de dados foi gentilmente cedido por Paul Santi e todas as informações que constam no banco foram levantadas por Santi e seus colaboradores.

- **Probabilidade**

Conforme as informações adicionais que constam no banco de dados, no momento do levantamento de campo, parte do talude estava seco, mas era possível perceber sinais de infiltração sazonal; outra parte se encontrava molhada. Esse talude de rocha cristalina é cortado por pequenas (<15cm) zonas de fraqueza, como xistosidade e zonas de cisalhamento. Tanto o grau de descalçamento quanto o tamanho dos blocos se encontram na faixa entre 0,3m a 0,6m, mostrando coerência entre esses valores. Há uma família de descontinuidades persistente mergulhando para fora do talude. A abertura observada em campo está entre 1mm a 5mm e é preenchida com material granular. A descontinuidade mais crítica possui parede ondulada. Na Tabela A.6 estão as notas atribuídas às variáveis P1 a P8, que foram dadas de acordo com essas características descritas, em conformidade com a Tabela A.1.

**Tabela A.6** – Notas atribuídas às variáveis relacionadas com a probabilidade.

Variável	P1	P2	P3	P4	P5	P6	P7	P8
Nota	3	3	2	2	4	3	3	2

Para padronizar as variáveis P1 a P8, basta inserir as notas atribuídas na Equação A.1, sendo os valores das médias e desvio padrão para cada variável aqueles mostrados na Tabela A.2. Segue o cálculo para as duas primeiras variáveis (P1 e P2) como exemplo. Para as demais (P3 a P8, basta seguir os mesmos passos. Na Tabela A.7 estão os valores padronizados.

$$Z_1 = \frac{P_1 - \mu_1}{\sigma_1} = \frac{3 - 1,450}{0,684} = 2,266$$

$$Z_2 = \frac{P_2 - \mu_2}{\sigma_2} = \frac{3 - 2,659}{0,775} = 0,440$$

**Tabela A.7** – Notas padronizadas das variáveis relacionadas com a probabilidade.

Variável	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8
Nota	2,266	0,440	-0,209	-0,359	0,946	-1,284	0,452	0,710

Essas notas padronizadas estão prontas para serem inseridas nas equações A.2 a A.9 a fim de se obter as componentes  $Y_i$ . Segue novamente o exemplo de cálculo para as duas primeiras componentes ( $Y_1$  e  $Y_2$ ) utilizando, portanto, as Equações A.2 e A.3; para as demais ( $Y_3$  a  $Y_8$ ) basta seguir os mesmos passos, utilizando as Equações A.4 a A.9. Na Tabela A.8 estão os valores para todas as componentes  $Y_i$ .

$$Y_1 = 0,405 (2,266) + 0,254 (0,440) + 0,546 (-0,209) + 0,518 (-0,359) + 0,334 (0,946) + 0,280 (-1,284) + 0,078 (0,452) + 0,098 (0,710) = 0,791$$

$$Y_2 = 0,238 (2,266) + 0,468 (0,440) - 0,096 (-0,209) - 0,337 (-0,359) + 0,057 (0,946) - 0,278 (-1,284) + 0,702 (0,452) + 0,165 (0,710) = 1,732$$

**Tabela A.8** – Componentes principais para cálculo da a probabilidade.

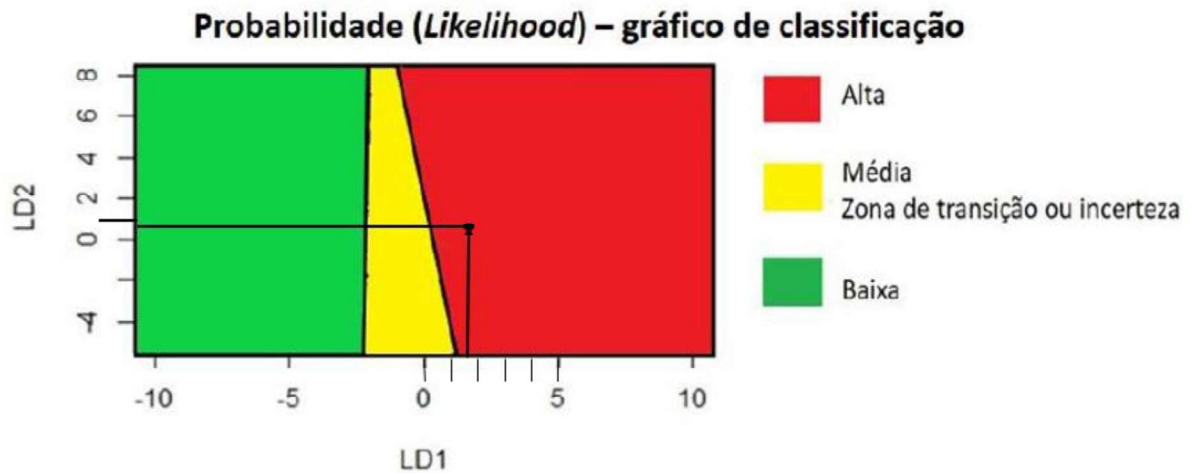
Componente	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
Valor	0,791	1,732	1,363	0,840	1,104	0,297	1,118	-0,055

Os valores obtidos para as componentes são então inseridos nas equações discriminantes LD1 e LD2 (Equações A.10 e A.11). Os valores obtidos para LD1 e LD2 são plotados no gráfico discriminante, onde é possível perceber que o talude do exemplo tem probabilidade alta de queda de blocos (Figura A.7).

$$LD_1 = 1,505 (0,791) + 0,317 (1,732) + 0,187 (1,363) + 0,031 (0,840) - 0,318 (1,104) - 0,113 (0,297) + 0,013 (1,118) + 0,106 (-0,055) = 1,644$$

$$LD_2 = 0,029 (0,791) + 0,136 (1,732) - 0,078 (1,363) - 0,325 (0,840) + 0,657 (1,104) - 0,820 (0,297) + 0,258 (1,118) + 0,313 (-0,055) = 0,632$$





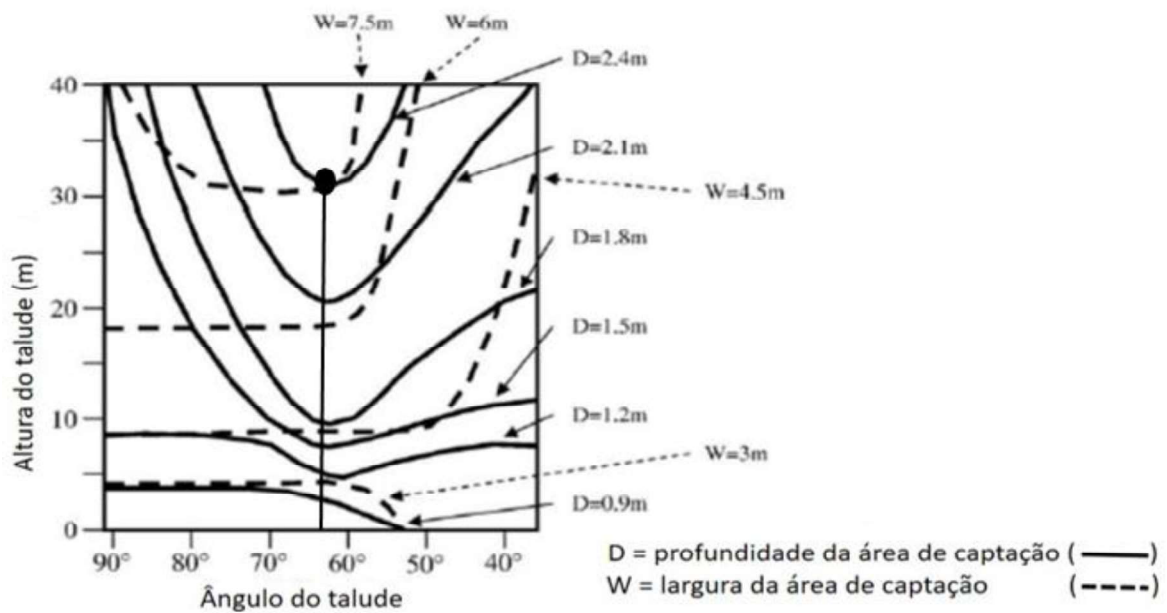
**Figura A.7** – probabilidade de queda de blocos do talude exemplo.

- **Consequência A**

De acordo com as informações inseridas no banco de dados, tanto o talude de corte quanto a porção acima do corte são fontes potenciais de blocos. Logo, a altura do talude considerada é de 91m. O ângulo de inclinação está na faixa de 62° a 65°. A superfície é irregular, com grandes feições de lançamentos. A largura medida para a área de captação está entre 1,83m a 3,66m, enquanto a profundidade medida é de 0,3m.

Para os parâmetros CA1 (altura do talude), CA2 (inclinação do talude) e CA3 (feições de lançamento), a atribuição das notas é simples e basta seguir o que indica a Tabela A.3. Para a a variável CA4 (área de captação), é necessário comparar as dimensões medidas com as dimensões recomendadas pelo ábaco de Ritchie (Ritchie, 1963) utilizando a Equação 2.19 do subitem 2.4.1 do Capítulo 2 e só então atribuir a nota mais adequada.

Observando o ábaco de Richie (Ritchie, 1963) (Figura A.8), nota-se que para a faixa de ângulos em questão, a largura recomendada ( $W_r$ ) para a área de captação é de 7,5m para taludes com altura maior que 30m. Já a profundidade recomendada ( $D_r$ ) é de 2,4m. Logo, considerando o valor mínimo medido para a largura, o somatório das dimensões medidas é 21,8% do somatório das dimensões recomendadas. Se considerar um valor médio para a largura medida, o somatório fica em torno de 30,8% do recomendado. As notas atribuídas para as variáveis CA1 a CA4 estão na Tabela A.9.



**Figura A.8** – Ábaco de Ritchie (1963) para projetar as dimensões das áreas de captação de blocos.

$$\text{Eficácia da área de captação} = 100 \times \frac{(Dm+Wm)}{(Dr+W_r)} = 100 \times \frac{(0,30+1,83)}{(2,40+7,50)} = 21,8\%$$

$$\text{Eficácia da área de captação} = 100 \times \frac{(Dm+Wm)}{(Dr+W_r)} = 100 \times \frac{(0,30+2,75)}{(2,40+7,50)} = 30,8\%$$

**Tabela A.9** – Notas atribuídas às variáveis relacionadas com a consequência A.

Variável	CA1	CA2	CA3	CA4
Nota	4	2	3	4

Para padronizar as variáveis CA1 a CA4, basta inserir as notas atribuídas na Equação A.12, sendo os valores das médias e desvio padrão para cada variável aqueles mostrados na Tabela A.4. Segue o cálculo para as duas primeiras variáveis (CA1 e CA2) como exemplo. Para as demais (CA3 e CA4), basta seguir os mesmos passos. Na Tabela A.10 estão os valores padronizados.

$$Z_{A_1} = \frac{CA_1 - \mu_{ca1}}{\sigma_{ca1}} = \frac{4 - 2,932}{1,091} = 0,980$$

$$Z_{A_2} = \frac{CA_2 - \mu_{ca2}}{\sigma_{ca2}} = \frac{2 - 2,261}{0,777} = -0,336$$

**Tabela A.10** – Notas padronizadas das variáveis relacionadas com a consequênciaA.

Variável	Z1	Z2	Z3	Z4
Nota	0,980	-0,336	0,258	0,499

Essas notas padronizadas estão prontas para serem inseridas nas equações A.13 a A.16 a fim de se obter as componentes  $Y_{ca_i}$ . Segue novamente o exemplo de cálculo para as duas primeiras componentes ( $Y_{ca1}$  e  $Y_{ca2}$ ) utilizando, portanto, as Equações A.13 e A.14; para as demais ( $Y_{ca3}$  a  $Y_{ca4}$ ) basta seguir os mesmos passos, utilizando as Equações A.15 e A.16. Na Tabela A.11 estão os valores para todas as componentes  $Y_{ca_i}$ .

$$Y_{ca1} = 0,641 (0,980) - 0,121 (-0,336) + 0,696 (0,258) + 0,299 (0,499) = 0,998$$

$$Y_{ca2} = 0,367 (0,980) + 0,740 (0,-0,336) + 0,032 (0,258) - 0,562 (0,499) = -0,161$$

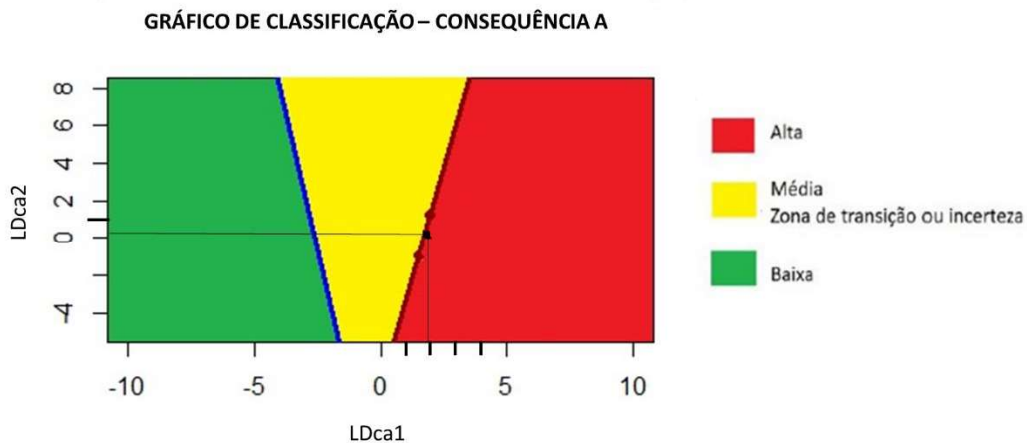
**Tabela A.8** – Componentes principais para cálculo da consequência A.

Componente	$Y_{ca1}$	$Y_{ca2}$	$Y_{ca3}$	$Y_{ca4}$
Valor	0,998	-0,161	-0,109	0,595

Os valores obtidos para as componentes são então inseridos nas equações discriminantes  $LD_{ca1}$  e  $LD_{ca2}$  (Equações A.17 e A.18). Os valores obtidos para  $LD_{ca1}$  e  $LD_{ca2}$  são plotados no gráfico discriminante, onde é possível perceber que a chance de um bloco que cai do talude do exemplo tem altas chances de atingir a rodovia (Figura A.9).

$$LD_{ca1} = 1,764 (0,998) + 0,808 (-0,161) - 0,876 (-0,109) + 0,254 (0,595) = 1,876$$

$$LD_{ca2} = 0,079 (0,998) + 0,348 (-0,161) + 1,028 (-0,109) + 0,688 (0,595) = 0,320$$



**Figura A.9** – Classificação do talude exemplo quanto à consequência A.

- **Consequência B**

As informações referentes ao tráfego também estão contidas no banco de dados. O Fluxo médio diário (FMD) é de 41200 carros por dia. O limite de velocidade (LV) no trecho onde o talude está é de 97km/h (ou 60 milhas por hora – mph). A distância de visibilidade medida ao se colocar um objeto de 15 cm na borda da rodovia é de 121,92m. Como o limite de velocidade da via é de 97km/h, pela Tabela 2.10 do Capítulo 2, toma-se que a distância de visibilidade recomendada deveria ser de 305m. A extensão do talude em questão é de 78,94m (0,07894 km).

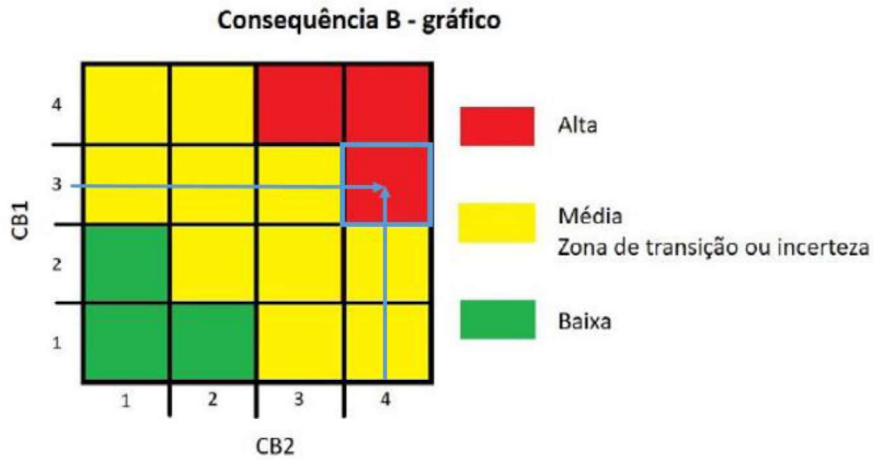
Com essas informações, é possível calcular o percentual de distância de visibilidade (PDV – CB1) conforme explicado no subitem 2.4.1 do Capítulo 2 e o risco médio ao veículo (RMV – CB2), utilizando a Equação 2.20 também mostrada no subitem 2.4.1 do Capítulo 2. Com os valores calculados, pode-se atribuir as notas a essas variáveis de acordo com a Tabela A.5.

$$PDV = \frac{\text{distância de visibilidade medida}}{\text{distância de visibilidade recomendada}} \times 100 = \frac{121,92}{305} \times 100 = 40\%$$

$$RMV = 100 \times \frac{(FMD \times EXT)/24}{LV} = 100 \times \frac{(41200 \times 0,07894)/24}{97} = 140\%$$

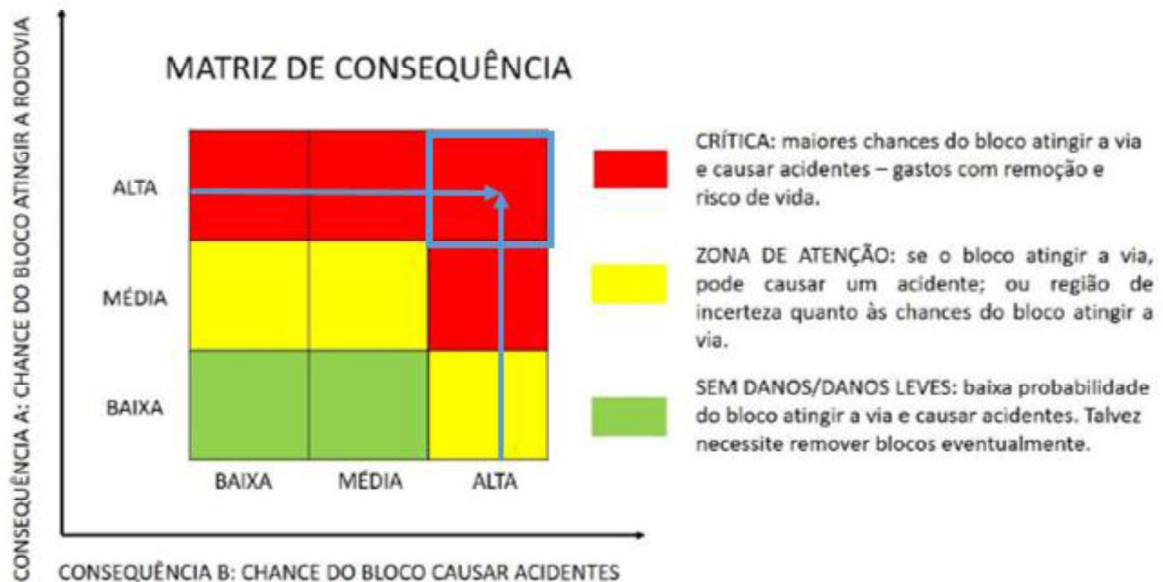
Como o parâmetro PDV (CB1) é 40%, a nota atribuída a essa variável de acordo com a Tabela A.5 é 3; como o parâmetro RMV (CB2) é maior que 100%, a nota atribuída a essa variável de acordo com a Tabela A.5 é 4. Sendo assim, é possível classificar esse talude quanto a consequência B utilizando a matriz de consequência B, onde fica

evidente que a chance de acontecer um acidente devido a queda de blocos nesse talude é alta (Figura A.10).

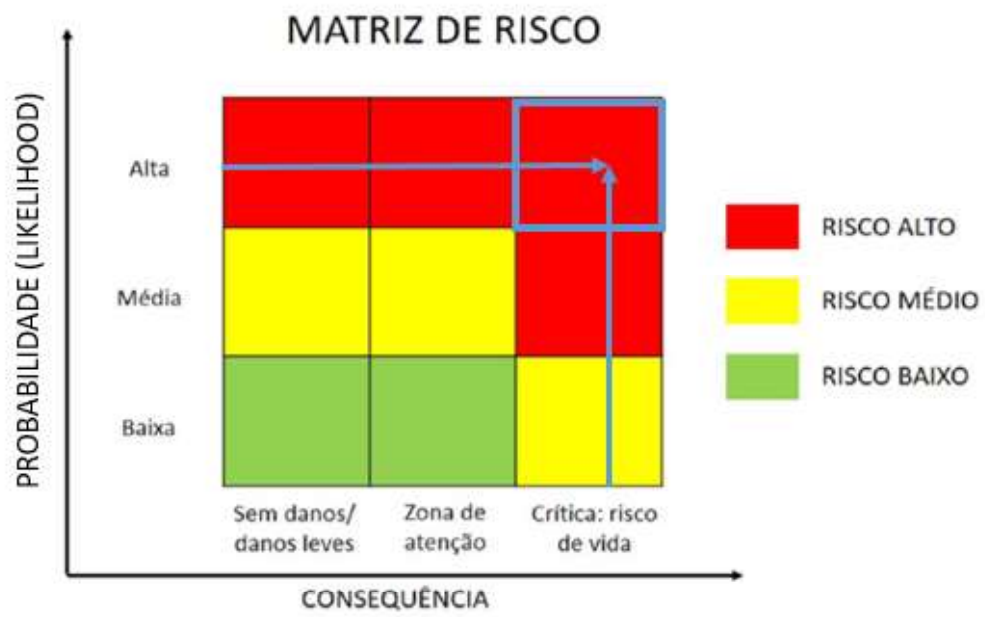


**Figura A.10** – Classificação do talude exemplo quanto à consequência B.

Sabendo que a consequência A é alta (Figura A.9) e que a consequência B também é alta (Figura A.10), é possível determinar que esse talude está em uma situação crítica quanto à consequência global, observando a matriz de consequências (Figura A.11). Como a probabilidade de queda de blocos para esse talude é alta (Figura A.7), tem-se, pela matriz de risco, que esse talude possui alto risco de queda de blocos (Figura A.12).



**Figura A.11** – Consequência global para o talude exemplo.



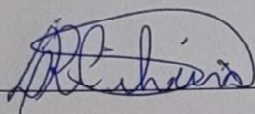
**Figura A.12** – Risco de queda de blocos para o talude exemplo.



## DECLARAÇÃO

Eu, Larissa Regina Costa Silveira, declaro que esta tese intitulada PROPOSTA DE SISTEMA DE ANÁLISE DE RISCO PARA QUEDA DE BLOCOS EM TALUDES RODOVIÁRIOS é inteiramente e exclusivamente de minha autoria e que, com exceção das citações diretas e indiretas claramente indicadas e referenciadas nesse trabalho, e do uso autorizado de banco de dados, seu texto, figuras, gráficos, quadros, tabelas, algoritmos e demais dados foram por mim obtidos e, portanto, não contêm plágio.

Ouro Preto, 18 de outubro de 2022.



---

Larissa Regina Costa Silveira

