

UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E
AMBIENTAL

MACHINE LEARNING APLICADO À PREVISÃO DE
***PIPING* EM BARRAGENS DE TERRA HOMOGÊNEAS**

MATHEUS LIMA DE BARROS

ORIENTADOR: PROF. ANDRÉ LUÍS BRASIL
CAVALCANTE, PhD

MONOGRAFIA DE PROJETO FINAL 2 EM
ENGENHARIA CIVIL

BRASÍLIA / DF: DEZEMBRO/2020

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**MACHINE LEARNING APLICADO À PREVISÃO DE *PIPING* EM
BARRAGENS DE TERRA HOMOGÊNEAS**

MATHEUS LIMA DE BARROS

MONOGRAFIA DE PROJETO FINAL SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM ENGENHARIA CIVIL.

APROVADA POR:

**ANDRÉ LUÍS BRASIL CAVALCANTE, PhD (ENC-UnB)
(ORIENTADOR)**

**ANDRÉ PACHECO DE ASSIS, PhD (ENC-UnB)
(EXAMINADOR INTERNO)**

**LENILDO SANTOS DA SILVA, PhD (ENC-UnB)
(EXAMINADOR EXTERNO)**

DATA: BRASÍLIA/DF, 16 de dezembro de 2020.

FICHA CATALOGRÁFICA

BARROS, MATHEUS LIMA

Machine Learning aplicado à Previsão de *Piping* em Barragens de Terra Homogêneas. Distrito Federal, 2020.

xiii, 97p, 210 x 297 mm (ENC/FT/UnB, Bacharel, Engenharia Civil, 2020).

Trabalho de Projeto Final – Universidade de Brasília, Faculdade de Tecnologia, Departamento de Engenharia Civil e Ambiental.

- | | |
|----------------------------|------------------------------|
| 1. Barragens de Terra | 2. <i>Piping</i> |
| 3. <i>Machine Learning</i> | 4. Redes Neurais Artificiais |
| 5. Algoritmo Boruta | 6. Testes Estatísticos |

I. ENC/FT/UnB II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

BARROS, M.L. (2020). *Machine Learning* aplicado à Previsão de *Piping* em Barragens de Terra Homogêneas. Trabalho de Projeto Final. Departamento de Engenharia Civil e Ambiental, Faculdade de Tecnologia, Universidade de Brasília, DF, 97p.

CESSÃO DE DIREITOS

NOME DO AUTOR: Matheus Lima de Barros

TÍTULO DO TRABALHO DE PROJETO FINAL: *Machine Learning* aplicado à Previsão de *Piping* em Barragens de Terra Homogêneas

GRAU/ANO: Bacharel em Engenharia Civil / 2020.

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta monografia de Projeto Final e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia pode ser reproduzida sem a autorização por escrito do autor.

Matheus Lima de Barros
matheuslb.eng@gmail.com
Brasília/DF – Brasil.

AGRADECIMENTOS

À Deus, por ter me dado a oportunidade de ingressar em uma universidade federal e participar de diversas iniciativas acadêmicas e de extensão.

À minha família, parentes e amigos que tanto me apoiaram nestes anos de graduação e que me permitiram dedicar ao máximo nas atividades, em especial Soraia, Gabriela, Alberto e Lorena. Meu muito obrigado por todo o suporte e motivação.

Ao meu orientador, prof. André Brasil, pelo apoio no desenvolvimento do trabalho e por todos os ensinamentos transmitidos a mim durante a graduação. Um professor que faz a diferença para os alunos e para a comunidade a sua volta.

Aos alunos do grupo de pesquisa Geofluxo, pelas revisões e sugestões para enriquecimento do trabalho durante e fora os seminários semanais. Em especial, Camilla, Manuelle, Nicholas, Moisés e Patrícia.

Aos professores do Departamento de Engenharia Civil e Ambiental da UnB que puderam contribuir com a minha formação, enquanto profissional e ser humano. Um agradecimento especial para as professoras, Fabiana Arruda, Cláudia Gurjão e Eugênia Fonseca, pelo acompanhamento, carinho e dedicação em me ajudar na jornada universitária. Além dos professores, Arthur Schleicher e Pastor Taco.

Aos projetos de extensão do ENC e aos estágios, que muito me fizeram crescer e conhecer pessoas comprometidas com o que fazem, o Programa de Educação Tutorial (PET), a empresa júnior Concreta Consultoria e Serviços, o grupo de pesquisa Geofluxo, o Clube de Consultoria da UnB, a empresa RHUMB Planejamento e Urbanismo e o Ministério da Defesa.

Aos colegas dos cursos de Engenharia Ambiental, em especial a turma 2014/1, e de Engenharia Civil que tive a chance de conhecer, obrigado por deixarem os momentos mais difíceis mais descontraídos.

Por fim, a todos que fizeram parte da minha graduação, muito obrigado!

RESUMO

As barragens de terra são estruturas que desempenham um papel fundamental na economia de um país. No entanto, estas estruturas estão associadas a vários riscos ambientais e humanos. O fenômeno de *piping* é conhecido mundialmente como uma das principais causas de falhas e rupturas de barragens de terra. O *piping* está associado a inúmeros fatores físicos, tais como parâmetros do solo (ângulo de atrito e coesão), aspectos geométricos das barragens (altura e seção típica), condições geológicas, e aspectos construtivos. Devido à complexidade do processo, a maioria dos modelos físicos e matemáticos não são suficientes para simular e prever a ocorrência do fenômeno. O algoritmo Boruta, cujo princípio é classificador *Random Forest*, e o teste t de hipóteses de Welch foram aplicados para identificar as variáveis mais relevantes para o *piping*. A técnica de redes neurais artificiais (RNA) foi utilizada para elaborar um modelo de previsão da ocorrência de *piping*. O banco de dados utilizado na pesquisa é composto por 23 barragens de terra homogêneas com e sem a ocorrência de *piping*. As variáveis analisadas foram o ângulo de inclinação do dreno, a altura da barragem, o nível de água a montante, a inclinação do talude a jusante, o diâmetro máximo efetivo das partículas de solo, o peso específico saturado do solo, o coeficiente de permeabilidade, a coesão efetiva do solo, e o ângulo de atrito interno efetivo. Uma análise paramétrica foi efetuada para compreender os efeitos das variações de parâmetros da simulação da RNA nos resultados obtidos. Pelas análises, as variáveis mais importantes identificadas para o *piping* são o diâmetro efetivo da partícula de solo e o ângulo de inclinação do dreno vertical. Com a aplicação da técnica de redes neurais artificiais (RNA), foram obtidos resultados satisfatórios para a previsão da ocorrência de *piping* em barragens de terra homogêneas, a partir do banco de dados e de barragens hipotéticas simuladas no *Geostudio*, considerando o mecanismo de erosão regressiva. Os resultados contribuem para a melhoria das ferramentas de previsão de ocorrência de *piping*, para apoiar o monitoramento e a instrumentação de barragens de terra e para enaltecer a importância de modelos de análise de risco e segurança de barragens.

Palavras-chave: Barragens de Terra, Segurança de Barragens, Fenômeno de Piping, Estatística, Aprendizado de Máquina, Redes Neurais Artificiais, Boruta.

ABSTRACT

Dams are structures with a fundamental role in a country's economy. However, these structures are associated with various environmental and human risks. The phenomenon of piping is known worldwide as one of the main causes of failures and ruptures of homogeneous earth dams. Piping is associated with numerous physical factors such as soil parameters (angle of friction and cohesion), geometric aspects of dams (height and typical section), geological conditions, and construction aspects. Due to the complexity of the process, most physical and mathematical models are not sufficient to simulate and predict the occurrence of the phenomenon. The Boruta algorithm, whose principle is the Random Forest classifier, and Welch's t hypothesis test were applied to identify the most relevant variables for piping. The Artificial Neural Network (ANN) technique was used to elaborate a model to predict the occurrence of piping. The database used in the research is composed by 23 homogeneous earth dams with and without the occurrence of piping. The variables analyzed were the inclination angle of the vertical drain, the dam height, the water level upstream, the slope downstream, the effective diameter of soil particles, the specific saturated soil weight, the permeability coefficient, the effective soil cohesion, and the effective internal friction angle. A parametric analysis was carried out to understand the effects of the RNA simulation parameters on the results obtained. By the analysis, the most important variables identified for piping are the effective diameter of the soil particle and the inclination angle of the vertical drain. With the application of the Artificial Neural Networks (ANN) technique, satisfactory results were obtained to predict the occurrence of piping in homogeneous earth dams, from the database and hypothetical dams simulated in Geostudio, for the backward erosion mechanism. The results contribute to the improvement of piping prediction tools, to support the monitoring and instrumentation of dams and to highlight the importance of the evaluation of dam risk and safety analysis models.

Key words: Embankment Dams, Homogeneous Earth Dams, Dam Safety, Piping, Statistics, Machine Learning, Artificial Neural Networks, Boruta Algorithm.

SUMÁRIO

1. INTRODUÇÃO	1
1.1. CARACTERIZAÇÃO DO PROBLEMA	2
1.2. JUSTIFICATIVA	3
1.3. OBJETIVOS	6
1.3.1. Objetivo Geral	6
1.3.2. Objetivos Específicos	6
1.4. DESCRIÇÃO DOS CAPÍTULOS	6
2. REVISÃO BIBLIOGRÁFICA	8
2.1. BARRAGENS DE TERRA	8
2.2. PRINCIPAIS CAUSAS DE RUPTURA DE BARRAGENS DE TERRA	13
2.3. EROÇÃO INTERNA	15
2.4. FLUXO EM MEIO POROSO	19
2.5. ANÁLISE DE ESTABILIDADE	24
2.6. TESTES ESTATÍSTICOS	28
2.7. BORUTA E FLORESTAS ALEATÓRIAS	30
2.8. REDES NEURAIS ARTIFICIAIS (RNA)	35
2.9. RNA APLICADAS PARA ANÁLISES DE PIPING	43
3. METODOLOGIA	46
3.1. REVISÃO DA LITERATURA E DOS SOFTWARES	47
3.2. OBTENÇÃO DE DADOS DE ESTUDOS ANTERIORES	48
3.3. ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES	48
3.4. ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA) ..	49
3.4.1. Escolha de Parâmetros de Entrada e Treinamento	49
3.4.2. Simulações e Validações do Modelo de RNA	49
3.5. ANÁLISE DE RESULTADOS E CONCLUSÕES	52
4. BANCO DE DADOS	53
5. RESULTADOS E DISCUSSÕES	55
5.1. ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES	55
5.1.1. Matriz de Correlação	55
5.1.2. Algoritmo Boruta	56
5.1.3. Testes Estatísticos	57
5.2. ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA) ..	59
5.2.1. Análise Paramétrica – Neurônios	60
5.2.2. Análise Paramétrica – Camadas	63
5.2.3. Análise Paramétrica – Algoritmos de Otimização	65
5.2.4. Análise Paramétrica – Funções de Ativação	66

5.2.5.	Retroanálise do Banco de Dados	67
5.2.6.	Validação da RNA	69
6.	CONCLUSÕES E SUGESTÕES.....	75
6.1.	ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES.....	75
6.2.	ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA) ..	77
6.3.	SUGESTÕES PARA TRABALHOS FUTUROS	78
	REFERÊNCIAS BIBLIOGRÁFICAS	79

LISTA DE FIGURAS

Figura 1.1 – Evolução da ocorrência de piping em barragens de terra. Fonte: Okeke & Wang (2016).	3
Figura 1.2 – Impactos do desastre de Mariana – MG. Fonte: Agência Brasil (2019).	4
Figura 2.1 – Tipos de Barragens de Terra. Fonte: Adaptado de Zhang et al. (2009).	9
Figura 2.2 – Seção tipo de uma barragem de terra homogênea, Vigário, Brasil. Fonte: Assis et al. (2003).	9
Figura 2.3 – Modos de Falhas em Barragens de Terra. Fonte: The Constructor (2019).	14
Figura 2.4 – Tipos de Erosão Interna em Barragens. Fonte: Adaptado de Molinder (2016).	18
Figura 2.5 – Representação da Lei de Darcy. Fonte: Das (2019).	19
Figura 2.6 – Forças atuantes em uma fatia de solo. Fonte: Geostudio (2012).	25
Figura 2.7 – Variação do parâmetro λ e fatores de segurança. Fonte: Geostudio (2012).	28
Figura 2.8 – Exemplos de precisão e exatidão. Fonte: GM Metrologia (2020).	32
Figura 2.9 – Passo 1 do Algoritmo Boruta. Fonte: Kursu & Rudnicki (2010).	34
Figura 2.10 – Representação de um neurônio biológico. Fonte: Mundo Educação (2019).	36
Figura 2.11 – Estrutura do Perceptron. Fonte: Adaptado de Rosenblatt (1957).	37
Figura 2.12 – Exemplos de step functions.	38
Figura 2.13 – Funções de ativação e a variação do termo bias.	38
Figura 2.14 – Estrutura de uma RNA com o parâmetro bias. Fonte: Adaptado de Towards Data Science (2019).	39
Figura 2.15 – Algoritmo de Backpropagation e a função erro. Fonte: Gunther & Fritsch (2010).	41
Figura 2.16 – Resultados do treinamento. Fonte: Tayfur et al. (2005).	43
Figura 2.17 – Resultados da simulação. Fonte: Tayfur et al. (2005).	44
Figura 2.18 – Resultados da simulação com GRNN. Fonte: Xue et al. (2014).	45
Figura 3.1 – Metodologia Geral do Projeto Final.	46
Figura 3.2 – Metodologia Específica da Etapa 1.	46
Figura 3.3 – Metodologia Específica da Etapa 2.	47
Figura 3.4 – Barragem Padrão no Geostudio para a Validação da RNA.	50
Figura 4.1 – Resumo de métricas estatísticas das variáveis.	53
Figura 5.1 – Matriz de correlação das variáveis.	55
Figura 5.2 – Resultados do algoritmo Boruta.	56
Figura 5.3 – Arquitetura da RNA modelo da simulação.	59
Figura 5.4 – Resultados com uma camada oculta (N = 1 e N = 2 neurônios).	60
Figura 5.5 – Resultados com uma camada oculta (N = 3 e N = 4 neurônios).	61
Figura 5.6 – Resultados com uma camada oculta (N = 5 e N = 6 neurônios).	61
Figura 5.7 – Resultados com uma camada oculta (N = 7 e N = 8 neurônios).	62
Figura 5.8 – Resultados com uma camada oculta (N = 9 e N = 10 neurônios).	62
Figura 5.9 – Resultados com mais de uma camada oculta (N = 4 e N = 4,4 neurônios).	63
Figura 5.10 – Resultados com mais de uma camada oculta (N = 5 e N = 5,5 neurônios).	64
Figura 5.11 – Resultados com mais de uma camada oculta (N = 4 e N = 4,4,4 neurônios).	64

Figura 5.12 – Resultados com mais de uma camada oculta (N = 5 e N = 5,5,5 neurônios).....	65
Figura 5.13 – Resultados dos Algoritmos de Otimização.....	66
Figura 5.14 – Resultados das Funções de Ativação.....	67
Figura 5.15 – Arquitetura da RNA após Treinamento.....	68
Figura 5.16 – Resultados da Previsão da RNA após Treinamento.....	68
Figura 5.17 – Resultados da Simulação no SLOPE/W.....	70
Figura 5.18 – Resultados da Simulação no SEEP/W.....	70
Figura 5.19 – Barragem Hipotética com Dreno Vertical a 90.....	71
Figura 5.20 – Barragem Hipotética com Dreno Vertical a 45.....	71
Figura 5.21 – Validação da RNA com Resultados do <i>Geostudio</i>	74

LISTA DE TABELAS

Tabela 2.1 – Relatos de erosões internas nos RSB de 2011 a 2019. Fonte: ANA, 2020.	11
Tabela 2.2 – Escala de evidência do valor p. Fonte: Wasserman (2004).....	30
Tabela 3.1 – Critérios da simulação do SLOPE/W. Fonte: Adaptado de Gameleira e Amaral (2018).....	51
Tabela 3.2 – Critérios da simulação do SEEP/W. Fonte: Adaptado de Gameleira e Amaral (2018).....	51
Tabela 4.1 – Banco de dados de barragens de terra homogêneas. Fonte: Xue et al. (2014).	54
Tabela 5.1 – Resultados do teste t e do valor p.	57
Tabela 5.2 – Comparação de Resultados da Base de Dados com os Previstos.	69
Tabela 5.3 – Barragens Hipotéticas da Validação.	72
Tabela 5.4 – Dimensões de Elementos das Barragens Hipotéticas.	72
Tabela 5.5 – Resultados do SEEP/W e do SLOPE/W para as Barragens Hipotéticas. .	73
Tabela 5.6 – Comparação de Resultados da Base de Dados com os Previstos.	74

LISTA DE SIGLAS

ANA	Agência Nacional de Águas
Aneel	Agência Nacional de Energia Elétrica
ANM	Agência Nacional de Mineração
ANN	<i>Artificial Neural Network</i>
BP	<i>Backpropagation</i>
CEB	Companhia Energética de Brasília
CNEN	Comissão Nacional de Energia Nuclear
CNRH	Conselho Nacional de Recursos Hídricos
CRI	Categoria de Risco
DPA	Dano Potencial Associado
EPE	Empresa de Pesquisa Energética
GRNN	<i>Generalized Neural Networks</i>
ICOLD	<i>International Commission on Large Dams</i>
LTU	<i>Linear Threshold Unit</i>
MEF	Método dos Elementos Finitos
MLP	<i>Multilayer Perceptron</i>
MME	Ministério de Minas e Energia
PNSB	Política Nacional de Segurança de Barragens
PSB	Plano de Segurança da Barragem
RBP	<i>Resilient Backpropagation</i>
RNA	Rede Neural Artificial
RSB	Relatório de Segurança de Barragens
SEDM	<i>Standard Error of the Difference of Means</i>
SEM	<i>Standard Errors of the Mean</i>
SNISB	Sistema Nacional de Informações sobre Segurança de Barragens

1. INTRODUÇÃO

O desenvolvimento socioeconômico de um país está diretamente relacionado com a geração de energia. Países desenvolvidos apresentam uma oferta de energia elétrica per capita superior aos países em desenvolvimento, onde situa-se o Brasil (GOLDEMBERG, 2000). Segundo a EPE (2019), a matriz elétrica brasileira de 2018 foi composta por 66,6% de energia oriunda de usinas hidrelétricas, um total de 423,9 TWh. Além disso, conforme o balanço energético realizado pela EPE (2019), no ano de 2018 a participação de fontes renováveis de energia na matriz elétrica brasileira foi de 83,3%.

A importância das usinas hidrelétricas é notável frente à produção de energia elétrica no Brasil, dado o alto potencial hidráulico dos corpos hídricos presentes no território nacional. A construção de barramentos com a finalidade de formação de reservatórios é prática comum também em outros países, visto que a energia hidrelétrica apresenta um baixo custo de produção e utiliza uma fonte renovável de energia, a água. Outra utilização comumente adotada é a contenção de rejeitos industriais e de mineração. Em que cabe destacar que segundo o MME (2019), a mineração contribui com o Produto Interno Bruto (PIB) nacional em cerca de 4%, com 180 mil trabalhadores diretos e 2 milhões indiretos.

Dessa maneira, as barragens são estruturas geotécnicas projetadas para diversos usos, em que se pode citar a acumulação de água, a geração de energia elétrica, a regularização de vazão, a disposição de rejeitos de mineração, a disposição de rejeitos industriais e o uso recreativo. Por tratarem de obras de significativo impacto ambiental e econômico, o planejamento, a execução e o monitoramento das barragens devem ser realizados de modo a garantir o cumprimento de requisitos mínimos de segurança estabelecidos por órgãos fiscalizadores. Como empreendimentos de grande porte, o risco potencial de ruptura deve ser avaliado e monitorado de maneira rigorosa, visto que um rompimento pode ocasionar em desastres ambientais, prejuízos econômicos e perda de vidas humanas.

No Brasil, as barragens são reguladas por diferentes órgãos governamentais de acordo com o uso. Quanto ao uso de disposição de rejeitos de mineração, essas são de responsabilidade da Agência Nacional de Mineração (ANM). As hidrelétricas, por sua vez, são fiscalizadas pela Agência Nacional de Energia Elétrica (Aneel). Os barramentos

utilizados para retenção de água para abastecimento são regulados pela Agência Nacional de Águas (ANA). Se a barragem está relacionada com produção de energia nuclear, Comissão Nacional de Energia Nuclear (CNEN) é o órgão fiscalizador. Os demais casos são de responsabilidade de órgãos locais estaduais ou de entidades que fornecem licenças ambientais de instalação e operação. Caso a barragem apresentar determinados requisitos, essa deve seguir as diretrizes da Lei 12.334/2010 no qual estabelece a Política Nacional de Segurança de Barragens (PNSB).

O monitoramento e a fiscalização do comportamento dessas estruturas são fundamentais para garantir a prevenção de acidentes, considerando a quantidade de riscos associados à sua construção e operação. A ruptura pode ocorrer por distintos fenômenos, dentre eles, galgamento, erosão interna, liquefação e instabilidade de taludes. Segundo Ferdos (2016), os fenômenos de erosão interna constituem uma das principais causas de ruptura de barragens de terra. Esses fenômenos, denominados genericamente de *piping*, ocorrem pela infiltração de água no solo de fundação ou no próprio corpo da barragem, no qual provoca o carreamento de finos, que por conseguinte, causa a instabilidade.

1.1. CARACTERIZAÇÃO DO PROBLEMA

Foster *et al.* (2000) realizou um estudo estatístico de 11.192 barragens construídas entre 1800 e 1986, contemplando 136 casos de rompimento, em que foi constatado que o *piping* é a maior causa das rupturas das barragens em operação, com cerca de 49,0%. Outros modos de falha apresentaram menores ocorrências, liquefação com cerca de 1,7% e escorregamentos de terra com 4,3%. Os demais casos estão associados com o fenômeno de galgamento, quando ocorre a transposição do barramento por uma massa de água.

Conforme Ferdos (2016), a erosão interna ocorre de maneira silenciosa e sob condições específicas, os processos acontecem de maneira restrita ao solo de fundação ou ao próprio corpo da barragem. Dessa forma, apenas detecta-se o *piping* quando são observados danos visíveis na estrutura pelo carreamento do material poroso constituinte ou pela emissão de alertas de sistemas de monitoramento. O estudo de novas técnicas de identificação da ocorrência de erosão interna em barragens torna-se fundamental para evitar acidentes.

As partículas de solo, na ocorrência do fenômeno de *piping*, são transportadas localmente, geralmente em suspensão, pelas forças hidrodinâmicas da água (Zhang *et al.*, 2016). O movimento do material particulado através da estrutura ou da fundação da barragem forma uma espécie de conduto interno, que pode progredir até causar a ruptura total. A modelagem do fenômeno é complexa e o seu surgimento está relacionado a diversos fatores da barragem e do meio (ZHANG *et al.*, 2016; FOSTER *et al.*, 2000). Na Figura 1.1, ilustra-se a evolução do *piping* em uma barragem de terra até a fase de colapso.

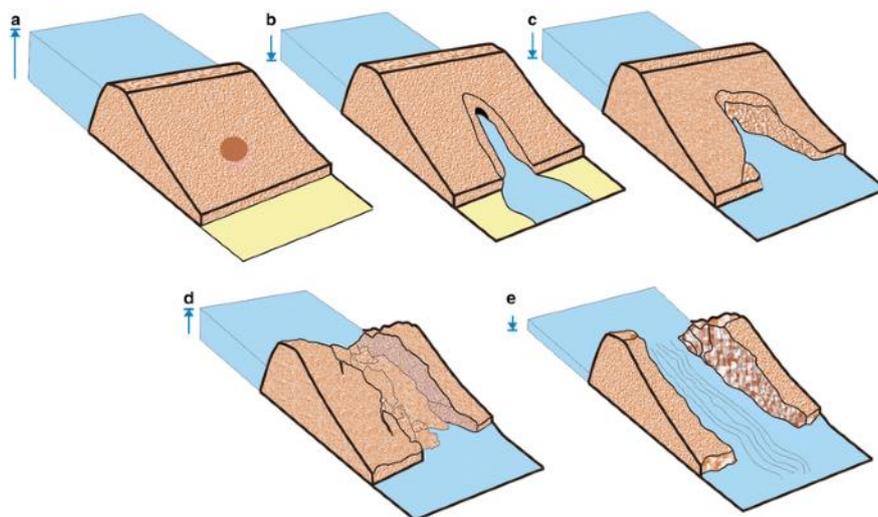


Figura 1.1 – Evolução da ocorrência de *piping* em barragens de terra. Fonte: Okeke & Wang (2016).

De acordo com a ANA (2018), no Relatório de Segurança de Barragens de 2017, cerca de 76% das barragens cadastradas no Sistema Nacional de Informações sobre Segurança de Barragens (SNISB) não apresentaram informações suficientes para que seja possível determinar se devem se submeter ou não à PNSB. A falta de dados confiáveis aliada à ausência de fiscalização das condições de segurança das barragens impossibilita o gerenciamento efetivo dos riscos de acidentes.

1.2. JUSTIFICATIVA

De acordo Cielen *et al.* (2016), a ciência de dados é definida pela utilização de métodos para analisar grandes quantidades de dados e extrair informações a partir disso. Como analogia, o autor compara os dados não tratados com óleo bruto, enquanto a ciência de dados seria a refinaria que permite transformar a matéria prima em produtos para tomada de decisão. Os sistemas de monitoramento de segurança de estruturas complexas, como barragens e pontes, geram um volume massivo de informações diariamente. Ao unir os

dados de uma estrutura às outras semelhantes, considerando também o histórico dessas informações e a sua evolução ao longo do tempo, a ciência de dados apresenta-se como uma alternativa para produzir informações relevantes a partir de números brutos.

A aplicação de técnicas de análise de dados é recorrente em diversos campos do conhecimento, onde há a necessidade de tratar volumes densos de informações. No que tange a engenharia geotécnica, essas novas metodologias facilitam o monitoramento de estruturas complexas, permitem simulações numéricas considerando ensaios em campo, além de possibilitarem prever o comportamento e propriedades de elementos ao levar em conta séries históricas. O processo de previsão de informações a partir de um conjunto de dados de entrada é comumente denominado de *machine learning*, traduzido como aprendizado de máquina para o português (CIELEN *et al.*, 2016).

No Brasil, a temática de segurança de barragens ganhou destaque nos últimos anos pela ruptura das barragens de rejeitos em Mariana – MG (G1, 2020) e Brumadinho – MG (G1, 2019). As rupturas foram causadas por outros fenômenos e ocasionaram em perdas incalculáveis, ambientais e de vidas humanas, conforme a Figura 1.2. Os acontecimentos induziram uma maior notoriedade do assunto perante a opinião pública, por um acompanhamento mais efetivo das estruturas por parte de órgãos fiscalizadores.



Figura 1.2 – Impactos do desastre de Mariana – MG. Fonte: Agência Brasil (2019).

No Distrito Federal (DF), a Companhia Energética de Brasília (CEB), em colaboração com pesquisadores da UnB, firmou um projeto de pesquisa e desenvolvimento para o estudo da ocorrência de *piping* em barragens. Trata-se do projeto AINOA – Sistema de Monitoramento de Patologias Internas em Barragens de Terra e Enrocamento Fundamentado em Inteligência Artificial e Internet das Coisas: Caso de Estudo da Barragem do Paranoá – DF.

O objetivo do projeto é desenvolver um sistema de monitoramento de erosão interna em barragens de terra e enrocamento, fundamentado nos conceitos de inteligência artificial e internet das coisas. O monitoramento proposto é por meio de sensores piezoelétricos e geofones, em que será elaborado um sistema integrado de gerenciamento de dados em tempo real. Além da criação de modelos reduzidos em laboratório, para testes de sensores e compreender o fenômeno de *piping*.

O sistema será aplicado na barragem do Paranoá no Distrito Federal (DF), utilizada principalmente para geração de energia e classificada pelo SNISB (2020) com categoria de risco baixo e dano potencial associado alto. O projeto está em andamento e possui uma publicação internacional no *Fourth International DAM WORLD Conference*, advinda deste trabalho final de graduação (BARROS *et al.*, 2020).

Este projeto final, portanto, ao considerar o tema recorrente de segurança de barragens no contexto do projeto AINOA, visa trazer os benefícios das técnicas modernas de aprendizado de máquina para a obtenção de informações relevantes e a previsão do comportamento de barramentos de terra. Para isso, optou-se pelo uso de técnicas que utilizam um algoritmo criado a partir de árvores de decisão e das redes neurais artificiais (RNA). Com o intuito de, a partir de um banco de dados estruturado de barragens de terra homogêneas, identificar as variáveis que mais afetam a ocorrência de *piping* e elaborar um modelo de previsão da ocorrência.

1.3. OBJETIVOS

1.3.1. Objetivo Geral

O objetivo geral deste projeto final é utilizar algoritmos de aprendizado de máquina (*machine learning*) e análises estatísticas para realizar a identificação dos fatores chave que mais contribuem para a ocorrência de *piping* em barragens de terra homogêneas e propor um modelo numérico de previsão do fenômeno, considerando o mecanismo de erosão regressiva.

1.3.2. Objetivos Específicos

Os objetivos específicos a serem desenvolvidos são:

- a) Documentar parâmetros físicos e bases de dados de *piping* em barragens de terra homogêneas a partir da consulta à literatura;
- b) Avaliar o grau de relevância de cada parâmetro para a ocorrência do fenômeno, de acordo com a base de dados utilizada, por meio de algoritmos de aprendizado de máquina;
- c) Comparar os resultados obtidos pelo algoritmo de aprendizado de máquina com análises estatísticas;
- d) Realizar análises paramétricas do modelo de rede neural artificial (RNA) para possibilitar uma maior sensibilidade do seu funcionamento;
- e) Comparar os resultados obtidos pela RNA com modelos de estabilidade e simulação de fluxo de água através de barragens de terra homogêneas, nos módulos SLOPE/W e SEEP/W do *Geostudio*.

1.4. DESCRIÇÃO DOS CAPÍTULOS

O Capítulo 1 visa apresentar o tema a ser trabalhado neste projeto final, a partir da contextualização de sua relevância e a aplicação de técnicas de aprendizado de máquina na engenharia. Além disso, abordou os objetivos propostos geral e específicos, além da estrutura do projeto final.

O Capítulo 2 enuncia a revisão da literatura técnica com foco do tema a ser pesquisado, em que se contemplam tópicos a respeito de conceitos básicos das barragens de terra homogêneas, principais causas de ruptura, conceituação e modelagem do fenômeno de *piping*, assim como, revisão dos conceitos de algoritmos de aprendizado de máquina, testes estatísticos, definições de RNA e estudos anteriores.

O Capítulo 3 apresenta a metodologia proposta para o desenvolvimento dos objetivos detalhados no Capítulo 1, por meio de diagramas explicativos e a descrição das etapas envolvidas.

No Capítulo 4, o banco de dados de barragens de terra homogêneas utilizado na pesquisa é apresentado e caracterizado quanto a métricas estatísticas convencionais.

Nos Capítulos 5 e 6, são elencados os principais resultados alcançados e conclusões gerais, a partir da metodologia adotada e as simulações computacionais realizadas com o banco de dados. Por fim, serão destacados os fatores limitantes, com sugestões para trabalhos futuros, com o intuito de contribuir para a continuação da linha de pesquisa.

2. REVISÃO BIBLIOGRÁFICA

2.1. BARRAGENS DE TERRA

A Lei nº 12.334 de 20 de setembro de 2010 da Casa Civil, estabelece a Política Nacional de Segurança de Barragens (PNSB) e cria o Sistema Nacional de Informações sobre Segurança de Barragens (SNISB). A Lei nº 12.334/2010 apresenta a seguinte definição de uma barragem:

Qualquer estrutura em um curso permanente ou temporário de água para fins de contenção ou acumulação de substâncias líquidas ou de misturas de líquidos e sólidos, compreendendo o barramento e as estruturas associadas.

Segundo a *International Commission on Large Dams – ICOLD* (2019), as barragens de terra representam 65% das 57.985 estruturas cadastradas, seguidas das barragens de concreto convencional por gravidade (14%) e de enrocamento (13%). A ICOLD (2019) contempla apenas as denominadas grandes barragens, aquelas que apresentam pelo menos 15 m do ponto mais baixo das fundações até o topo da crista ou um reservatório com capacidade total equivalente maior ou igual a 3 milhões de metros cúbicos. A Lei nº 12.334/2010 aplica-se para barragens que atendam pelo menos uma das características citadas pela ICOLD (2019) para definir grandes barragens. Aplica-se também caso o barramento possua reservatório com resíduos perigosos ou quando o dano potencial associado (DPA) é considerado médio ou alto.

Dentre os tipos de barragens, Assis (2003) lista os principais fatores para a tomada de decisão da tipologia mais adequada, como a existência e proximidade de material de qualidade para construção, as condições geológicas e geotécnicas da região e a topografia. Além disso, a escolha é realizada com base em outros parâmetros, tais como, disponibilidade de solo ou rocha, natureza das fundações e condições climáticas (ASSIS, 2003). De acordo com Zhang *et al.* (2009), as barragens de terra podem ser subdivididas em quatro categorias típicas, conforme a Figura 2.1: (a) barragens de terra homogêneas, (b) barragens de terra zonadas, (c) barragens de terra zonadas com núcleo impermeável e (d) barragens de terra com face do concreto.

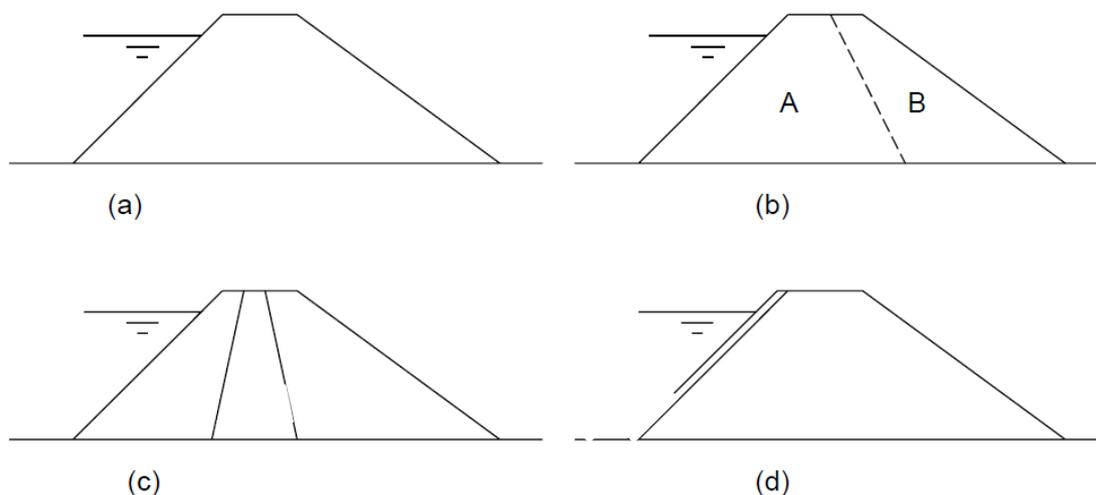


Figura 2.1 – Tipos de Barragens de Terra. Fonte: Adaptado de Zhang *et al.* (2009).

As barragens de terra homogêneas apresentam um tipo de solo principal em sua seção, e conforme Assis (2003), exigem menores declividades dos taludes de montante e jusante, conseqüentemente, o que acarreta o maior volume do corpo da barragem. As barragens de terra zonadas possuem distintos tipos de solo em sua seção, separados em regiões. No caso de barragens de terra com núcleo impermeável, normalmente de argila ou de concreto, a estrutura garante um maior controle do fluxo de água através do corpo. Na última categoria, é construída uma face impermeável de concreto à montante da barragem, também com o intuito de minimizar a infiltração pelo corpo do barramento (ASSIS, 2003; ZHANG *et al.*, 2009).

A seção típica de uma barragem de terra pode ser vista na Figura 2.2, em que há um material predominante constituinte do corpo, além da presença de elementos típicos como drenos e reforços de enrocamento.

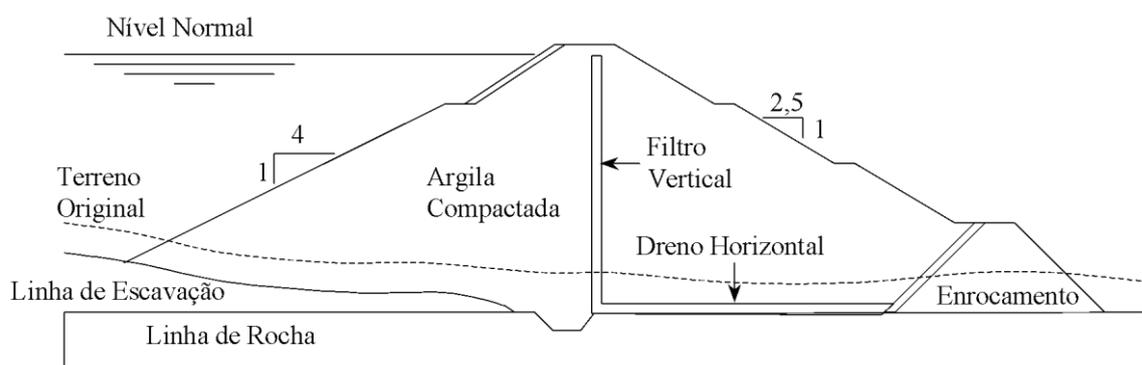


Figura 2.2 – Seção típica de uma barragem de terra homogênea, Vigário, Brasil. Fonte: Assis *et al.* (2003).

Segundo a ANA (2020), as barragens são definidas como “estruturas físicas que represam um curso de água”. A Lei nº 12.334/2010 distribui a responsabilidade de fiscalização dos barramentos, conforme o uso. Para fins de acumulação de água, geração de eletricidade, disposição final ou temporária de resíduos de mineração ou disposição de resíduos industriais, a fiscalização está sob responsabilidade, respectivamente, do órgão que outorgou o uso dos recursos hídricos, que autorizou o uso do potencial hidráulico, que outorgou os direitos minerários ou que forneceu a autorização da licença ambiental para instalação e operação.

A mesma lei define a ANA como responsável pelo gerenciamento do Sistema Nacional de Informações sobre Segurança de Barragens (SNISB), no qual emite um relatório de segurança dos barramentos do país anualmente, o RSB. Conforme o art. 7º da Lei 12.334/2010 e o Relatório de Segurança de Barragens de 2019, as barragens são classificadas a partir de três critérios principais:

- Categoria de Risco (CRI): relacionada principalmente à estrutura da barragem. É função das características técnicas do barramento, do estado de conservação e o atendimento ao Plano de Segurança da Barragem (PSB). Varia entre alta, média ou baixa.
- Dano Potencial Associado (DPA): relacionado principalmente na área afetada a jusante da barragem. É função do potencial de perda de vidas humanas e impactos ambientais, sociais e econômicos de uma possível ruptura. Varia entre alto, médio e baixo.
- Volume do Reservatório: a resolução do Conselho Nacional de Recursos Hídricos (CNRH) de nº 143/2012 define a graduação do volume do reservatório conforme o uso da barragem, com acúmulo de água ou disposição de rejeitos industriais ou minerais.

De acordo com o Relatório de Segurança de Barragens (RSB) de 2019 (ANA, 2020), estão cadastradas no SNISB ao todo 19.388 barragens, distribuídas nos estados brasileiros e no Distrito Federal. Cerca de 56% dessas, desconsiderando os cadastros sem informação, são constituídas de barragens de terra. Das 19.388 estruturas, 1.096 foram classificadas como Categoria de Risco (CRI) e Dano Potencial Associado (DPA) altos, ou seja, possuem uma probabilidade de falha elevada e as consequências de uma eventual ruptura seriam graves.

A ANA, desde 2016, solicita aos órgãos fiscalizadores, quando da elaboração do RSB anual, uma lista de barragens com nível significativo de atenção. Dessa forma, 156 barragens no ano de 2019 foram elencadas nessa situação por 26 dos 33 órgãos fiscalizadores, por conta do estado de conservação das estruturas, ausência de documentos, deficiências de projetos, dentre outros. No RSB de 2018, 68 barragens estavam nesta situação, segundo relatos. O aumento em comparação com o ano de 2019 ocorreu em virtude do aumento do número de fiscalizações pelos órgãos responsáveis, o que permitiu uma maior caracterização do estado das barragens.

Segundo o RSB de 2017, os órgãos fiscalizadores apontaram 45 barragens em estado de atenção, por conta de insuficiência do vertedor, baixo nível de conservação e a falta de documentos que comprovem a estabilidade da estrutura. No RSB de 2016, foram apontadas 25 barragens com algum comprometimento, além de 695 estruturas classificadas com DPA e CRI altos. Nota-se, ao longo dos anos, o aumento do número de casos preocupantes para os órgãos fiscalizadores, o que exige maiores estudos para identificar e remediar falhas em conformidade com a PNSB.

A Tabela 2.1 apresenta um resumo de consultas de relatos de ocorrências de erosões internas ou *piping* em barragens nos Relatórios de Segurança de Barragens de 2011 a 2019 (ANA, 2020).

Tabela 2.1 – Relatos de erosões internas nos RSB de 2011 a 2019. Fonte: ANA, 2020.

RSB	Nº de Relatos de Erosões Internas	Órgão Fiscalizador	Localidade
RSB 2011	0	-	-
RSB 2012/2013	3	ANA	Barragem Jaburu I/CE e Barragem Capoeira/PB
		ANEEL	PCH Arvoredo/SC
RSB 2014	4	ANEEL	Barragem UHE Jirau/RO
		ANA	Barragem Fazenda Boa Vista do Uru/GO e Barragem Mineradora Herculano/MG
		SERHMACT/PB	Barragem Gramame/PB
RSB 2015	2	ANEEL	PCH Inxu/MT
		DRHI/SC	Barragem na propriedade do senhor Nei Carlos Pedro Zampieri/SC
RSB 2016	6	Instituto das Águas do Paraná	Barragem na Fazenda Guavirova/PR

RSB	Nº de Relatos de Erosões Internas	Órgão Fiscalizador	Localidade
RSB 2016	6	DNPM	Barragem Dique B3/MG e Barragem Itabiruçu/MG
		NATURATINS/TO	Barragem Taboca/TO
		SEDAM/RO	Barragem Capa Zero/RO
		ANA	Jaburu I/CE
RSB 2017	10	ANA	Barragem Cacimba Nova/PE e Barragem Barreiros/PE
		AESA/PB	Açude Imãos/PB
		ANM	Dique de Sela/MG
		DRH e SEMA /RS	Barragem do Capané/RS e Santa Bárbara/RS
		INEMA/BA	Afligidos/BA
		AGERH/ES	Santa Julia/ES
		IGARN/RN	Riacho do Meio/RN
RSB 2018	7	ANA	Jaburu I/CE
		AGERH / ES	Barragem Alto Santa Julia/ES
		IGARN/RN	Riacho do Meio/RN
		SEMA /RS	Barragem do Capané/RS, Santa Bárbara/RS e Assentamento Filho de Sepé/RS
		DAEE/SP	Barragem da Lagoa Rica Iaras/SP
RSB 2019	16	NATURATINS/TO	Barragem Balneário Iracema de Mito/TO e Morada do Sol/TO
		SEMAD/GO	Barragem B2/GO
		IGAM/MG	Peão/MG
		INEMA/BA	Rio da Caixa/BA
		ANA	Granjeiro/CE, Jaburu I/CE e Cacimba Nova/PE
		AGERH/ES	Duas Bocas/ES
		APAC/PE	Lagoa do Barro/PE
		IGARN/RN	Algodoeira São Miguel/RN e Riacho do Meio/RN
		SEMA /RS	Barragem do Capané/RS e Filhos de Sepé/RS
		DAEE/SP	Espelho 594/SP
		SEMAD/GO	Barragem 01/GO
INEA/RJ	Açude do Camorim/RJ		

Quanto à ocorrência de fenômenos de erosão interna, no RSB de 2019 nos relatos dos órgãos fiscalizadores, existem casos reportados no Rio de Janeiro e no Tocantins. No Rio de Janeiro, no Açude do Camorim, após uma vistoria *in loco*, foi constada a necessidade de reparos emergenciais em um talude danificado por erosões internas. No caso da barragem Balneário Iracema de Mito em Tocantins, houve um solapamento da crista e da

estrutura do vertedouro, acredita-se que o incidente ocorreu por conta de infiltrações de água em regiões próximas às manilhas do barramento (ANA, 2020).

O Anexo I do RSB de 2017 apresenta os acidentes e os incidentes ocorridos no período de avaliação do relatório. Cabe destacar que a ANA (2018) define acidente como evento que ocasiona em ruptura, enquanto incidentes são eventos de menor dano potencial. Dentre os acidentes, destacam-se o caso da Barragem Cacimba Nova e da Barragem Barreiros. Ambas romperam por fenômenos de erosão interna e eram formadas por um material homogêneo no maciço.

Em Cacimba Nova, segundo a ANA (2018), o enchimento do reservatório e consequentemente o aumento da carga de água provocou o carreamento de partículas finas do solo de fundação. Em Barreiros, o gradiente hidráulico gerado pelo enchimento induziu o fluxo de água pelo maciço do barramento, provocando o arraste partículas sólidas do aterro. No RSB de 2016, é relatado um incidente de uma barragem no estado de Rondônia no qual sofreu o processo de erosão interna e carreamento de partículas, em um canal de cerca de 5 cm de diâmetro.

2.2. PRINCIPAIS CAUSAS DE RUPTURA DE BARRAGENS DE TERRA

A Figura 2.3 apresenta os principais modos de falhas em barragens de terra, em que destacam-se os processos de infiltração (*seepage*), de erosão interna pelo corpo da barragem, pela fundação e ao redor de condutos que atravessam o corpo da barragem (*internal erosion*), deterioração do concreto (*concrete deterioration*), fissuramentos (*cracking*), rompimento de condutos no corpo da barragem (*leakage out of conduit*) e galgamento (*settlement in crest of dam*). Nota-se, dessa forma, que um barramento está susceptível a diferentes riscos, de origem biológica, química e física, nos quais podem ocasionar em falhas.

A ICOLD (2019) aponta o galgamento como o principal precursor de ruptura de grandes barragens, incluindo todos os tipos de barramentos, de terra, de enrocamento e de concreto. O fenômeno pode ocorrer por conta de insuficiência do vertedouro por falhas de projeto ou erros de execução, bloqueio de vertedouros por conta de detritos e recalques

da crista da barragem. A segunda maior causa são defeitos nas fundações, o que contempla instabilidade de taludes e recalques.

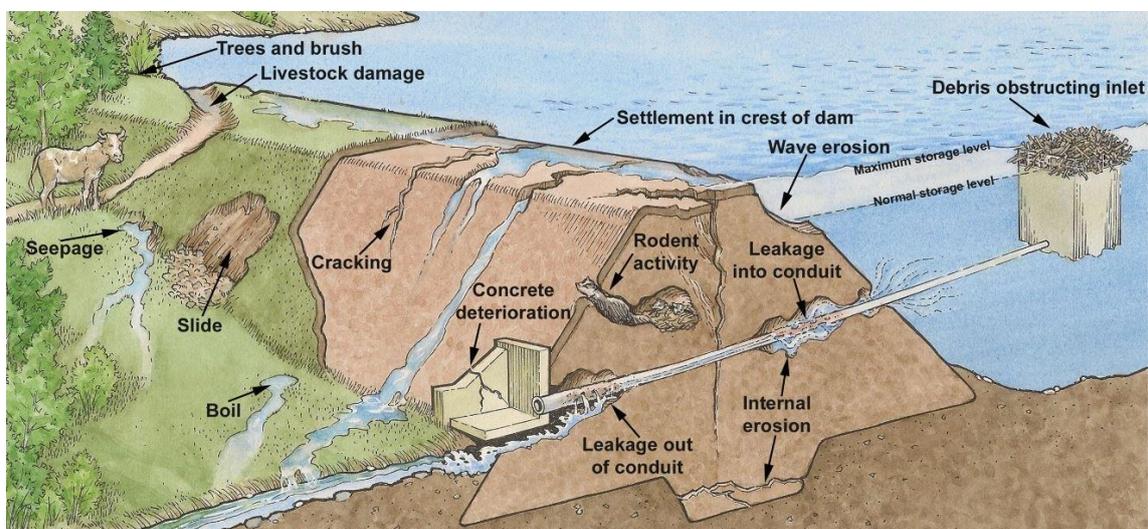


Figura 2.3 – Modos de Falhas em Barragens de Terra. Fonte: The Constructor (2019).

O *piping* é listado como a terceira maior causa, definido pela ICOLD (2019), como “erosão interna por infiltração”, no qual ocorre em volta de estruturas hidráulicas tais como vertedouros e condutos que atravessam o corpo do barramento, e através de fissuras causadas por recalques ou raízes de vegetação, irregularidades nas fundações e tocas de animais. Outras causas listadas também são falta de manutenção de elementos do barramento e uso de materiais inadequados na construção.

De acordo com Foster *et al.* (2000), a principal causa de ruptura de barragens de terra são os processos de erosão interna, cerca de 49% dos casos estudados. No estudo estatístico realizado com 11.192 barragens construídas entre 1800 e 1986, incluindo 136 rupturas, Foster *et al.* (2000) elencou três principais modos de falha: galgamento; *piping*, no qual considera os casos de *piping* através do corpo da barragem, pelo solo de fundação ou do corpo da barragem para o solo de fundação; e escorregamentos em taludes, no qual contempla escorregamentos de montante e jusante, efeito de sismos e liquefação.

De maneira análoga a Foster *et al.* (2000), Zhang *et al.* (2009) fez uma análise de 900 casos de rompimento de barragens pelo mundo, excluindo a China, sendo realizado um levantamento com as principais causas de rupturas e os fenômenos envolvidos em cada uma. Dentre a população de 900 barramentos, 593 correspondem às barragens de terra,

aproximadamente 66% do total. Zhang *et al.* (2009) constata que a principal causa das rupturas das barragens de terra está relacionada com problemas de qualidade das estruturas, com 42,5%, seguida do galgamento com 36,4%, desastres com 3,0% e gerenciamento não adequado com 1,3%. Os demais casos estão classificados como outros ou desconhecidos.

Dentro os problemas de qualidade nas estruturas, Zhang *et al.* (2009), assim como Foster *et al.* (2000), identifica os processos de erosão interna como majoritários, com 61,1% de ocorrência entre os casos analisados. Nas análises foram considerados três cenários, *piping* pelo corpo da barragem, pelo solo de fundação e ao redor do vertedouro. Escorregamentos no corpo ou na fundação estão associados a 18,3% das rupturas, enquanto falhas relacionadas com o vertedouro representam 6,3% dos casos analisados. Os 14,3% restantes são desconhecidos. Zhang *et al.* (2009) conclui que as principais causas de rompimento de barragens de terra estão associadas ao galgamento e ao *piping* pelo corpo do barramento ou solo de fundação.

O zoneamento da estrutura tem uma influência direta no início dos processos de erosão interna e instabilidade de taludes (Foster *et al.*, 2000). Nesse caso, barramentos com dispositivos ineficientes ou não existentes de controle de percolação e de poropressão, tais como drenos verticais e inclinados, drenos de pé e tapetes drenantes (MEIRELLES, 2019), tendem a apresentar maiores ocorrências de *piping*.

2.3. EROSÃO INTERNA

Por meio da análise de casos históricos, Foster *et al.* (2000) aponta que a ocorrência de *piping* pelo corpo da barragem é duas vezes maior que pelo solo de fundação e vinte vezes maior que pela fundação com início no corpo da barragem. Nos casos de rompimento por *piping* no corpo da barragem, o local de maior ocorrência do início do fenômeno é próximo ou em volta de condutos. Enquanto nos casos de *piping* pela fundação, o local de início mais frequente é em irregularidades no solo ou no maciço da fundação.

Foster *et al.* (2000) elencou os fatores que influenciam na ocorrência de erosão interna, conforme a seção da barragem. Rupturas em barragens de terra homogêneas estão associadas à ocorrência de *piping* ao redor de condutos, em materiais com compactação

não adequada ou materiais dispersivos utilizados no aterro. Zhang *et al.* (2009) afirma que a ocorrência do fenômeno depende de fatores tais como, o arranjo da barragem, a qualidade da construção da barragem e as condições geológicas. Destaca ainda que quaisquer pontos no corpo ou solo de fundação são locais potenciais.

Dos casos de barragens zonadas de terra e de enrocamento, observou-se que, na maioria dos casos, a barragem não possuía meios filtrantes e o material do núcleo foi carregado para jusante (Foster *et al.*, 2000). Zhang *et al.* (2009) aponta os mesmos riscos, nesse caso, que os das barragens de terra homogêneas, ou seja, potencial de ocorrência do *piping* através do solo de fundação, corpo da barragem ou ambos.

Segundo Foster *et al.* (2000), para barragens de terra com núcleo impermeável, a maioria dos casos de ruptura por *piping* esteve associada à presença de um núcleo fino no qual foi carregado para o aterro de jusante ou em rachaduras em condutos antigos que atravessavam o barramento. Zhang *et al.* (2009) conclui que para esse tipo de estrutura, a presença de um núcleo impermeável possibilita um controle maior da infiltração pelo corpo da barragem. Enfatiza, entretanto, que o núcleo também está suscetível ao *piping* por conta da ocorrência do surgimento de fraturas de origem hidráulica e irregularidades na interface de elementos com o corpo da barragem. Os pontos potenciais para a ocorrência do *piping* seriam o solo de fundações e a interface da fundação com o corpo da barragem.

Em barragens de terra com face de concreto, a ruptura por erosão interna também ocorreu ao redor de condutos, por conta problemas de conexões com a face de montante e ruptura de condutos no interior do corpo da estrutura (FOSTER *et al.*, 2000).

Foster *et al.* (2000) conclui que um número significativo de falhas por *piping* ocorrem durante o primeiro enchimento. Além disso, para os três cenários de *piping* analisados, pelo solo de fundação, pelo corpo da barragem e do corpo para a fundação, aproximadamente 67% das rupturas ocorrem entre os cinco primeiros anos de operação da barragem.

Os fenômenos de erosões internas, que ocorrem a partir da percolação de água em barragens de terra, são objeto de estudo da geotecnia desde Terzaghi (1931) e Casagrande (1937). Fell *et al.* (2015) afirma que os processos de erosão interna são importantes de

serem avaliados em barragens de pequeno, médio e grande porte e em diques, ainda, define que os processos de erosão interna podem ser analisados em quatro fases: 1) iniciação do processo erosivo interno; 2) continuação do processo erosivo interno; 3) evolução e formação de um conduto interno; e 4) formação do mecanismo de falha da barragem.

Ainda segundo Fell *et al.* (2015), a iniciação e a evolução da erosão interna em barragens de terra estão diretamente relacionadas com o tipo de solo do corpo e da fundação, além da existência de um gradiente hidráulico elevado. Os tipos de erosão interna podem ser classificados em quatro mecanismos principais, conforme Fell *et al.* (2015): erosão em fugas concentradas (*concentrated leaks*), sufusão (*suffusion*), erosão de contato (*contact erosion*) e erosão regressiva (*backward erosion*).

A erosão por fugas concentradas ou *concentrated leaks* acontece a partir da formação de um canal preferencial de percolação. A partir disso, forças de cisalhamento provocadas pelo escoamento nas paredes do canal, o material constituinte da barragem sofre erosão. Esse tipo de erosão pode ocorrer por distintos fatores, dentre recalques diferenciais durante e após a construção, trincas de origem hidráulica, colapso de materiais com grau de compactação não adequados ou ao redor de condutos que passam por dentro da barragem (FELL *et al.*, 2015).

A erosão por sufusão ou *suffusion* origina-se a partir da percolação de água através de solos não coesivos e instáveis internamente, com granulometria constituída de partículas grossas ou com granulometria descontínua. As partículas mais finas são transportadas pelos poros das partículas mais grossas e a partir disso, o material apresenta maior condutividade hidráulica, velocidades de percolação mais acentuadas e mudanças no gradiente hidráulico, fatores que contribuem para acelerar ainda mais o processo erosivo (FELL *et al.*, 2015).

O solo é considerado estável quando todas as partículas resistem aos esforços solicitados. Em solos instáveis, apenas as partículas maiores estão sujeitas aos esforços e com isso, as menores podem ser transportadas sem ocasionar em variações volumétricas (MOLINDER, 2016).

A erosão de contato ou *contact erosion* ocorre em regiões de transição de materiais de diferentes granulometrias e condutividades hidráulicas. Nessas zonas de contato, a percolação de água pelo solo de maior condutividade hidráulica, como areias e seixos, pode erodir solos de granulometria mais fina, como argilas e siltes. O escoamento é paralelo à zona de contato (FELL *et al.*, 2015). Isso ocorre em zonas de transição entre o solo do corpo da barragem e o material de drenos verticais, o que ressalta a importância da execução para evitar segregação.

A erosão regressiva ou *backward erosion* ocorre quando gradientes hidráulicos elevados no pé da barragem provocam a erosão de partículas, o que pode induzir a formação de um canal de jusante para montante. Caso a camada de solo superior ao canal erosivo não tenha coesão suficiente para sustentar a presença dos vazios causados pelo escoamento, a camada sofre um colapso interrompendo a formação do canal. No momento em que o canal se forma, a instabilidade da estrutura é acentuada a partir do carregamento de materiais para jusante. Um dos sinais mais comuns da ocorrência de erosão regressiva é a presença de materiais não coesivos fluando na área de jusante (FELL *et al.*, 2015). A Figura 2.4 ilustra os distintos tipos de erosões internas.

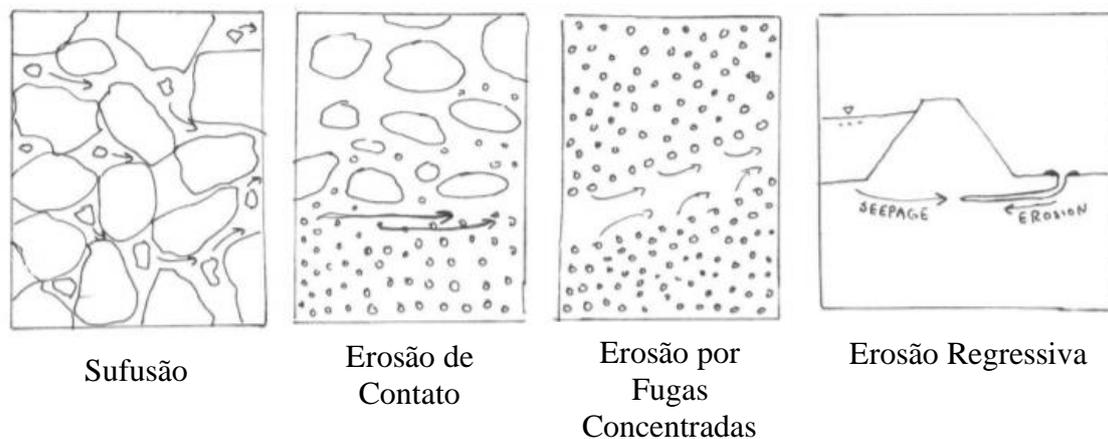


Figura 2.4 – Tipos de Erosão Interna em Barragens. Fonte: Adaptado de Molinder (2016).

Segundo aponta Fell *et al.* (2015), solos não coesivos, como areias e siltes, estão mais susceptíveis a sofrer erosões regressivas, erosões de contato e sufusão, a depender da distribuição granulométrica. Enquanto solos coesivos, como argilas, estão mais susceptíveis aos processos de erosão por fugas concentradas e erosões de contato. Sufusão e erosões regressivas podem ser observados também, no caso de solos coesivos, em

situações de gradientes hidráulicos locais elevados, o que normalmente não ocorre em barragens (FELL *et al.*, 2015).

O complexos fenômenos de erosões internas ou *piping* em barragens de terra estão associados com propriedades hidráulicas do solo, assim, devem ser revisados os conceitos de fluxo em meios porosos saturados e não saturados, conforme o item a seguir.

2.4. FLUXO EM MEIO POROSO

Segundo Darcy (1856), a vazão de infiltração através de um elemento de solo com área transversal A pode ser definida pela equação:

$$q = kiA \quad (2.1)$$

onde:

q = vazão (m/s);

k = condutividade hidráulica (m/s);

i = gradiente hidráulico (adimensional);

A = área transversal perpendicular à direção do escoamento (m²).

A Figura 2.5 apresenta a representação da Lei de Darcy, segundo Das (2019). Um elemento de solo em que se analisa a diferença de potencial Δh entre os pontos A e B, distanciados por um comprimento L .

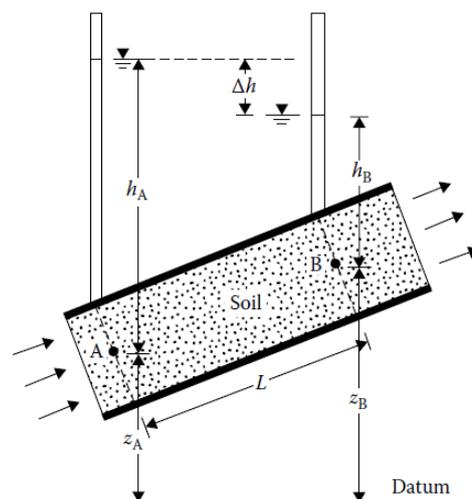


Figura 2.5 – Representação da Lei de Darcy. Fonte: Das (2019).

Segundo Das (2019), o gradiente hidráulico i pode ser calculado a partir da equação:

$$i = \frac{\Delta h}{L} \quad (2.2)$$

onde:

Δh = diferença de carga hidráulica entre os pontos A e B (m);

L = distância entre os pontos A e B (m).

A carga entre os pontos é obtida por meio do Teorema de Bernoulli (DAS, 2019) em que a parcela referente a velocidade é desprezada em meios porosos. A carga hidráulica em um ponto representativo A, portanto, é dada por:

$$A = z_A + h_A \quad (2.3)$$

onde:

z_A = carga de elevação (m);

h_A = carga de pressão (m).

Das (2019) define que, em estruturas hidráulicas, a ocorrência de *piping* está relacionada com o gradiente hidráulico. Em situações que o gradiente hidráulico aplicado é maior ou igual ao gradiente hidráulico crítico, há a ocorrência de processos erosivos caso o fluxo for ascendente. O gradiente hidráulico crítico i_{cr} é obtido a partir da seguinte equação:

$$i_{cr} = \frac{\gamma_{sub}}{\gamma_w} \quad (2.4)$$

onde:

γ_{sub} = peso específico submerso do solo (kN/m³);

γ_w = peso específico da água (kN/m³).

O peso específico submerso do solo pode ser calculado, conforme Das (2019), da seguinte forma:

$$\gamma_{sub} = \gamma_{sat} - \gamma_w \quad (2.5)$$

onde:

γ_{sat} = peso específico saturado do solo (kN/m³).

O gradiente hidráulico crítico pode ser calculado também a partir da densidade relativa dos grãos e do índice de vazios, conforme as equações a seguir:

$$\begin{aligned}\gamma_{sub} &= \gamma_{sat} - \gamma_w \\ \gamma_{sub} &= \frac{G_s \gamma_w + e \gamma_w}{1 + e} - \gamma_w \\ \gamma_{sub} &= \frac{(G_s - 1) \gamma_w}{1 + e}\end{aligned}\quad (2.6)$$

Logo,

$$i_{cr} = \frac{G_s - 1}{1 + e}\quad (2.7)$$

onde:

G_s = densidade relativa dos grãos (adimensional);

e = índice de vazios (adimensional).

Segundo Das (2019), para as combinações usuais dos dois índices físicos em solos, o gradiente crítico varia entre o intervalo de 0,85 a 1,1.

Harza (1935), a partir de pesquisas em segurança de estruturas hidráulicas, estabeleceu um fator de segurança contra *piping*, conforme a equação:

$$F_S = \frac{i_{cr}}{i_{saída}}\quad (2.8)$$

onde:

F_S = fator de segurança contra *piping* (adimensional);

$i_{saída}$ = gradiente hidráulico máximo de saída (adimensional).

O gradiente hidráulico máximo de saída pode ser obtido por meio de análises de percolação e redes de fluxo. Considera-se que o fator de segurança entre 3 e 4 é adequado para uma operação segura da estrutura para a ocorrência de *piping* (DAS, 2019). O gradiente hidráulico máximo de saída é determinado em simulações do SEEP/W do *Geostudio*, a partir das equações matemáticas detalhadas a seguir.

Ao considerar um processo de infiltração bidimensional em que a condutividade hidráulica nas direções é distinta, ou seja, em um meio anisotrópico, a modelagem de percolação no SEEP/W do *Geostudio* (2012), considera a seguinte relação, obtida a partir do fluxo de Darcy, do Teorema de Bernoulli e da Lei de Continuidade de Laplace:

$$\frac{\partial}{\partial x} \left(k_x \frac{\partial H}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y \frac{\partial H}{\partial y} \right) + Q = \frac{\partial \theta}{\partial t} \quad (2.9)$$

$$H = \frac{u_w}{\gamma_w} + z \quad (2.10)$$

onde:

H = carga hidráulica total de um ponto (m);

k_x = condutividade hidráulica na direção x (m/s);

k_y = condutividade hidráulica na direção y (m/s);

Q = fluxo de contorno aplicado (m³/s);

θ = teor de umidade volumétrica (adimensional);

t = tempo (s);

u_w = poropressão pela água (kPa);

γ_w = peso específico da água (kN/m³);

z = carga de elevação (m).

Ao analisar um volume elementar representativo em um tempo específico, a Equação (2.9) indica que a soma das taxas de variação de fluxo nas direções em x e y e o fluxo externo aplicado equivale à taxa de variação do teor de umidade volumétrica do solo. No caso do escoamento laminar, o termo de variação do teor de umidade volumétrica é nulo visto que o fluxo de entrada e o de saída é constante (SEEP/W GEOSTUDIO, 2012).

Fredlund e Morgenstern (1976) definem que variações no teor de umidade volumétrica estão associadas a variações das propriedades do solo e do estado de tensões. O estado de tensões nas condições saturada e não saturada, dessa forma, é descrito a partir de duas variáveis de estado: $\sigma - u_a$ e $u_a - u_w$, onde σ = tensão total; u_a = poropressão pelo ar; u_w = poropressão pela água (FREDLUND & MORGENSTERN, 1976).

Na modelagem transiente, o módulo SEEP/W do *Geostudio* (2012) considera que a tensão total nos elementos é constante, ou seja, não há carregamentos ou descarregamentos na simulação. Além disso, assume-se que a poropressão pelo ar é constante e igual às condições atmosféricas. Assim, a condição da $\sigma - u_a$ é constante. Enquanto na condição $u_a - u_w$, o termo u_a relativo à poropressão pelo ar é constante. Logo a variação do teor de umidade volumétrica depende exclusivamente da poropressão pela água u_w , conforme a seguinte equação:

$$\partial\theta = m_w \partial u_w \quad (2.11)$$

onde:

m_w = inclinação da curva de retenção do solo (adimensional).

Substituindo as Equações (2.10) e (2.11) na Equação (2.9), considerando que a derivada em relação ao tempo da carga relativa à elevação do ponto de análise é nula por ser constante, obtém-se a relação a seguir (SEEP/W GEOSTUDIO, 2012):

$$\frac{\partial}{\partial x} \left(k_x \frac{\partial H}{\partial x} \right) + \frac{\partial}{\partial y} \left(k_y \frac{\partial H}{\partial y} \right) + Q = m_w \gamma_w \frac{\partial H}{\partial t} \quad (2.12)$$

O fluxo não saturado e transiente em meios porosos, por efeito da gravidade, é definido pela equação de Richards, demonstrada por Richards (1931). Pelo balanço de massa de entrada e saída de um volume elementar representativo, a equação é demonstrada, considerando também a Lei de Darcy-Buckingham para meios não saturados. Na simulação de percolação de água do SEEP/W do *Geostudio*, a curva de retenção do solo pode ser estimada a partir da relação estabelecida por van Genuchten (1980), dada por:

$$\theta(\psi) = \theta_r + (\theta_s - \theta_r) \left[\frac{1}{(1 + (\alpha\psi)^n)^m} \right] \quad (2.13)$$

$$m = 1 - \frac{1}{n} \quad (2.14)$$

onde:

θ = teor de umidade volumétrica (adimensional);

θ_s = teor de umidade volumétrica saturada (adimensional);

θ_r = teor de umidade volumétrica residual (adimensional);

ψ = sucção matricial (cm);

α, n, m = parâmetros de ajuste.

A condutividade hidráulica não saturada pode ser estimada pela seguinte equação (VAN GENUCHTEN, 1980):

$$k(\psi) = k_s \left[\frac{1}{1 + (\alpha\psi)^n} \right]^m \left\{ 1 - \left[\frac{(\alpha\psi)^n}{1 + (\alpha\psi)^n} \right]^m \right\}^p \quad (2.15)$$

onde:

k = condutividade hidráulica não saturada (cm/s);

k_s = condutividade hidráulica saturada (cm/s);

α, n, m, p = parâmetros de ajuste.

2.5. ANÁLISE DE ESTABILIDADE

Uma usual análise realizada em taludes de barragens é a de estabilidade. Conforme será abordado no Capítulo 3, barragens hipotéticas serão simuladas e verificadas se atendem aos critérios mínimos esperados.

O SLOPE/W do *Geostudio* (2012) utiliza o método do equilíbrio limite para forças e momentos no cálculo do fator de segurança (FS). O FS é o fator que relaciona a resistência de cisalhamento do solo com os esforços atuantes. A resistência ao cisalhamento ao diminuir, provoca a movimentação da massa de solo ao longo de uma superfície de ruptura em um estado de equilíbrio limite (SLOPE/W GEOSTUDIO, 2012). Das & Sobhan (2014) definem a resistência ao cisalhamento de solos, considerando a análise de tensões efetivas, a partir da equação:

$$\tau = c' + (\sigma_n - u_w) \tan(\varphi') \quad (2.16)$$

onde:

τ = resistência ao cisalhamento (kPa);

c' = coesão efetiva (kPa);

σ = tensão total (kPa);

u_w = poropressão pela água (kPa);

φ' = ângulo de atrito efetivo (graus).

O fator de segurança, em termos da resistência de cisalhamento do solo, é dado pela equação a seguir (DAS & SOBHAN, 2014).

$$FS = \frac{\tau_{resistente}}{\tau_{atuante}} \quad (2.17)$$

onde:

FS = fator de segurança, em termos da resistência ao cisalhamento (adimensional);

$\tau_{resistente}$ = resistência média ao cisalhamento do solo (kPa);

$\tau_{atuante}$ = esforços médios de cisalhamento aplicados ao longo da superfície de ruptura (kPa).

Segundo Das & Sobhan (2014), um fator de segurança unitário ($FS = 1$) indica a ruptura iminente. Os autores aconselham um valor mínimo de 1,5 para projetos em análises de estabilidade de taludes.

Nas análises, o módulo SLOPE/W do *Geostudio* (2012) simula uma superfície de ruptura no corpo do talude e calcula o FS . Distintas superfícies são avaliadas com o objetivo de identificar àquelas com os menores fatores de segurança, ou seja, as superfícies críticas. De acordo com o método de estabilidade utilizado, a superfície é dividida em fatias verticais. O método do equilíbrio limite considera que o valor de FS é o mesmo em todas as fatias e ainda, que o FS é igual, em termos de ângulo de atrito efetivo e coesão, para os distintos solos contemplados pela superfície de ruptura (SLOPE/W GEOSTUDIO, 2012). A Figura 2.6 representa uma superfície de ruptura e as distintas forças atuantes em uma fatia de solo.

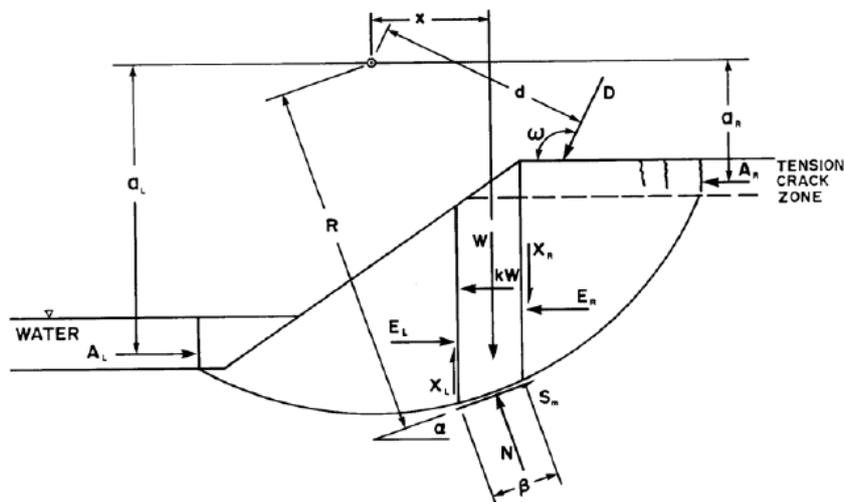


Figura 2.6 – Forças atuantes em uma fatia de solo. Fonte: *Geostudio* (2012).

Para a análise de estabilidade de taludes, foi utilizado o método de Morgenstern-Price (1965), que considera forças normais e de cisalhamento entre as fatias, além de satisfazer as condições de equilíbrio de forças e momentos. As forças entre as fatias estão relacionadas pela equação:

$$X = E\lambda f(x) \quad (2.18)$$

onde:

X = força de cisalhamento entre as fatias (kN);

E = força normal entre as fatias (kN);

λ = porcentagem da função utilizada (adimensional);

$f(x)$ = função arbitrária, em que nesse caso foi adotada a função *half-sine*.

Quando o valor de λ é nulo, tem-se o método de fatias simplificado de Bishop. Enquanto quando $f(x)$ é constante, ou seja, a relação entre as forças de cisalhamento e normais entre as fatias é a mesma em todas as fatias, tem-se o método de Spencer (SLOPE/W GEOSTUDIO, 2012).

O método do equilíbrio limite generalizado, considerado no módulo SLOPE/W do *Geostudio* (2012), calcula o fator de segurança com respeito aos momentos e às forças, de acordo com as equações a seguir:

$$F_m = \frac{\sum(c'\beta R + (N - u_w\beta)R \tan(\varphi'))}{\sum Wx - \sum Nf \pm \sum Dd} \quad (2.19)$$

$$F_f = \frac{\sum(c'\beta \cos(\alpha) + (N - u_w\beta) \tan(\varphi') \cos(\alpha))}{\sum N \operatorname{seno}(\alpha) - \sum D \cos(\omega)} \quad (2.20)$$

em que:

$$N = \frac{W + (X_R - X_L) - \frac{c'\beta \operatorname{seno}(\alpha) + u_w\beta \tan(\varphi') \operatorname{seno}(\alpha)}{F}}{\cos(\alpha) + \frac{\tan(\varphi') \operatorname{seno}(\alpha)}{F}} \quad (2.21)$$

onde:

W = peso da fatia de solo (kN);

N = força normal na base da fatia (kN);

D = carregamento concentrado (kN);

ω = ângulo de aplicação da carga concentrada (graus);

α = inclinação da base da fatia (graus);

β, R, x, f, d = parâmetros geométricos, conforme Figura 2.6;

F = fator de segurança quando a expressão de N é substituída nas equações as equações (2.19) ou (2.20);

X_R e X_L = forças de cisalhamento entre as fatias, respectivamente direita e esquerda.

A variável W , referente ao peso da fatia de solo, pode ser obtida por meio da seguinte equação, segundo Das & Sobhan (2014), considerando um comprimento unitário (1 metro) na direção perpendicular ao plano:

$$W = \gamma_{nat} LH \quad (2.22)$$

onde:

γ_{nat} = peso específico natural do solo (kN/m³);

L = largura da fatia de solo (m);

H = altura da fatia de solo (m).

O peso específico natural do solo γ_{nat} , segundo Das & Sobhan (2014), pode ser obtido por meio de parâmetros do solo. Quando a umidade é nula ($w = 0\%$), tem-se o peso específico do solo seco γ_d . Enquanto em condições saturadas ($w = 100\%$) tem-se o peso específico do solo saturado γ_{sat} . A equação do peso específico natural é dada por:

$$\gamma_{nat} = \frac{(1 + w)G_S\gamma_w}{1 + e} \quad (2.23)$$

onde:

w = umidade do solo (adimensional);

G_S = densidade relativa dos grãos (adimensional);

γ_w = peso específico da água (kN/m³);

e = índice de vazios (adimensional).

O módulo SLOPE/W do *Geostudio* (2012) efetua as análises de estabilidade de talude em três etapas, conforme resumido a seguir.

1. Na primeira iteração, as forças normais e de cisalhamento entre fatias são desprezadas. Calcula-se os fatores de segurança para as forças e os momentos. Esses valores são utilizados como dados de entrada na segunda etapa.
2. O parâmetro λ é zerado, ou seja, desconsidera-se as forças de cisalhamento entre fatias. Normalmente, em 4 a 6 iterações, as equações dos fatores de segurança para as forças e os momentos convergem.
3. Na última etapa, consideram-se distintos valores para o parâmetro λ são adotados e os fatores de segurança mais uma vez são calculados. O fator de segurança que satisfaz o equilíbrio de forças e de momentos é selecionado, o cruzamento entre as linhas do gráfico da Figura 2.7 no caso do método de Morgenstern-Price.

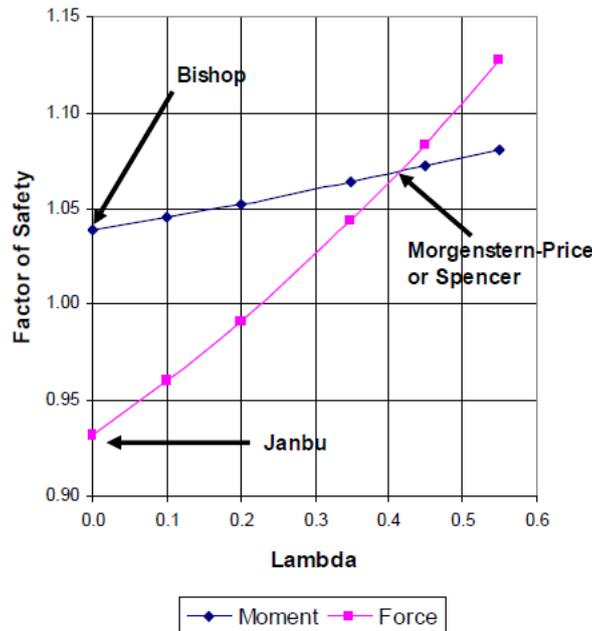


Figura 2.7 – Variação do parâmetro λ e fatores de segurança. Fonte: *Geostudio* (2012).

2.6. TESTES ESTATÍSTICOS

O teste t de Welch foi desenvolvido como uma generalização da hipótese estatística conhecida como teste t de Student, para o caso em que estão envolvidas diferentes variâncias de amostra (WELCH, 1947). Também conhecido como "teste t de variâncias distintas", o teste de Welch é um teste paramétrico de duas amostras utilizado para avaliar a hipótese de que a média de uma variável é igual nas duas amostras.

Tal como o teste t clássico, para duas amostras $(x_{11}, x_{12}, \dots, x_{1n})$ e $(x_{21}, x_{22}, \dots, x_{2n})$, assume-se que ambos seguem a distribuição normal $N(\mu_1, \sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$, onde μ é a média e σ^2 é a variância de cada amostra. Deseja-se testar a hipótese nula de que as duas médias são iguais ($\mu_1 = \mu_2$). Se a hipótese nula não for satisfeita, a hipótese alternativa é aceita ($\mu_1 \neq \mu_2$). Os parâmetros μ e σ são estimados pela média amostral \bar{x} e o desvio padrão s (WELCH, 1947).

Uma vez que a abordagem de Welch assume que os desvios não são idênticos, os Erros Padrões da Média – *Standard Errors of the Mean* (SEM) – são calculados separadamente para as amostras 1 e 2 da seguinte forma de acordo com Welch (1947):

$$SEM_1 = \frac{s_1}{\sqrt{n_1}} \quad (2.24)$$

$$SEM_2 = \frac{s_2}{\sqrt{n_2}} \quad (2.25)$$

onde:

s = desvio padrão da amostra;

n é o tamanho da amostra.

Após a definição do SEM de ambas as amostras, o Erro Padrão da Diferença das Médias – *Standard Error of the Difference of Means* (SEDM) – é determinado pela equação a seguir:

$$SEDM = \sqrt{SEM_1^2 + SEM_2^2} \quad (2.26)$$

Assim, o valor t é calculado pela equação (WELCH, 1947):

$$t = \frac{x_2 - x_1}{SEDM} \quad (2.27)$$

Com este procedimento, a distribuição pode ser aproximada por uma distribuição t com um número de graus de liberdade calculado a partir de cada desvio padrão da amostra e do número de elementos das amostras (DALGAARD, 2008).

O valor p é definido como a probabilidade de obter um valor numericamente grande ou maior do que o valor t observado. A hipótese nula da média da população igual é rejeitada quando $p\text{-valor} \leq \alpha$, onde α é o nível de significância. Assim, o valor p é considerado como uma medida de evidência contra a hipótese nula. Quanto menor o valor p , maior a evidência contra a hipótese nula (WASSERMAN, 2004). Normalmente, utiliza-se a escala apresentada na Tabela 2.2.

Tabela 2.2 – Escala de evidência do valor p . Fonte: Wasserman (2004).

valor p	Evidência
$< 0,01$	Evidência muito forte contra a hipótese nula
$0,01 - 0,05$	Evidência forte contra a hipótese nula
$0,05 - 0,10$	Evidência fraca contra a hipótese nula
$> 0,1$	Pouca ou nenhuma evidência contra a hipótese nula

Segundo Wasserman (2004), a correlação é uma medida do quão forte é a relação linear entre duas variáveis. A correlação cor pode ser calculada a partir da relação, considerando as variáveis x e y :

$$cor = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.28)$$

onde:

x_i e y_i = variáveis de análise;

\bar{x} e \bar{y} = médias das variáveis, respectivamente, x e y .

2.7. BORUTA E FLORESTAS ALEATÓRIAS

O aprendizado de máquina ou *machine learning* é um ramo da inteligência artificial que permite que simulações computacionais possam atingir resultados melhores por meio de treinamentos com exemplos (ERTEL, 2017). Em aplicações na engenharia, a partir de dados de entrada de ensaios de campo e de laboratório, os algoritmos de aprendizado de máquina podem ser calibrados para simular processos físicos. Assim, contribuem no reconhecimento de padrões para previsões e classificações, sendo implementados por linguagens de programação de computadores.

Nas barragens de terra homogêneas, diversos parâmetros geotécnicos já foram investigados quanto a sua relação com a ocorrência de *piping*, tais como, a altura da

barragem, a inclinação do talude a jusante, o ângulo de inclinação do dreno, o nível de água, o volume do reservatório, o gradiente hidráulico, a viscosidade do fluido, a coesão efetiva do solo, o ângulo de atrito interno, o grau de saturação do solo, o diâmetro efetivo das partículas, a porosidade, a massa específica e a condutividade hidráulica (XUE *et al.*, 2014; KOENDERS & SELLMEIJER, 1992).

Devido a quantidade de variáveis envolvidas nos estudos de *piping*, uma análise estatística dos parâmetros mais relevantes é essencial para reduzir a quantidade de variáveis de entrada e otimizar os modelos de previsão (PATHAK, 2018). Uma vez que modelos enxutos reduzem custos computacionais de simulações e permitem avaliações mais ágeis para tomada de decisão. Vários tipos de modelos foram propostos na literatura nas últimas décadas, a partir de simulações numéricas, como Sellmeijer (2006), Rotunno *et al.* (2018) e Chen *et al.* (2019).

Recentemente, a abordagem estatística por modelos de aprendizado de máquina (*machine learning*) tem-se revelado promissora na previsão de *piping*, uma vez que permite incluir análises probabilísticas de grandes volumes de dados (*big data*) e realizar o reconhecimento de padrões com base em estudos de casos históricos. Nesse caso, pode-se citar pesquisas como Tayfur *et al.* (2005), Yongbiao (2012), Xue *et al.* (2014) e Jamel (2018), em que os resultados dos algoritmos são comparados com ensaios de campo e de laboratório, e com modelos de simulação de barragens.

O algoritmo florestas aleatórias (*Random Forest*) é um algoritmo de classificação de conjuntos fundamentado no modelo de árvore de decisão (BREIMAN, 2001). O algoritmo extrai diferentes subconjuntos de dados a partir de uma base de dados de entrada. As árvores de decisão são componentes classificadores não enviesados e independentes, construídas pela formação dos subconjuntos, utilizando uma amostragem por *bootstrap* (KURSA & RUDNICKI, 2010; CHEN *et al.*, 2016).

O resultado da floresta, formada pelas árvores de decisão, é obtido pela média dos resultados individuais, no caso de regressões, ou pela decisão obtida mais vezes, no caso de classificações (KURSA & RUDNICKI, 2010). O que diferencia esta técnica de outras é que os componentes das árvores de decisão são formados aleatoriamente (POLAMURI, 2017).

O algoritmo Boruta é construído em torno das florestas aleatórias (*Random Forest*) e foi implementado em um pacote do *RStudio*, por Kursá & Rudnicki (2010). O nome do algoritmo origina-se da mitologia eslava, em que *Boruta* é o nome dado ao deus das florestas. Trata-se de um algoritmo de seleção de variáveis do tipo empacotador (*wrapper*), em que se utiliza subconjuntos ou pacotes de um banco de dados para o treinamento do modelo. A partir de um resultado preliminar, o algoritmo decide em adicionar ou remover variáveis do subconjunto (PATHAK, 2018).

Ao tratar de seleção de variáveis, segundo Kursá & Rudnicki (2010), nas florestas aleatórias a medida de importância de uma variável em relação a outra é obtida pelo do cálculo da perda de acurácia de classificação, que ocorre por conta da permutação aleatória dos valores das variáveis entre as observações. A variável dependente, no caso desse estudo a ocorrência de *piping*, é avaliada a partir do cálculo da importância das variáveis independentes, demais variáveis do banco de dados.

É importante destacar que a acurácia ou exatidão está relacionada com a diferença entre um valor obtido com o valor esperado, portanto, um valor acurado é aquele próximo ao valor real. Enquanto a precisão, está relacionada com a variação entre os valores obtidos, valores precisos são aqueles em que há pouca variação. A Figura 2.8 apresenta as diferenças entre os conceitos (GM METROLOGIA, 2020).

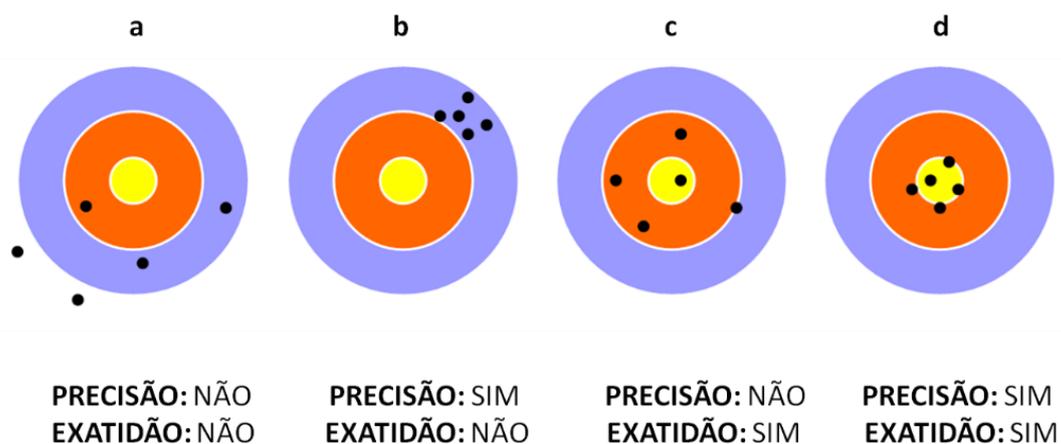


Figura 2.8 – Exemplos de precisão e exatidão. Fonte: GM Metrologia (2020).

Cada árvore de decisão apresenta um resultado de importância da variável, a partir do respectivo subconjunto utilizado. Assim, a média e o desvio padrão dos erros são calculados, e determina-se, então o valor de Z , conforme a equação a seguir:

$$Z = \frac{x - \mu}{\sigma} \quad (2.29)$$

onde:

x = erro obtido na classificação;

μ = média simples da população de erros da classificação;

σ = desvio padrão da população de erros da classificação.

O parâmetro Z obtido da floresta aleatória, contudo, não pode ser utilizado como parâmetro de significância das variáveis por não apresentar uma distribuição normal, média nula e desvio padrão unitário $N(0,1)$. No algoritmo Boruta, utiliza-se ainda sim o parâmetro Z , por esse considerar flutuações da média dos erros de acurácia da simulação (KURSA & RUDNICKI, 2010).

Visto que o parâmetro Z , por si só, não pode ser utilizado para avaliar a significância estatística, é necessário um parâmetro de referência externo para que o algoritmo avalie se uma variável é importante ou não em relação a outra. Para isso, o algoritmo Boruta cria as variáveis sombra (*shadow attributes*).

As variáveis sombra são obtidas por meio da cópia das observações do banco de dados para uma determinada variável e o embaralhamento aleatório dos valores entre as linhas. Com isso, todas as correlações aleatórias entre a variável e a variável de análise são perdidas. A classificação, pela floresta aleatória, no algoritmo Boruta é realizada nesse banco de dados estendido com cópias embaralhadas das variáveis (KURSA & RUDNICKI, 2010). A importância calculada, pelo parâmetro Z , é comparada entre as colunas originais e as duplicadas.

Em suma, Boruta é um algoritmo de seleção de variáveis que identifica aquelas mais importantes para uma variável de análise. Com base nas florestas aleatórias, o algoritmo calcula o grau de importância das variáveis em múltiplas iterações com amostras

aleatórias, o que reduz o impacto de flutuações e correlações redundantes (PATHAK, 2018; KURSA & RUDNICKI, 2010).

A seguir, será apresentado um passo a passo da aplicação do algoritmo Boruta na avaliação de variáveis relevantes para o *piping*, conforme Kursá & Rudnicki (2010) e Pathak (2018).

1. A partir de um banco de dados de variáveis que podem afetar a ocorrência de *piping*, o algoritmo analisa a relação individual de cada variável com o fenômeno. Considere a variável nível de água da barragem e a variável binária que indica a ocorrência de *piping* quando for igual a 1. A partir disto, é criada uma coluna duplicada da variável de nível de água, em que os valores são embaralhados para eliminar correlações, conforme a Figura 2.9.

Nível de Água (m)	Ocorrência de <i>Piping</i>		Nível de Água (m)	Ocorrência de <i>Piping</i>	Nível de Água ' (m)
40	1	➔	40	1	70
90	0		90	0	35
70	1		70	1	40
35	0		35	0	100
100	0		100	0	90

Figura 2.9 – Passo 1 do Algoritmo Boruta. Fonte: Kursá & Rudnicki (2010).

2. Neste passo, aplica-se o algoritmo *Random Forest* e determina-se o parâmetro Z da variável nível de água, quanto maior o seu valor, maior a importância da variável para a ocorrência de *piping*.
3. Verificação do critério $Z_{variável} > Z_{sombra}$. A variável nível de água é considerada importante se e somente se o valor do parâmetro Z no banco de dados original é maior que a da coluna duplicada. Caso for, é registrada uma pontuação (hit) para a variável. Após isso, a coluna duplicada é eliminada do banco de dados.
4. Os passos anteriores são realizados novamente até que todas as variáveis do banco de dados estejam com o parâmetro Z calculado. Para a variável nível de água, a partir do número de iterações, são elencadas diferentes colunas duplicadas

embaralhadas com o intuito de obter os máximos valores possíveis para Z nas variáveis sombra e comparar com o resultado da coluna original. Assim, para cada iteração verifica-se o critério do passo 3 e são os registrados os *hits*.

5. As variáveis com mais *hits* são as mais importantes, utiliza-se uma distribuição binomial na comparação com as variáveis sombra. Identifica-se, dessa forma, o maior valor de Z como sendo a importância da variável independente (nível de água) para a variável dependente (ocorrência de *piping*).

2.8. REDES NEURAIS ARTIFICIAIS (RNA)

O conceito de RNA foi introduzido por McCulloch & Pitts (1943), no artigo *A Logical Calculus of the Ideas Immanent in Nervous Activity*. Foi a primeira estrutura proposta para uma rede neural artificial, um modelo computacional simples que simula como os neurônios biológicos funcionam e reagem à problemas complexos utilizando lógica proposicional (Géron, 2017). Segundo Tayfur *et al.*(2005), uma RNA é capaz de identificar relações a partir de treinamentos com dados de entrada, isso permite a resolução de problemas complexos envolvendo reconhecimento de padrões, modelagens não lineares, classificações e associações.

Géron (2017) destaca que a evolução computacional dos anos 1990 permitiu o treinamento de redes neurais complexas em um intervalo razoável de tempo, o que possibilita a resolução de problemas complexos. Além disso, o autor relata que, como se acreditava inicialmente, simulações realizadas com RNA não tendem mais a se restringir aos máximos locais por conta da otimização dos algoritmos de treinamento. Nessa situação, Géron (2017) cita que mesmo se essa situação ocorrer, o máximo local encontrado é relativamente próximo ao ponto de máximo global.

McCulloch & Pitts (1943) propuseram um modelo simples para o neurônio biológico, com um ou mais elementos binários de entrada e um elemento binário de saída. Por meio da lógica proposicional, o sistema é capaz de computar problemas complexos, utilizando operadores tais como, *and* (e), *or* (ou) e *not* (não). Trata-se de um modelo de um neurônio artificial, idealizado com inspiração biológica no que tange conexão entre neurônios e estruturação em camadas.

A Figura 2.10 apresenta um neurônio biológico com os seus principais componentes. A estrutura é formada por um corpo celular com o núcleo com ramificações denominadas de dendritos, que são responsáveis por transmitir os impulsos para a célula. Pelo axônio, o neurônio conduz o impulso nervoso para o próximo neurônio da rede, sendo sinapse o nome dado para a zona de comunicação entre neurônios (MUNDO EDUCAÇÃO, 2019). Conforme Géron (2017), a organização dos neurônios no córtex cerebral de animais se dá em camadas, em que um neurônio se comunica com o semelhante da próxima camada pelo axônio. A partir de um número determinado de impulsos, um neurônio pode emitir o próprio impulso e assim transmitir o sinal para os demais componentes da rede.

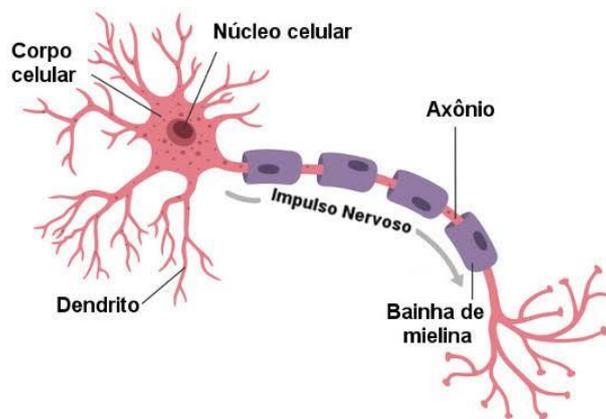


Figura 2.10 – Representação de um neurônio biológico. Fonte: Mundo Educação (2019).

Uma das estruturas mais simples de RNA é o denominado *Perceptron*, proposta por Rosenblatt (1957), conforme apresentado na Figura 2.11. Conforme Géron (2017), os dados de entrada nesse caso não são mais binários, mas sim números, organizados em uma estrutura denominada de *linear threshold unit* (LTU). Cada entrada (x) é associada a um peso correspondente (w), assim, a partir da multiplicação de cada *input* pelo seu peso, somando esses termos obtém-se a soma ponderada (z). Essa operação pode ser representada matricialmente, de acordo com a equação:

$$\mathbf{z} = \mathbf{w}^T \cdot \mathbf{x} = w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}_{n \times 1} \cdot [x_1 \quad x_2 \quad \dots \quad x_n]_{1 \times n} \quad (2.30)$$

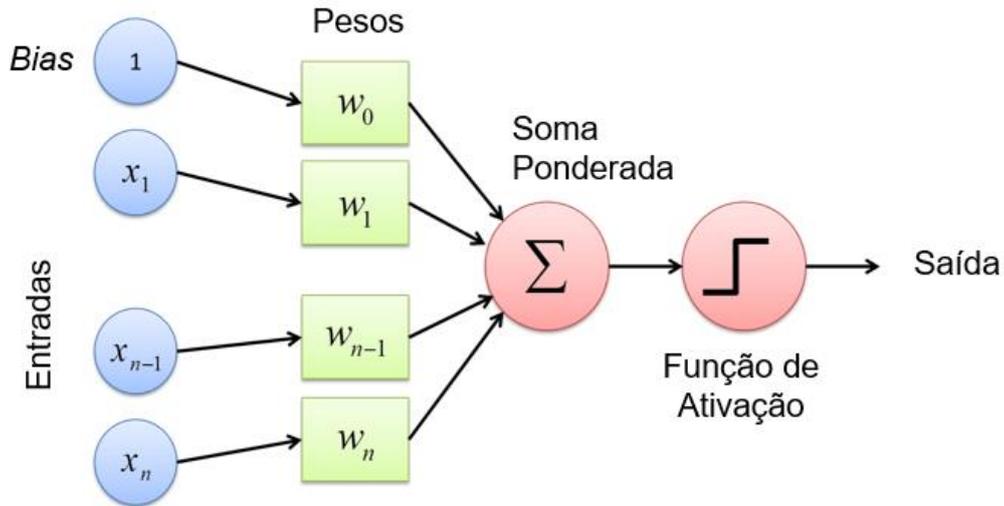


Figura 2.11 – Estrutura do *Perceptron*. Fonte: Adaptado de Rosenblatt (1957).

Por sua vez, essa soma ponderada (z) é inserida como dado de entrada para uma função, denominada função de ativação, cujo objetivo é transformar o resultado em um intervalo definido. Comumente são adotadas funções que são fixas entre intervalos como $(0,1)$ e $(-1,1)$, por exemplo, Géron (2017) cita a função tangente hiperbólica. Trata-se de uma função que o gráfico apresenta um formato em S , sendo contínua e diferenciável. Para quaisquer valores de *input*, o *output* da função tangente hiperbólica está situado entre o intervalo $(-1,1)$. Géron (2017) afirma que isso contribui para a convergência do treinamento dos pesos, visto que a função normaliza os resultados em torno de zero.

Tayfur *et al.* (2005) utiliza a função sigmoide, que também contínua e diferenciável. A função apresenta um comportamento semelhante à função tangente hiperbólica, porém, restringe os valores de *output* entre $(0,1)$. A função sigmoide é dada por:

$$f(x) = \frac{1}{1 + \exp(-x)} \quad (2.31)$$

O neurônio indicado como *constant*, na Figura 2.11, representa o termo *bias*. Conforme Géron (2017), esse termo também é referido como termo de intercepto, comparado ao termo independente de equações lineares. De acordo com GeeksforGeeks (2019), o *bias* induz o atraso na função de ativação da RNA. Nesse caso, assume-se um valor inicial constante para o *bias neuron* e durante o treinamento, esse auxilia no melhor ajuste dos *outputs* da RNA conforme os dados de entrada, deslocando a função de ativação para a

esquerda ou para a direita. O *bias* é somado à soma ponderada (z) que é o parâmetro de entrada função de ativação.

A Figura 2.12 traz de maneira gráfica as funções de ativação mencionadas, função tangente hiperbólica e a função sigmoide. A Figura 2.13 apresenta a influência da variação do termo *bias* na função de ativação tangente hiperbólica. Nota-se que à medida que o valor subtraído da soma de pesos aumenta, a função desloca-se no sentido positivo do eixo horizontal.

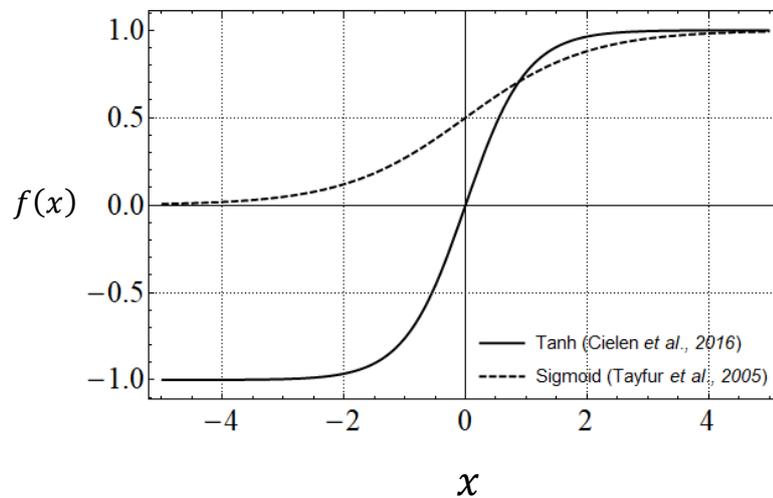


Figura 2.12 – Exemplos de *step functions*.

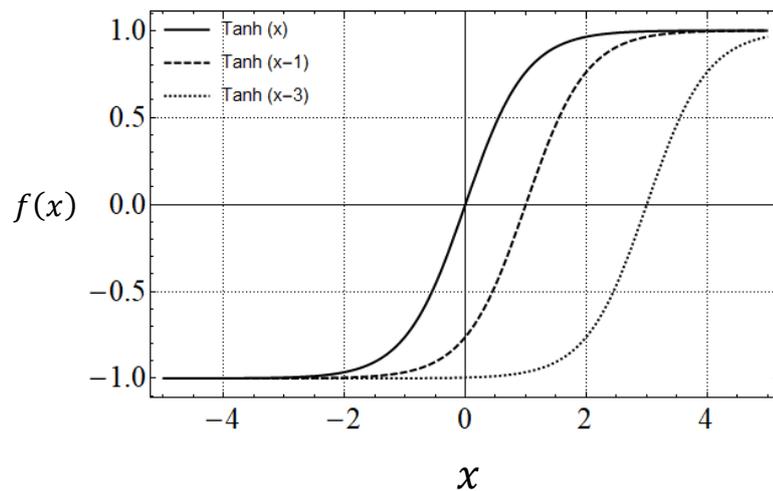


Figura 2.13 – Funções de ativação e a variação do termo *bias*.

Para o *Perceptron*, conforme o que foi explicitado, a equação aplicada pela rede neural artificial (RNA) é a seguinte:

$$o(x) = f(z) = f\left(b + \sum_{i=1}^n w_i x_i\right) = f(b + \mathbf{w}^T \mathbf{x}) \quad (2.32)$$

onde:

o = *output* ou resultado da RNA;

f = função de ativação;

z = soma ponderada;

\mathbf{x} = matriz dos valores das variáveis de entrada;

\mathbf{w}^T = matriz transposta dos pesos;

b = parâmetro *bias* ou viés.

Ao adicionar mais neurônios, no qual cada um possui um valor próprio da soma ponderada (z), tem-se a estrutura da Figura 2.14, denominada *multilayer perceptron* (MLP) ou *perceptron* de múltiplas camadas. Nesse exemplo de estrutura, existem três dados de entrada x_n , três neurônios na camada intermediária com as suas respectivas funções de ativação a_n , camada essa comumente denominada de *hidden layer*, e um neurônio de saída com a função de ativação a_1 .

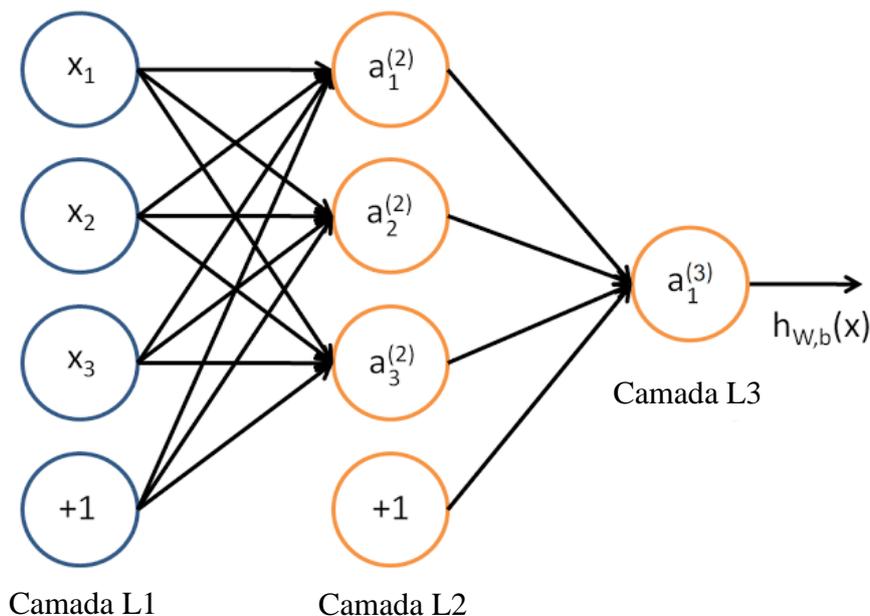


Figura 2.14 – Estrutura de uma RNA com o parâmetro *bias*. Fonte: Adaptado de Towards Data Science (2019).

Os pesos (w) são representados pelas linhas e o treinamento da rede visa otimizar os seus respectivos valores, com o intuito de diminuir a função erro que relaciona o *output* $h_{w,b}(x)$ da RNA com o dado de treinamento. Nota-se ainda a presença dos termos *bias* na camada intermediária e na de saída.

No caso de uma RNA do tipo MPL ou *multilayer perceptron*, com a inclusão de uma camada oculta com J neurônios, a função aplicada nas equações a seguir, sem e com a notação matricial (GUNTHER & FRITSCH, 2010).

$$o(x) = f \left(b + \sum_{j=1}^J w_j f \left(b_j + \sum_{i=1}^n w_{ij} x_i \right) \right) \quad (2.33)$$

$$o(x) = f \left(b + \sum_{j=1}^J w_j f(b_j + \mathbf{w}_j^T \mathbf{x}) \right) \quad (2.34)$$

onde:

o = *output* ou resultado da RNA;

f = função de ativação;

z = soma ponderada;

\mathbf{x} = matriz dos valores das variáveis de entrada;

\mathbf{w}_j^T = matriz transposta dos pesos do neurônio j da camada oculta;

b = parâmetro *bias* ou viés;

b_j = parâmetro *bias* ou viés do neurônio j da camada oculta.

Existem distintas técnicas para a otimização dos valores dos pesos (w), conforme Géron (2017). Uma delas é a técnica de retropropagação ou *backpropagation* (BP), cujo objetivo é otimizar os pesos para que a RNA possa aprender e mapear corretamente entradas e saídas. Segundo Gunther & Fritsch (2010), é definida uma função erro (E):

$$E = \frac{1}{2} \sum_{l=1}^L \sum_{h=1}^H (o_{lh} - y_{lh})^2 \quad (2.35)$$

onde:

o = *output* da RNA;

y = valor alvo do *output* da RNA, ou seja, o dado de treinamento;

L = número de padrões de treinamento;

H = número de neurônios de *output*.

A RNA é simulada para obter o *output*, com base nos dados de entrada. A partir disso, computa-se o erro associado, conforme a Equação (2.35). O erro é então propagado para a camada de *input*, em que se modifica os pesos. Conforme Gunther & Fritsch (2010), para modificar os pesos conforme o resultado da função erro, utiliza-se o método do gradiente descendente, dado por:

$$w_k^{(t+1)} = w_k^{(t)} - \delta \left(\frac{\partial E}{\partial w_k^{(t)}} \right) \quad (2.36)$$

onde:

t = número da iteração;

k = número do peso;

δ = taxa de aprendizado, normalmente adotada com o menor valor possível para evitar que o treinamento fique restrito a mínimos locais.

A otimização do termo *bias* é realizada com o mesmo método. Com isso, a RNA é simulada a partir dos dados de entrada e o algoritmo de BP otimiza os pesos e os termos *bias*, com o intuito de minimizar o erro calculado entre o *output* e o dado de treinamento. Segundo Gunther & Fritsch (2010), o algoritmo de BP modifica os pesos com o intuito de encontrar o valor mínimo da função erro. O termo $\partial E / \partial w$ é calculado para que se possa obter a raiz da função erro. A partir da taxa de aprendizado adotada, o algoritmo adiciona ou subtrai um incremento aos pesos, conforme a Figura 2.15.

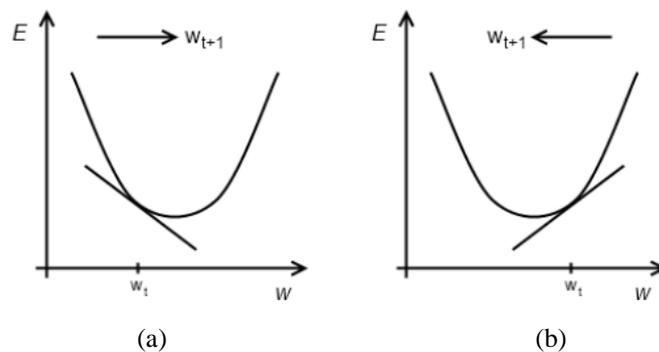


Figura 2.15 – Algoritmo de *Backpropagation* e a função erro. Fonte: Gunther & Fritsch (2010).

Dessa forma, quando a derivada parcial do erro em relação ao peso é negativa, adiciona-se um incremento para que o resultado se aproxime do mínimo local, conforme a Figura 2.15 (a). Em situações em que a derivada parcial for positiva, subtrai-se um incremento do peso, para também aproximar do mínimo local, conforme a Figura 2.15 (b) (GUNTHER & FRITSCH, 2010).

Em simulações de redes neurais artificiais no *Rstudio*, é possível também utilizar o algoritmo de retropropagação resiliente ou *resilient backpropagation* (RBP). A diferença do algoritmo BP tradicional é que no RBP cada peso da rede neural tem uma taxa de aprendizado distinta (GUNTHER & FRITSCH, 2010). Dessa forma, a equação de otimização dos pesos modifica-se para:

$$w_k^{(t+1)} = w_k^{(t)} - \delta_k^{(t)} \text{ sinal} \left(\frac{\partial E}{\partial w_k^{(t)}} \right) \quad (2.37)$$

onde:

t = número da iteração;

k = número do peso;

δ = taxa de aprendizado.

No RBP, a taxa de aprendizado aumenta caso a derivada parcial do erro em relação ao peso mantenha-se com o mesmo sinal. Se houver a troca de sinal, a taxa de aprendizado diminui, visto que o mínimo local foi perdido com o incremento no peso. Assim, esses procedimentos são seguidos para que a convergência para o mínimo local da função erro seja atingida em menos iterações. A técnica de *weight backtracking* é utilizada também que consiste em desfazer a iteração anterior e diminuir o valor de incremento do peso na próxima iteração, o que também contribui para evitar que o algoritmo perca os mínimos locais (GUNTHER & FRITSCH, 2010).

Sendo assim, com a definição da técnica de redes neurais artificiais (RNA), procede-se para um breve relato de trabalhos anteriores envolvendo *piping* e os algoritmos de aprendizado de máquina.

2.9. RNA APLICADAS PARA ANÁLISES DE PIPING

Tayfur *et al.* (2005) utilizou uma RNA para estimar o nível de piezômetros em uma seção de uma barragem na Polônia, na ocorrência de *piping* pelo corpo da barragem. A RNA utilizada possuía três camadas, com três neurônios na camada de *input*, cinco na camada intermediária ou *hidden*, e um neurônio na camada de *output*. As variáveis de entrada utilizadas foram o nível de água a montante do barramento, o nível de água a jusante do barramento e a identificação do piezômetro.

Para o treinamento da RNA, foram utilizados valores entre 0,02 e 0,9 para os pesos e -1 para o *bias*. O algoritmo de otimização dos parâmetros foi o *backpropagation* (BP), abordado no item 2.8. Tayfur *et al.* (2005) também utilizou um modelo pelo método dos elementos finitos (MEF) para a previsão dos dados. Para o piezômetro de número 39, a Figura 2.16 apresenta o treinamento comparando os dados medidos com a simulação pela RNA (*artificial neural network* – ANN) e pelo MEF (*finite element method* – FEM).

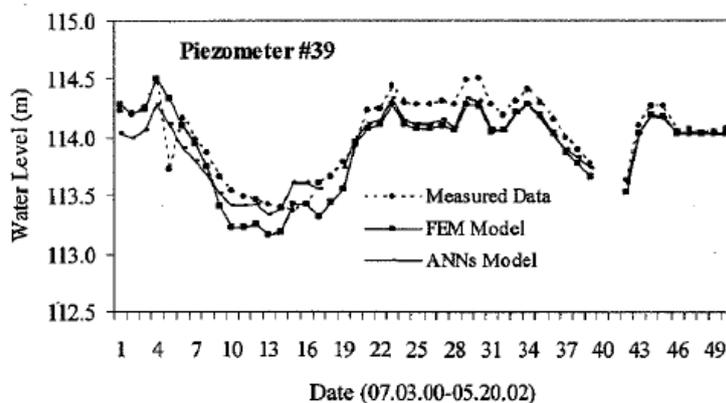


Figura 2.16 – Resultados do treinamento. Fonte: Tayfur *et al.* (2005).

Tayfur *et al.* (2005) conclui que o modelo por MEF apresenta uma capacidade de prever a variação de parâmetros no espaço e no tempo, enquanto a RNA avalia a sua variação apenas no tempo. Em alguns piezômetros analisados, a RNA mostrou-se com mais acurácia do que o MEF para prever a variação temporal dos níveis de água. Conforme Tayfur *et al.* (2005), o caminho de infiltração pode ser facilmente determinado a partir de análises dos comportamentos de variáveis em termos do espaço e do tempo. A Figura 2.17 apresenta a comparação dos resultados previstos com os medidos.

Além disso, Tayfur *et al.* (2005) afirma que a RNA pode estimar a infiltração utilizando apenas os níveis de água de montante e de jusante do barramento, o que exige menos aquisição de dados de campo quando comparado ao MEF. A RNA, diferentemente do MEF, é específica para o local, no qual exige uma série histórica de dados para que a previsão seja embasada com um coeficiente de determinação R^2 satisfatório.

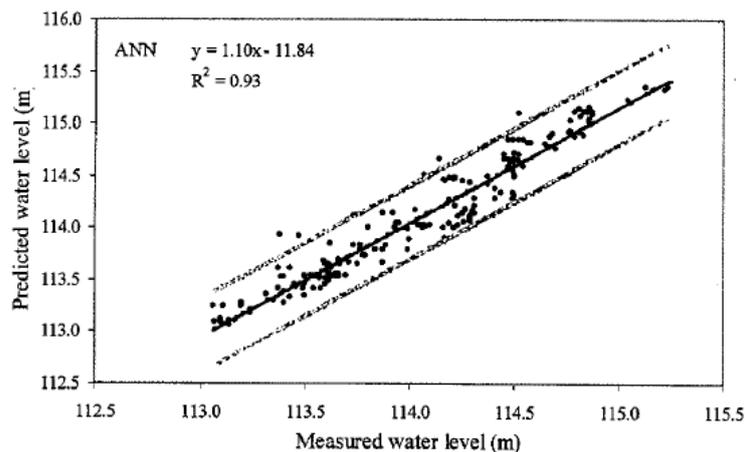


Figura 2.17 – Resultados da simulação. Fonte: Tayfur *et al.* (2005).

Xue *et al.* (2014) estimou o potencial de *piping* em barragens de terra utilizando redes neurais generalizadas (*generalized neural networks - GRNN*). O autor subdivide o fenômeno de *piping* em quatro etapas: iniciação, aparecimento de trincas e fissuras que podem tornar-se caminho para a infiltração; continuação, falta de dispositivos no barramento para prevenir que ocorra erosão de partículas e carreamento para jusante; progressão, o conduto alarga ocasionando no aumento de velocidade de infiltração e erosão; e ruptura, ocorrência do alargamento significativo do conduto no qual pode acarretar na instabilidade de taludes.

De modo paralelo à GRNN, Xue *et al.* (2014) utilizou o algoritmo de *backpropagation* (BP) para otimizar os pesos e os *bias* dos neurônios e comparou os resultados obtidos por uma rede neural convencional. Como entrada da rede neural, Xue *et al.* (2014) selecionou nove parâmetros: altura da barragem, nível de água, inclinação a jusante da barragem, ângulo efetivo de atrito interno, coesão efetiva, peso do solo saturado, condutividade hidráulica do solo, diâmetro máximo da partícula e inclinação do filtro de proteção quanto à infiltração. O *output* da rede neural é binário, em que o resultado indicado como unitário

significa que ocorreu o fenômeno de *piping*, enquanto o valor nulo representa a não ocorrência.

Xue *et al.* (2014) conclui que a GRNN é um método eficiente para a estimar a ocorrência de *piping* com uma alta acurácia, sendo um modelo com grande confiabilidade, mais satisfatório que o método de *backpropagation*. Além disso, o autor afirma que o desempenho da GRNN é dependente do parâmetros de suavização, uma otimização na sua determinação minimiza erros de classificação e conseqüentemente erros finais da rede neural. Os principais resultados obtidos são apresentados na Figura 2.18.

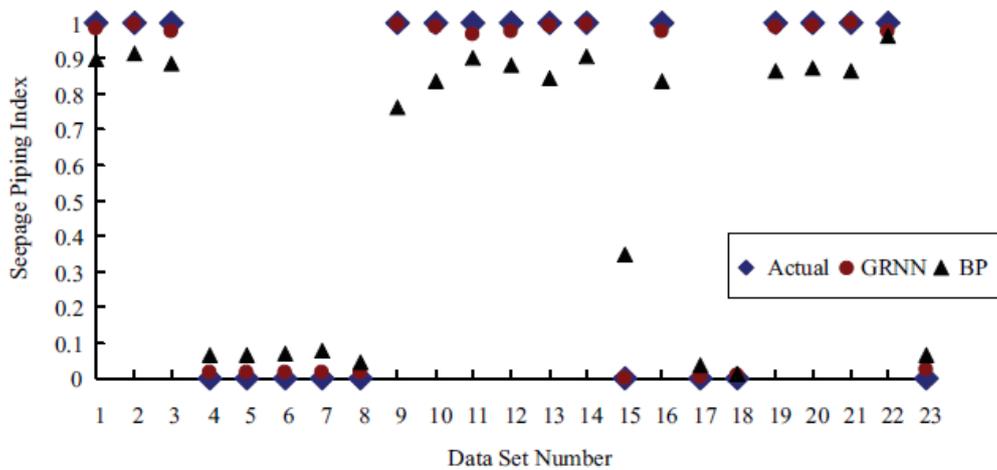


Figura 2.18 – Resultados da simulação com GRNN. Fonte: Xue *et al.* (2014).

Dessa forma, Tayfur *et al.* (2005) e Xue *et al.* (2014) utilizaram a técnica de rede neural artificial (RNA) para estimar o comportamento de dados de monitoramento e a ocorrência de erosões internas em barragens de terra, com resultados positivos que demonstram a acurácia das previsões.

3. METODOLOGIA

A metodologia geral proposta para o desenvolvimento deste projeto final está apresentada na Figura 3.1, dividida em duas etapas. A Figura 3.2 e a Figura 3.3 especificam as subetapas da metodologia geral. Nos itens a seguir, são detalhadas cada uma das subetapas, indicando os resultados esperados.

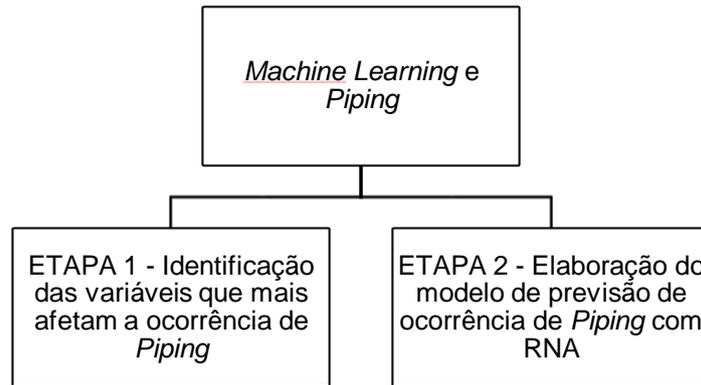


Figura 3.1 – Metodologia Geral do Projeto Final.

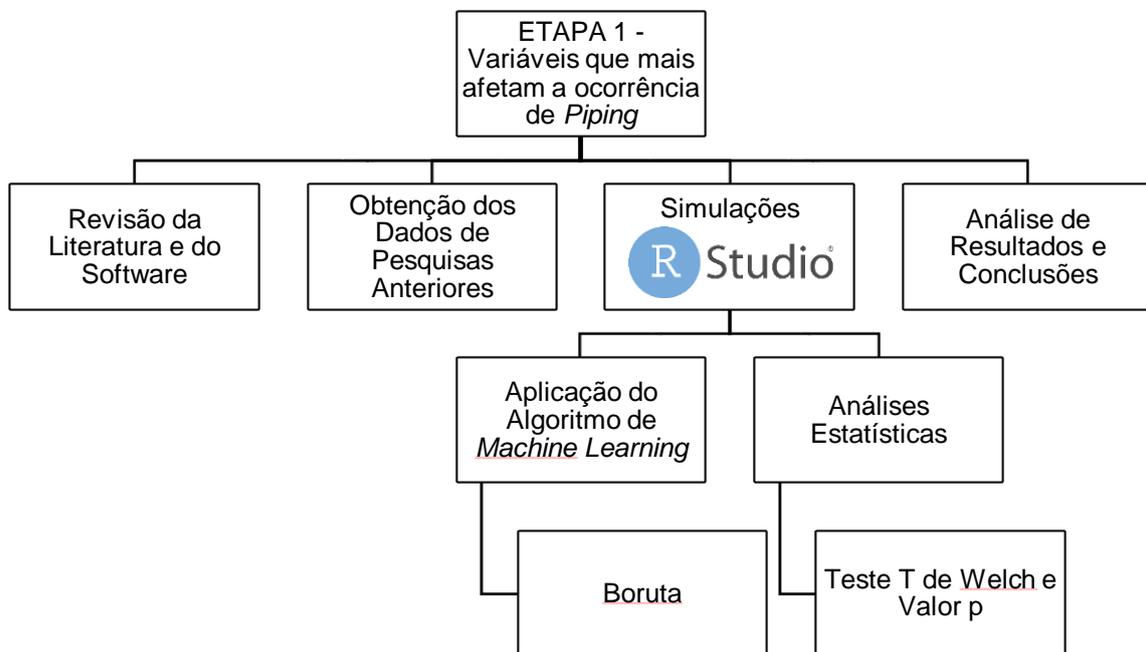


Figura 3.2 – Metodologia Específica da Etapa 1.

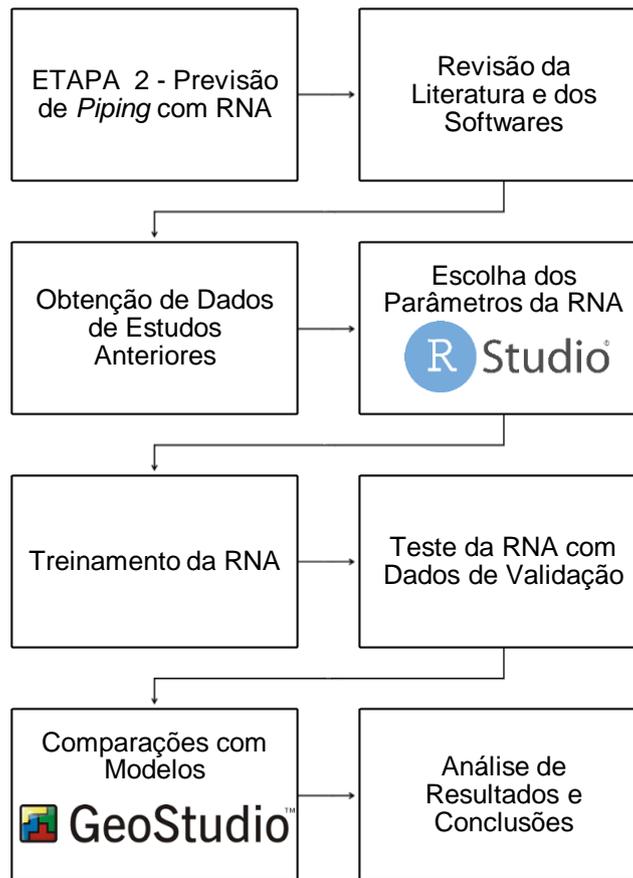


Figura 3.3 – Metodologia Específica da Etapa 2.

3.1. REVISÃO DA LITERATURA E DOS SOFTWARES

O primeiro passo da metodologia trata da revisão da literatura técnica, em que foram pesquisados artigos a respeito do tema estudado em bases reconhecidas, tais como *Science Direct*, *Google Scholar*, *SciELO*, *ASCE Library* e *Research Gate*. Isso permitiu que fosse realizado um levantamento do estado da arte de estudos sobre *piping* e do uso de técnicas de aprendizado de máquinas relacionados com a modelagem de fenômenos de erosão interna.

Uma pesquisa foi feita sobre usos do algoritmo Boruta e simulações de RNA no *RStudio*, exemplos de aplicações de regressões e mensuração de resíduos. O *Rstudio* é um *software* livre que permite executar análises estatísticas e simulações de maneira simples (DALGAARD, 2008). Por fim, uma revisão foi realizada dos manuais técnicos dos módulos SEEP/W e SLOPE/W do *Geostudio*, com o intuito de compreender as bases teóricas utilizadas nos modelos (SEEP/W & SLOPE/W GEOSTUDIO, 2012).

3.2. OBTENÇÃO DE DADOS DE ESTUDOS ANTERIORES

Um levantamento a partir de estudos anteriores foi feito com o intuito de identificar um banco de dados de propriedades e informações relevantes de barragens de terra homogêneas, contemplando dados de caracterização geotécnica de solos e informações geométricas e de operação das barragens, como nível de água a montante, altura da estrutura do barramento e inclinações de taludes.

Nesse projeto final, foi utilizado o banco de dados de Xue *et al.*, 2014), que será detalhado no Capítulo 4. A partir da função *summary* no software *Rstudio*, um resumo de métricas tradicionais da estatística foi elaborado para cada uma das variáveis. Além disso, por meio da função *cor.plot* do pacote *Psych*, também no *Rstudio*, foi gerada uma matriz de correlação para caracterizar as relações entre as variáveis.

3.3. ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES

A partir da importação da base de dados no software *RStudio*, foi considerada a ocorrência do fenômeno de *piping* como variável dependente e as outras como variáveis independentes. O algoritmo Boruta foi aplicado para identificar as variáveis mais importantes na ocorrência de *piping* nas barragens de terra homogêneas, pela função *boruta* do pacote *Boruta*. Para a variável V10 que indica a ocorrência ou não do fenômeno, cada uma das outras nove variáveis foram avaliadas quanto a sua respectiva importância. Foram considerados diferentes valores para o número máximo de iterações. A simulação final foi executada com 165 iterações, após esse valor, o *output* do algoritmo não apresentou mudanças significativas. Além disso, adotou-se o valor p igual a 0,01.

Para os testes t de Welch, os dados foram divididos em duas amostras: uma com a ocorrência do fenômeno de *piping* (amostra 1) e a outra com a não ocorrência (amostra 2). No software *RStudio*, foi utilizada a função *t.test*, assumindo a distribuição normal dos dados e as variâncias distintas de cada amostra. A hipótese nula testada é que as duas médias são iguais, como indica o teste de Welch. Foi utilizado o intervalo de confiança de 95%. Os testes t foram realizados para determinar os valores t e os valores p para as variáveis V1 a V9. O objetivo desta análise foi avaliar o desvio de cada variável em relação ao comportamento médio e assim identificar as maiores significâncias.

3.4. ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA)

3.4.1. Escolha de Parâmetros de Entrada e Treinamento

A partir do banco de dados obtido anteriormente, foram elencados os parâmetros de entrada da RNA, como sendo as variáveis especificadas no item 3.4.2, e os parâmetros mais adequados de acordo com experiências de estudos anteriores (TAYFUR *et al.*, 2005; GUNTHER & FRITSCH, 2010; XUE *et al.*, 2014).

Uma análise paramétrica foi efetuada para compreender os efeitos das variações de parâmetros da simulação da RNA nos resultados obtidos. Os parâmetros contemplados foram: número de camadas ocultas, número de neurônios por camada, função de ativação e algoritmos de otimização de pesos.

A partir disso, foi realizado o treinamento da RNA com o banco de dados e os parâmetros escolhidos para obter o melhor ajuste. Foi utilizada a função *neuralnet* do pacote *Neuralnet* para gerar os modelos de redes neurais artificiais e a função *compute* do pacote *Dplyr* para realizar as previsões a partir da RNA.

3.4.2. Simulações e Validações do Modelo de RNA

Com o treinamento realizado, as barragens de terra homogêneas do banco de dados utilizado foram simuladas, com o intuito de comparar a previsão da RNA com o observado nos estudos de caso.

A partir disso, foram realizadas validações da RNA simulada com os modelos simulados pelos módulos SEEP/W e SLOPE/W do *Geostudio* (2012). A determinação se há ocorrência ou não de *piping* no modelo de barragem de terra homogênea, pelo mecanismo de erosão regressiva, foi realizada a partir dos seguintes passos:

1. Cálculo do gradiente crítico da barragem, conforme a revisão da literatura;
2. Identificação do gradiente máximo X-Y ascendente pela simulação de percolação pelo SEEP/W;
3. Comparação entre os dois valores, caso o gradiente máximo seja maior do que o gradiente crítico, considera-se a ocorrência de *piping*;

4. Verificação da estabilidade da barragem hipotética, de acordo com o critério mínimo de fator de segurança de 1,5.

O resultado obtido pelo *Geostudio* é comparado com a previsão da RNA, o que permite uma validação e calibração dos parâmetros escolhidos. Trata-se, portanto, de um exercício de verificação do mérito metodológico do trabalho, haja vista a falta de disponibilidade de dados confiáveis de casos reais. Foi utilizada uma barragem de terra homogênea simplificada com os elementos básicos constituintes: corpo, fundação e dreno vertical, de acordo com a Figura 3.4.

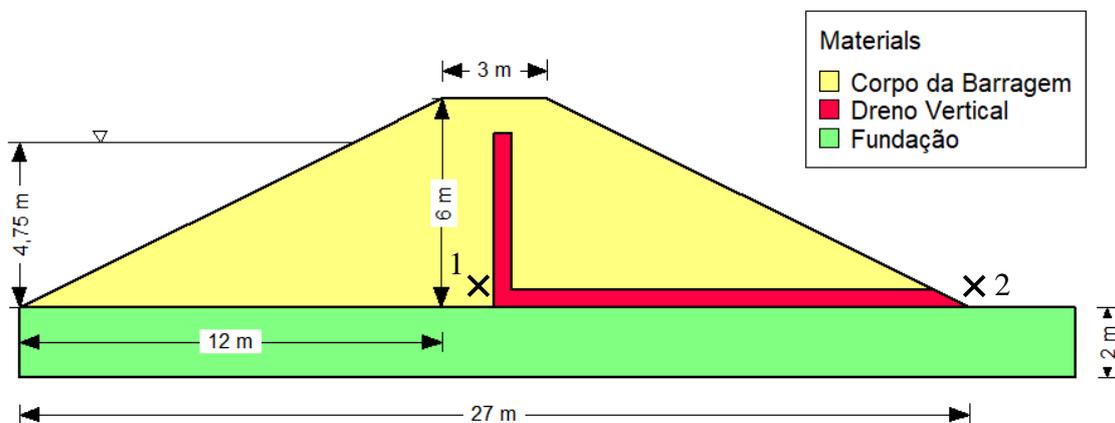


Figura 3.4 – Barragem Padrão no *Geostudio* para a Validação da RNA.

O banco de dados utilizado nesta pesquisa não contempla todas as informações suficientes para a modelagem no *Geostudio*. Para realizar um estudo de caso de um modelo de barragem pela RNA é necessário inserir todas as variáveis utilizadas no treinamento. Dessa forma, para o exercício proposto, foram consideradas as seguintes simplificações:

- Na modelagem no *Geostudio*, foram utilizadas as funções do *software* para estimar o teor de umidade volumétrica conforme o tipo de solo (areia siltosa). Sendo assim no treinamento da RNA de validação, foi retirada a variável V8, correspondente ao diâmetro máximo efetivo da partícula. Os dados de treinamento utilizados foram as observações 1 a 19, com as variáveis V1 a V7, V9 e V10;
- As dimensões adotadas a crista e as fundações estão apresentadas no Capítulo 5;
- Considerou-se a mesma inclinação de jusante, presente no banco de dados, para o talude de montante;

- Na análise de estabilidade do SLOPE/W, um parâmetro de entrada é o peso específico natural do solo. No banco de dados há apenas o peso específico do solo saturado. Adotou-se por simplificação, o peso específico saturado do solo como parâmetro de entrada do *Geostudio*;
- A condutividade hidráulica do banco de dados foi adotada como parâmetro de entrada da simulação de percolação no SEEP/W, para a estimativa das condutividades hidráulicas não saturadas;
- O gradiente máximo ascendente foi o maior valor observado dentre os pontos de análise 1 e 2 da Figura 3.4, respectivamente, próximo à base do dreno vertical e na saída do dreno vertical. Esses pontos foram adotados por apresentarem os maiores valores nas simulações;
- A Tabela 3.1 e a Tabela 3.2 apresentam os critérios utilizados para a simulação nos módulos SEEP/W e SLOPE/W do *Geostudio*. Os valores escritos como VAR foram variados nas simulações, em que cada combinação de parâmetros γ_{sat} , c' , ϕ' e k equivale a um material distinto do banco de dados original.

Tabela 3.1 – Critérios da simulação do SLOPE/W. Fonte: Adaptado de Gameleira e Amaral (2018).

Elemento	Classificação do Solo	γ_{nat} (kN/m ³)	c' (kPa)	ϕ' (°)
Maciço	Areia siltosa	VAR	VAR	VAR
Dreno Vertical	Areia natural fina	18	0	32
Fundação	Arenito	20	25	35

Tabela 3.2 – Critérios da simulação do SEEP/W. Fonte: Adaptado de Gameleira e Amaral (2018).

Elemento	θ_s	θ_r	k (m/s)
Maciço	0,50	0,05	VAR
Dreno Vertical	0,45	0,045	$1,0 \times 10^{-04}$
Fundação	-	-	$1,0 \times 10^{-06}$

No SEEP/W, foi adotada a condição de condutividade constante na fundação, enquanto para o dreno e para o corpo da barragem foram adotados os critérios apresentados de fluxo em solo saturado e não saturado. Para as condições de contorno, considerou-se regime permanente, com a condição hidráulica de carga do nível de água do reservatório para o talude de montante, pressão nula na saída da barragem no pé do dreno vertical, e o dreno vertical como face potencial de infiltração de água. Para a estabilidade de taludes, foi utilizado método de Morgenstern-Price (1965) no cálculo do fator de segurança.

A análise de percolação e o critério adotado para verificação da ocorrência de *piping* no SEEP/W do *Geostudio* não utiliza os parâmetros de resistência do solo como entrada, tais como o ângulo de atrito efetivo e a coesão. Dessa forma, apesar dos algoritmos de *machine learning* utilizarem todas as variáveis de entrada para a previsão de *piping*, as formulações de modelagem geotécnica apresentadas apenas utilizam os parâmetros hidráulicos do solo, tais como teor de umidade volumétrico e a condutividade hidráulica.

A identificação do gradiente crítico máximo ascendente foi realizada em regiões próximas às indicadas na Figura 3.4. Em uma primeira abordagem, todos os gradientes máximos na direção horizontal no sentido à jusante e na vertical no sentido ascendente foram determinados. Desse ponto, o gradiente máximo XY foi calculado a partir dos dois valores máximos obtidos. As diferenças observadas do gradiente ascendente XY obtido manualmente, calculado com os valores máximos em X e em Y, com o gradiente ascendente XY identificado direto do *Geostudio* nos pontos de análise não foram significativas. Portanto, por simplificação, foi utilizado o segundo.

A análise de estabilidade pelo SLOPE/W do *Geostudio* não foi realizada com o intuito de relacionar os valores obtidos de fator de segurança com os fenômenos de erosão interna. As simulações de estabilidade foram elaboradas para verificar o critério mínimo a ser atendido do fator de segurança de 1,5 para os taludes das barragens hipotéticas, além de promover o aprendizado na área no desenvolvimento do projeto final. Os parâmetros do solo variados entre as estruturas hipotéticas foram retirados do banco de dados original, o que seria equivalente a simular a utilização de um material do corpo de uma barragem em outra com dimensões geométricas e ângulo do dreno vertical distintos.

3.5. ANÁLISE DE RESULTADOS E CONCLUSÕES

Com relação ao algoritmo Boruta, será apresentado o resultado obtido a partir do processamento, acompanhado de análises críticas com base em experiências de estudos anteriores. Sobre a simulação da RNA, a partir dos testes e validações feitos, serão elencados os pontos chaves percebidos, de maneira gráfica e por meio de tabelas. Nas conclusões e sugestões serão listados os pontos limitantes do trabalho, além de sugestões de trabalhos futuros para a continuação da linha de pesquisa.

4. BANCO DE DADOS

Com base na revisão da literatura relativa à segurança das barragens e ao fenômeno de *piping* em barragens de terra homogêneas, a seguir é apresentada a base de dados utilizada neste trabalho (XUE *et al.*, 2014). As variáveis representam as principais informações relativas aos parâmetros geométricos de uma barragem de terra homogênea, aos parâmetros geotécnicos do solo e a ocorrência ou não de *piping*.

A base de dados é apresentada na Tabela 4.1, incluindo como variáveis a altura da barragem H (m), o nível de água H_w (m), a inclinação do talude a jusante λ , a coesão efetiva c' (kPa), o ângulo de atrito interno efetivo φ' ($^\circ$), o peso específico saturado γ_{sat} (kN/m³), a condutividade hidráulica k (m/s), o diâmetro efetivo da partícula de solo d_{10} (mm) e o ângulo de inclinação do dreno vertical δ ($^\circ$). A variável binária V10 possui o valor unitário (1) para a ocorrência de fenômeno de *piping* e nulo (0) para a não ocorrência.

A base de dados consiste em 23 estudos de casos de barragens de terra homogêneas chinesas com um total de 10 variáveis (V1 a V10), em que constam 14 barragens com a ocorrência de *piping*. Na Figura 4.1, apresenta-se um resumo de métricas estatísticas de cada uma das variáveis do banco de dados, incluindo os valores máximos e mínimos, as médias, as medianas, assim como o primeiro e o terceiro quartis.

V1		V2		V3		V4		V5	
Min.	: 6.00	Min.	: 4.75	Min.	:0.2490	Min.	: 12.00	Min.	:13.20
1st Qu.	: 28.50	1st Qu.	: 25.00	1st Qu.	:0.3015	1st Qu.	: 23.00	1st Qu.	:23.00
Median	: 39.00	Median	: 33.00	Median	:0.3640	Median	: 41.00	Median	:27.00
Mean	: 47.82	Mean	: 43.49	Mean	:0.3717	Mean	: 55.87	Mean	:26.63
3rd Qu.	: 53.75	3rd Qu.	: 48.00	3rd Qu.	:0.4450	3rd Qu.	: 80.00	3rd Qu.	:30.80
Max.	:133.10	Max.	:126.00	Max.	:0.5000	Max.	:157.00	Max.	:38.50
V6		V7		V8		V9		V10	
Min.	:19.50	Min.	:1.300e-08	Min.	:0.00400	Min.	:14.00	Min.	:0.0000
1st Qu.	:20.55	1st Qu.	:1.500e-07	1st Qu.	:0.00800	1st Qu.	:26.60	1st Qu.	:0.0000
Median	:20.80	Median	:2.900e-06	Median	:0.01000	Median	:53.10	Median	:1.0000
Mean	:20.82	Mean	:7.644e-04	Mean	:0.01291	Mean	:47.56	Mean	:0.6087
3rd Qu.	:21.20	3rd Qu.	:2.000e-05	3rd Qu.	:0.01250	3rd Qu.	:65.50	3rd Qu.	:1.0000
Max.	:22.90	Max.	:1.000e-02	Max.	:0.04000	Max.	:79.00	Max.	:1.0000

Figura 4.1 – Resumo de métricas estatísticas das variáveis.

Tabela 4.1 – Banco de dados de barragens de terra homogêneas. Fonte: Xue *et al.* (2014).

No.	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
	H (m)	H _w (m)	λ	c' (kPa)	ϕ' (°)	γ_{sat} (kN/m ³)	k (m/s)	d ₁₀ (mm)	δ (°)	Atual
1	133,0	123,00	0,455	40	27,0	21,3	$3,0 \times 10^{-09}$	0,09	79,0	1
2	87,5	80,00	0,400	12	28,9	21,2	$3,5 \times 10^{-09}$	0,08	75,0	1
3	35,5	31,00	0,295	60	15,6	20,5	$1,0 \times 10^{-09}$	0,08	43,0	1
4	31,0	29,00	0,249	20	26,7	20,8	$4,0 \times 10^{-06}$	0,10	14,0	0
5	31,0	29,00	0,249	15	26,7	20,1	$7,8 \times 10^{-07}$	0,09	14,0	0
6	29,0	25,00	0,435	30	31,6	20,8	$2,0 \times 10^{-07}$	0,10	23,5	0
7	39,0	35,50	0,466	109	21,2	20,7	$5,1 \times 10^{-08}$	0,40	29,0	0
8	39,0	35,50	0,466	76	13,8	20,7	$5,1 \times 10^{-08}$	0,40	25,0	0
9	28,0	25,00	0,286	157	13,2	20,3	$3,6 \times 10^{-09}$	0,09	60,0	1
10	28,0	25,00	0,286	153	24,8	21,2	$4,8 \times 10^{-10}$	0,09	60,0	1
11	96,0	90,00	0,417	20	26,0	21,0	$3,5 \times 10^{-10}$	0,12	65,0	1
12	56,0	49,00	0,364	30	29,0	19,5	$2,0 \times 10^{-09}$	0,13	59,0	1
13	51,0	47,00	0,308	42	34,5	21,2	$2,2 \times 10^{-10}$	0,12	66,0	1
14	133,1	126,00	0,476	41	32,0	21,7	$1,3 \times 10^{-08}$	0,04	76,0	1
15	13,0	10,50	0,364	44	38,4	22,9	$1,0 \times 10^{-04}$	0,20	26,6	0
16	6,7	5,50	0,400	109	21,2	20,7	$5,1 \times 10^{-08}$	0,04	67,5	1
17	6,0	4,75	0,500	51	38,5	20,9	$7,0 \times 10^{-05}$	0,24	26,6	0
18	87,5	80,00	0,256	14	27,0	21,2	$4,0 \times 10^{-08}$	0,08	28,0	0
19	51,5	46,00	0,455	100	19,3	21,0	$2,9 \times 10^{-08}$	0,07	45,0	1
20	39,5	33,00	0,347	32	27,2	20,2	$5,5 \times 10^{-10}$	0,10	67,0	1
21	29,0	26,00	0,315	26	27,8	20,8	$2,0 \times 10^{-07}$	0,10	65,0	1
22	42,5	39,00	0,361	84	32,2	20,6	$1,3 \times 10^{-10}$	0,04	53,1	1
23	7,0	5,60	0,400	20	30,0	19,5	$3,8 \times 10^{-07}$	0,17	26,6	0

A partir do banco de dados, foram seguidos os passos detalhados no Capítulo 3 para a aplicação do algoritmo de identificação das variáveis relevantes para a ocorrência de piping e o treinamento da rede neural artificial (RNA). Os principais resultados obtidos serão apresentados no Capítulo 5.

5. RESULTADOS E DISCUSSÕES

5.1. ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES

5.1.1. Matriz de Correlação

A matriz de correlação das variáveis do banco de dados é apresentada na Figura 5.1. Na escala de cores, quanto mais azul, maior a relação linear entre as variáveis no sentido positivo. Já em vermelho, maior a relação linear entre as variáveis no sentido negativo, o aumento do valor de uma variável implica na redução da outra.

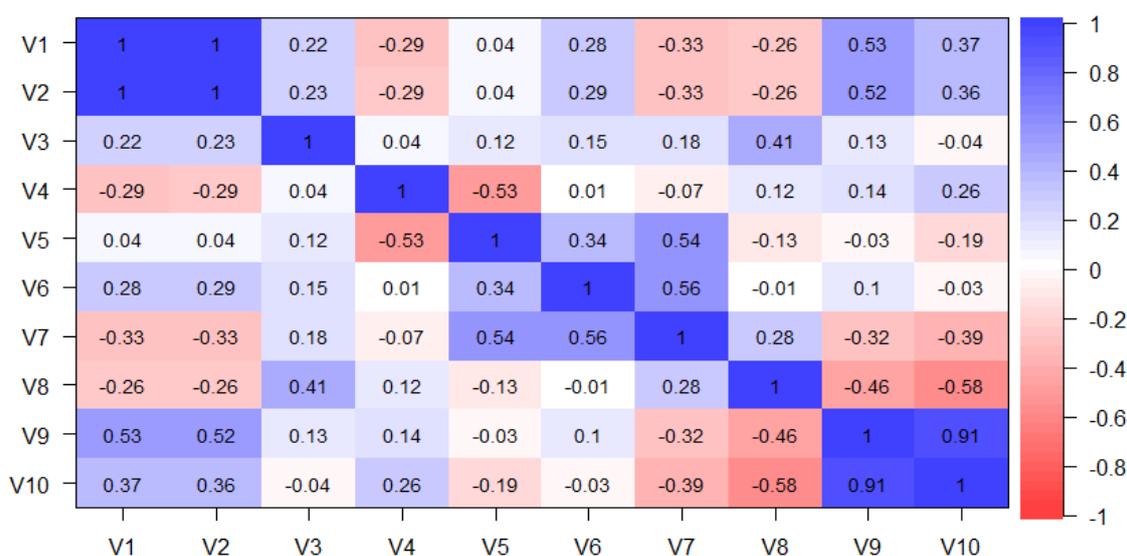


Figura 5.1 – Matriz de correlação das variáveis.

Uma relação esperada, que pode ser observada na matriz, é sobre as variáveis V1 (altura da barragem) e V2 (nível de água da barragem). Em uma barragem de altura maior, espera-se um nível de água maior, o que é indicado pela correlação unitária entre as duas variáveis. Nota-se uma relação positiva forte (+0,91) entre a V9 (ângulo de inclinação do dreno vertical) e a V10 (ocorrência de *piping*). Enquanto uma relação negativa regular (-0,58) entre a variável V8 (diâmetro efetivo da partícula de solo) e a V10 (ocorrência de *piping*).

Desta forma, a matriz aponta que quanto menor o diâmetro efetivo da partícula e maior o ângulo de inclinação do dreno vertical, maior será a tendência de surgirem processos de erosão interna. Tendo em vista que o *piping* está associado geralmente à solos de alta condutividade hidráulica (FOSTER *et al.*, 2000), e logo, com menos partículas finas, as

ocorrências do banco de dados podem ter ocorrido em volta de tubulações que atravessam o corpo das barragens ou por conta de falhas executivas, tais como trincas e materiais com compactação não adequada. O ângulo de inclinação do dreno, por sua vez, afeta a área de contato entre o material do corpo da barragem e o material não coesivo. Sendo assim, espera-se que ângulos próximos de 90° tornem a barragem menos suscetível às erosões internas, visto que diminuem as zonas de transição que podem induzir o mecanismo de erosão de contato. Contudo, especificamente para o banco de dados utilizado neste projeto final, isso não ocorre.

5.1.2. Algoritmo Boruta

A Figura 5.2 aponta que as variáveis com maior importância para a ocorrência ou não de *piping* (V10) são a relação de inclinação a jusante (V3), o diâmetro efetivo da partícula de solo (V8), a condutividade hidráulica (V7) e o ângulo de inclinação do dreno vertical (V9). As outras variáveis não são importantes para a variável V10, de acordo com a simulação. A variável menos importante foi o peso específico saturado (V6). As variáveis *shadowmin*, *shadowmean*, e *shadowmax* são utilizadas na simulação pelo algoritmo para determinar a pontuação de importância para as variáveis, correspondentes às colunas das variáveis sombra (KURSA & RUDNICKI, 2010).

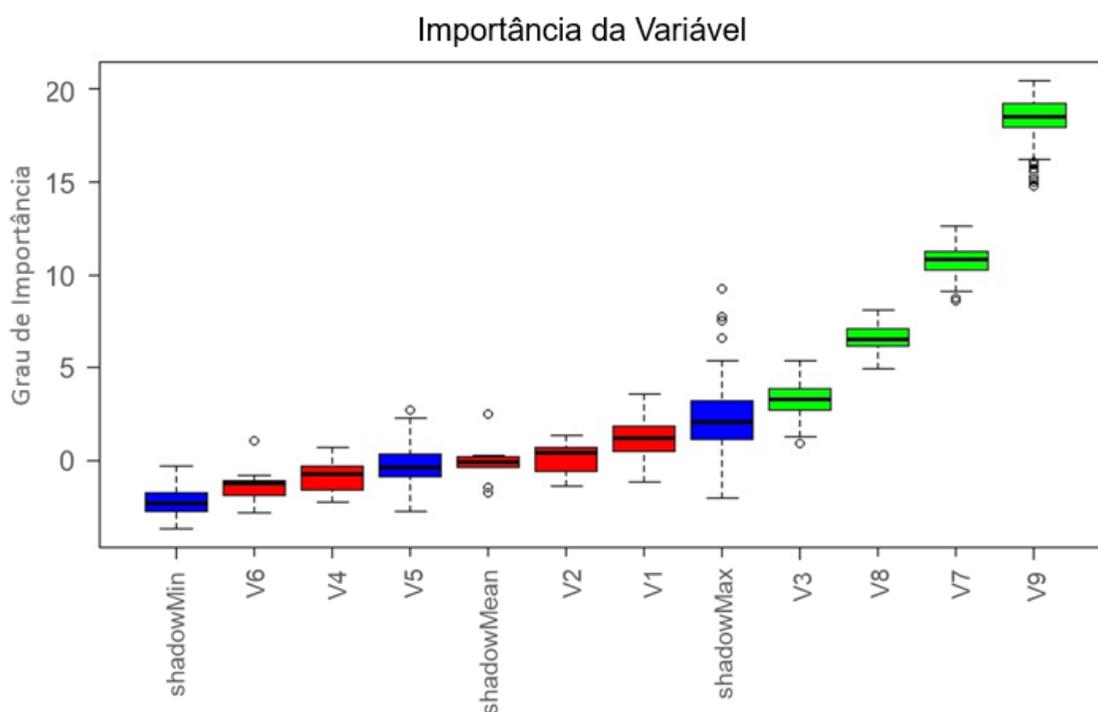


Figura 5.2 – Resultados do algoritmo Boruta.

Da análise dos dados na Tabela 4.1, é possível inferir que a ocorrência de *piping* nas barragens catalogadas está relacionada com ângulos de inclinação do dreno maiores que 43°. Para a base de dados utilizada neste trabalho, observa-se que quanto maior for o ângulo de inclinação do dreno, maior será a suscetibilidade da barragem de terra homogênea ao fenômeno de *piping*. Essa observação, contudo, não é esperada já que drenos verticais com ângulo próximo aos 90° diminuem as zonas de transição que podem induzir erosões internas de contato.

A condutividade hidráulica (V7), para essa base de dados, apresentou um comportamento inesperado, o valor médio para a amostra com ocorrência de *piping* é de $2,20 \times 10^{-8}$ m/s, e a amostra com não ocorrência apresentou um valor médio de $1,95 \times 10^{-5}$ m/s. O diâmetro efetivo da partícula de solo (V8) possui uma correlação negativa com o *piping*, ou seja, quanto menor o diâmetro, maior a tendência de ocorrência. Pela consulta à literatura, sabe-se que o fenômeno de *piping* ocorre tanto na presença de solos de diâmetro efetivo grande quanto pequeno. Entretanto, o comportamento destas variáveis é incomum, uma vez que a ocorrência de *piping* está geralmente associada à variabilidade das propriedades do solo, elevadas condutividades hidráulicas, e variações expressivas do diâmetro efetivo das partículas do solo (FOSTER *et al.*, 2000).

5.1.3. Testes Estatísticos

Os resultados para os valores t e os valores p para cada variável dos testes t são apresentados na Tabela 5.1.

Tabela 5.1 – Resultados do teste t e do valor p.

No.	Variável	valor t	valor t abs.	valor p
V9	δ (°)	-11,433	11,433	$2,30 \times 10^{-10}$
V8	d_{10} (mm)	2,620	2,620	0,029
V1	H (m)	-2,026	2,026	0,056
V2	H_w (m)	-1,989	1,989	0,060
V7	k (m/s)	1,541	1,541	0,162
V4	c' (kPa)	-1,347	1,347	0,193
V5	ϕ' (°)	0,837	0,837	0,416
V3	λ	0,188	0,188	0,854
V6	γ_{sat} (kN/m ³)	0,130	0,130	0,899

Como mencionado, foi considerada a amostra 1 com a ocorrência de *piping* e amostra 2 para normalidade, conforme a Equação (2.27). Conseqüentemente, para um dado parâmetro, o sinal negativo do valor t representa que a média desse parâmetro é mais alta na barragem com *piping* do que na barragem sem *piping*: se $\bar{x}_2 - \bar{x}_1 < 0$, então $\bar{x}_1 > \bar{x}_2$. Portanto, em uma primeira abordagem, espera-se que valores maiores do ângulo de inclinação do dreno vertical (V9), da altura da barragem (V1), do nível de água (V2), e da coesão efetiva (V4) possam causar *piping* nas barragens de terra homogêneas.

Esse resultado está de acordo com a literatura técnica, uma vez que existe um maior risco de percolação de água em barragens mais altas e sob níveis de água mais elevados. Além disso, esses aspectos contribuem para a instabilidade da estrutura e podem causar rupturas. O dreno vertical atua como uma barreira protetora para o *piping*, apesar disso, espera-se que drenos inclinados contribuam para uma menor estabilidade interna. Isso ocorre uma vez que as partículas do solo encontram mais dificuldade em permanecer na sua posição para o fluxo de água sob elevados gradientes hidráulicos, conforme Messerklinger (2013).

A formação de ruptura é mais provável em barragens com uma zona a jusante de materiais coesivos, tais como argila. Isso pode ser explicado por esses serem mais propensos a falhar devido ao desenvolvimento acelerado de *piping*, além de apresentarem baixa condutividade hidráulica e altas poropressões. Estes aspectos causam instabilidades de taludes, além de perdas de resistência de cisalhamento, o que contribui para a ocorrência da perda de borda livre e galgamentos (FOSTER & FELL, 1999).

O valor t também representa uma medida da diferença entre a média de um parâmetro nos dois conjuntos de amostras, tal como descrito na Equação (2.27). Assim, uma variável com um valor t mais elevado indica que existe uma grande discrepância da média dessa variável nas amostras. Isso evidencia que a variável é relevante para o evento de *piping*, uma vez que se afasta significativamente do valor médio global. O ângulo de inclinação do dreno vertical (V9) é a variável com o maior valor t absoluto, seguido do diâmetro efetivo da partícula (V8), sendo considerados, portanto, mais significativos.

Como apresentado na Tabela 2.2, quanto menor o valor p, maior a evidência de que o comportamento da variável nas amostras não é igual. Por conseguinte, a variável

apresenta um significado estatístico mais elevado. O valor p de V9 apresenta forte evidência de que os valores da variável na amostra 1 e na amostra 2 são distintos. Este resultado está alinhado o resultado do algoritmo Boruta, já que a V9 foi classificada como a variável mais importante.

Também se pode concluir que o ângulo de inclinação do dreno vertical (V9) e o diâmetro efetivo da partícula de solo (V8) são classificados como estatisticamente significativos, uma vez que os seus valores p são menores que a significância considerada de 5%. Consequentemente, implica que essas duas variáveis são preferíveis para serem utilizadas para treinar um modelo que simula a ocorrência de *piping*, é constatado que as correlações com o fenômeno não são aleatórias.

5.2. ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA)

Com o intuito de compreender os parâmetros de uma RNA e suas influências nas previsões de ocorrência de *piping* no banco de dados, foram realizadas simulações com o número de neurônios em uma camada, número de camadas ocultas, algoritmos de otimização dos pesos e funções de ativação.

O banco de dados foi dividido em duas partes, as observações 1 a 19 foram utilizadas no treinamento da RNA, enquanto as 4 observações restantes foram destinadas para a validação de resultados. Em toda as simulações, considerou-se uma RNA semelhante à da Figura 5.3.

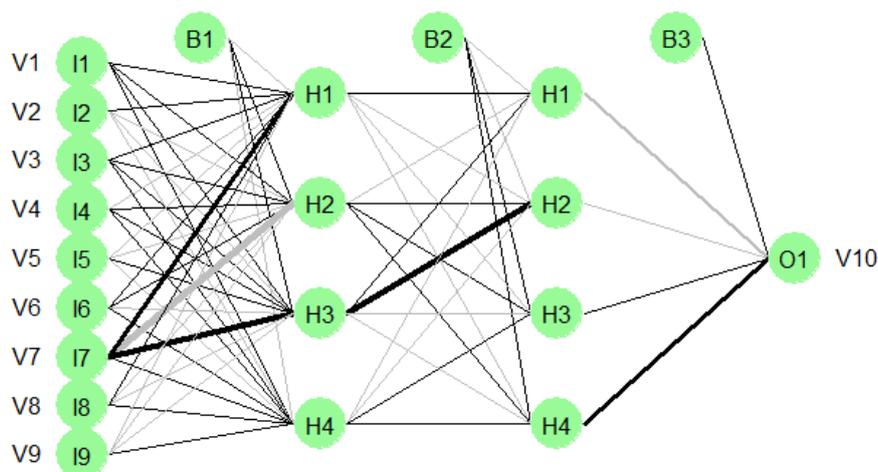


Figura 5.3 – Arquitetura da RNA modelo da simulação.

Em que I_1 representa o *input* da variável V_1 , B_1 o termo *bias* da primeira camada oculta, H_1 o primeiro neurônio da camada oculta e O_1 o *output* do modelo com a previsão de ocorrência de *piping*. A arquitetura da RNA da Figura 5.3, portanto, possui nove variáveis de entrada, duas camadas ocultas, com quatro neurônios em cada, e uma variável de saída.

5.2.1. Análise Paramétrica – Neurônios

Nas configurações da RNA, para avaliar os efeitos da variação do número de neurônios, foi adotada apenas uma camada oculta em que número de neurônios correspondentes era alterado. A função de ativação utilizada foi a sigmoide e o algoritmo de otimização dos pesos o RBP, considerando um total de 1000 repetições.

As figuras a seguir apresentam os resultados. Os pontos quadrados em azul são os observados na base de dados, em que o valor unitário representa a ocorrência de *piping*. Enquanto os pontos triangulares em verde e circulares em vermelho são os parâmetros de análise.

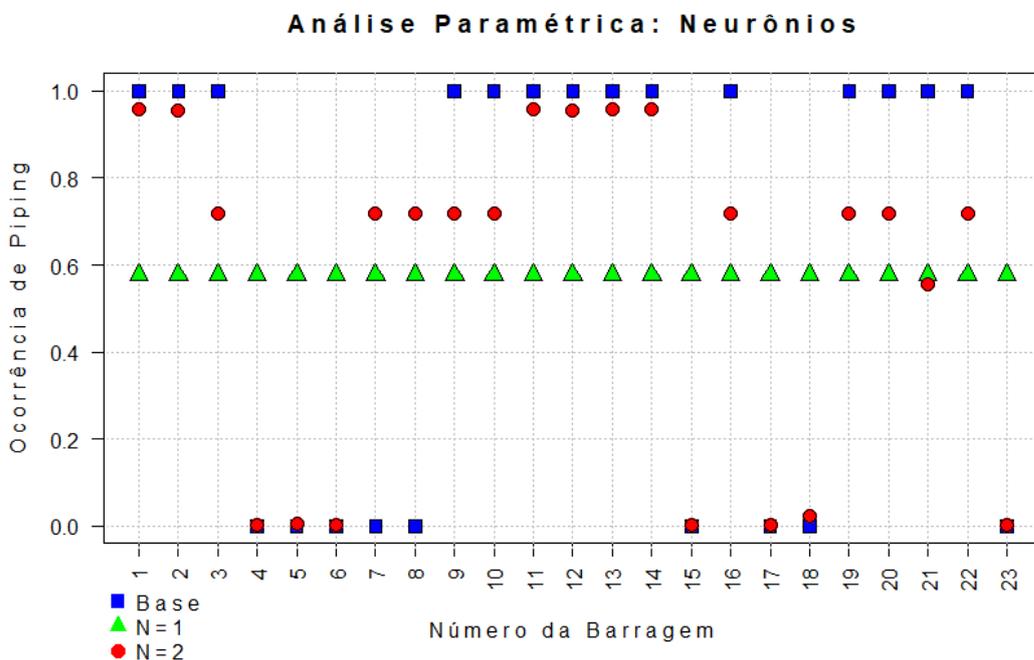


Figura 5.4 – Resultados com uma camada oculta ($N = 1$ e $N = 2$ neurônios).

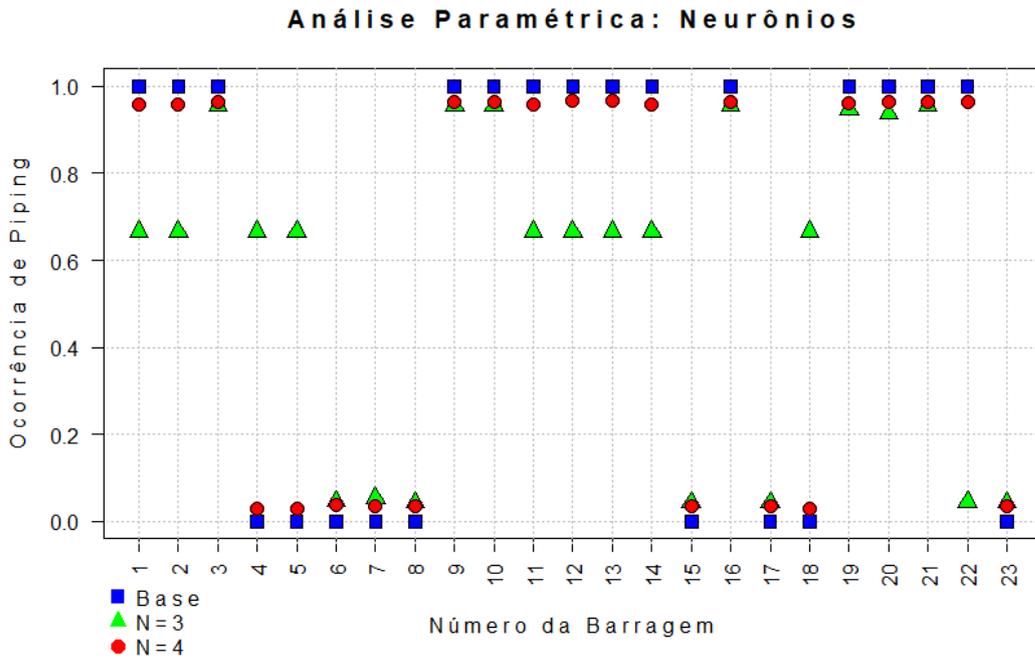


Figura 5.5 – Resultados com uma camada oculta (N = 3 e N = 4 neurônios).

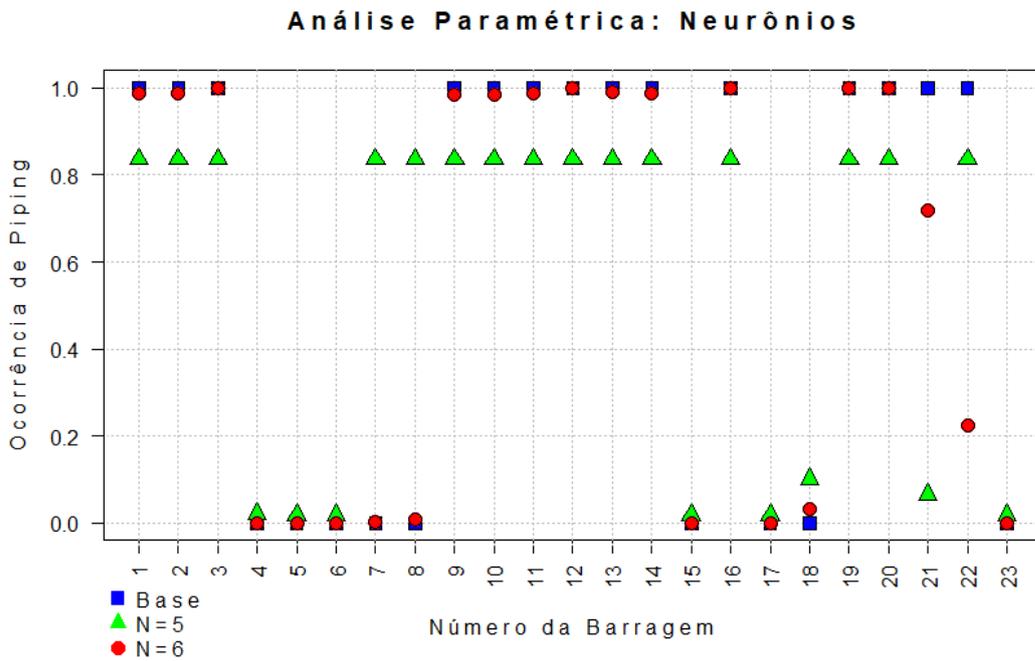


Figura 5.6 – Resultados com uma camada oculta (N = 5 e N = 6 neurônios).

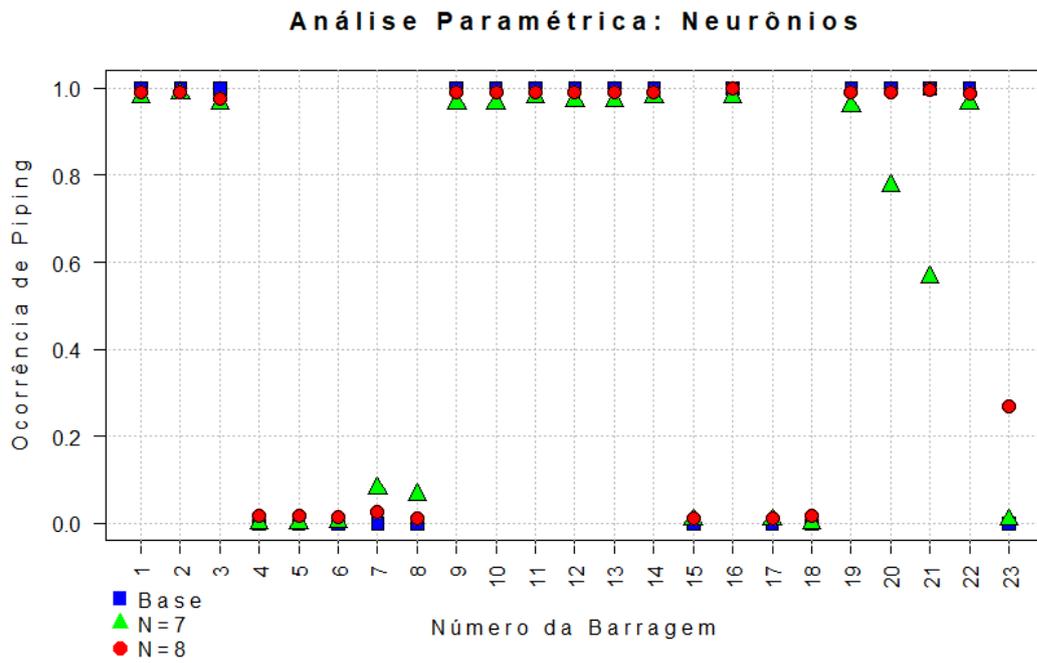


Figura 5.7 – Resultados com uma camada oculta (N = 7 e N = 8 neurônios).

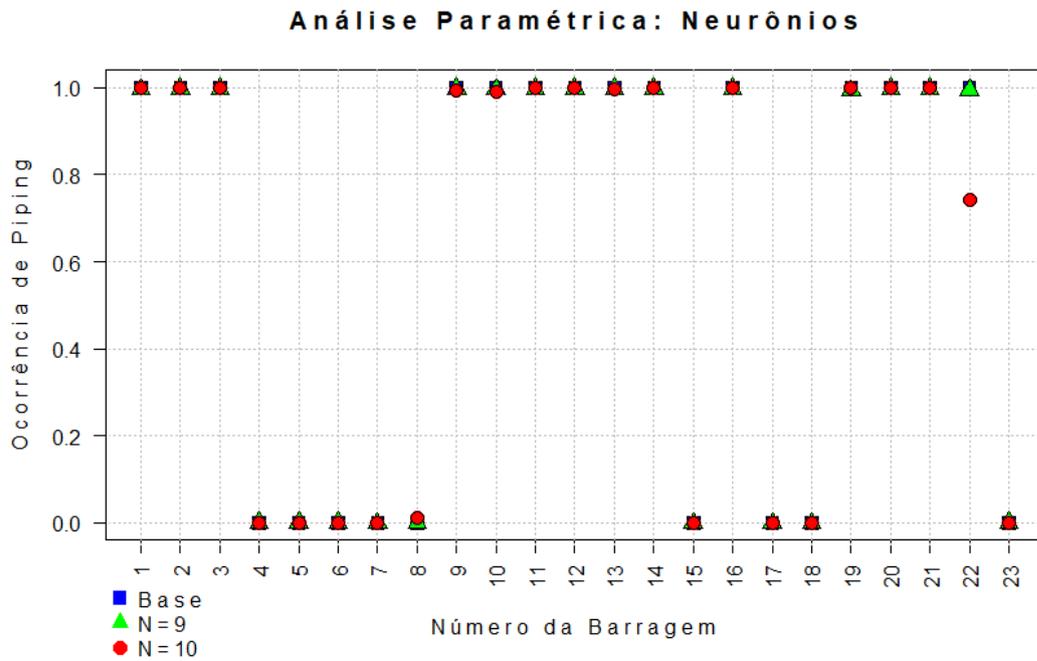


Figura 5.8 – Resultados com uma camada oculta (N = 9 e N = 10 neurônios).

Nota-se que para os cenários das simulações, as redes neurais artificiais com o número de neurônios inferiores a cinco, com a camada oculta única, apresentaram muitas previsões deslocadas dos valores esperados. O número baixo de neurônios, frente ao número de variáveis de entrada, fez que com que a acurácia de resultados não fosse satisfatória.

Para as simulações com mais de cinco neurônios na camada oculta única, as previsões ficaram mais próximas dos valores esperados. Observa-se uma acurácia maior, no entanto, os valores observados que tiveram maior discrepância são aqueles denominados de validação, ou seja, os dados que não foram utilizados no treinamento da RNA. Ainda, o aumento do número de neurônios não reflete diretamente em melhores resultados.

5.2.2. Análise Paramétrica – Camadas

Para a simulação das camadas, o número de camadas ocultas foi alterado de um a três. O número de neurônios por camada permaneceu fixo para avaliar exclusivamente a influência do aumento do número de camadas ocultas na previsão de resultados. Os parâmetros adotados de simulação da RNA foram os mesmos da análise do número de neurônios. Os resultados estão apresentados a seguir, em que a notação 4,4 representa duas camadas ocultas com 4 neurônios em cada.

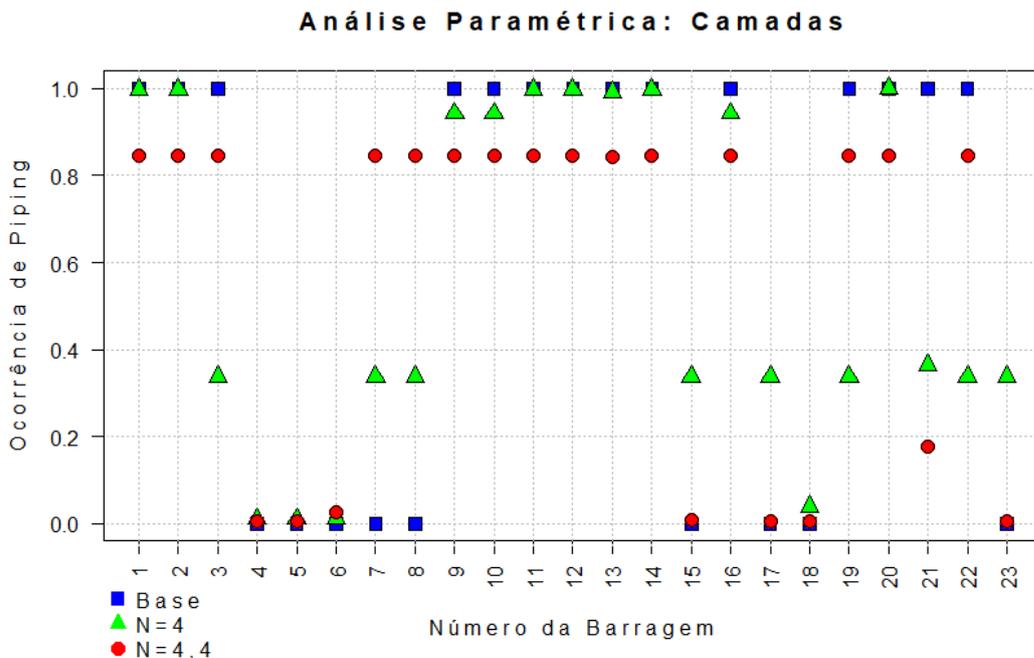


Figura 5.9 – Resultados com mais de uma camada oculta (N = 4 e N = 4,4 neurônios).

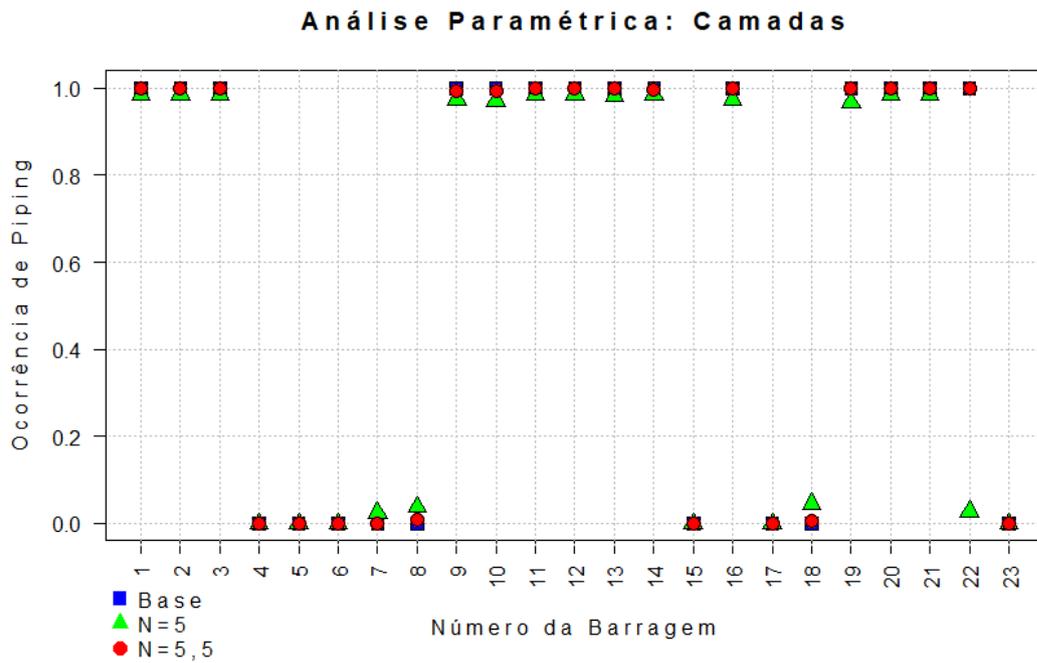


Figura 5.10 – Resultados com mais de uma camada oculta (N = 5 e N = 5,5 neurônios).

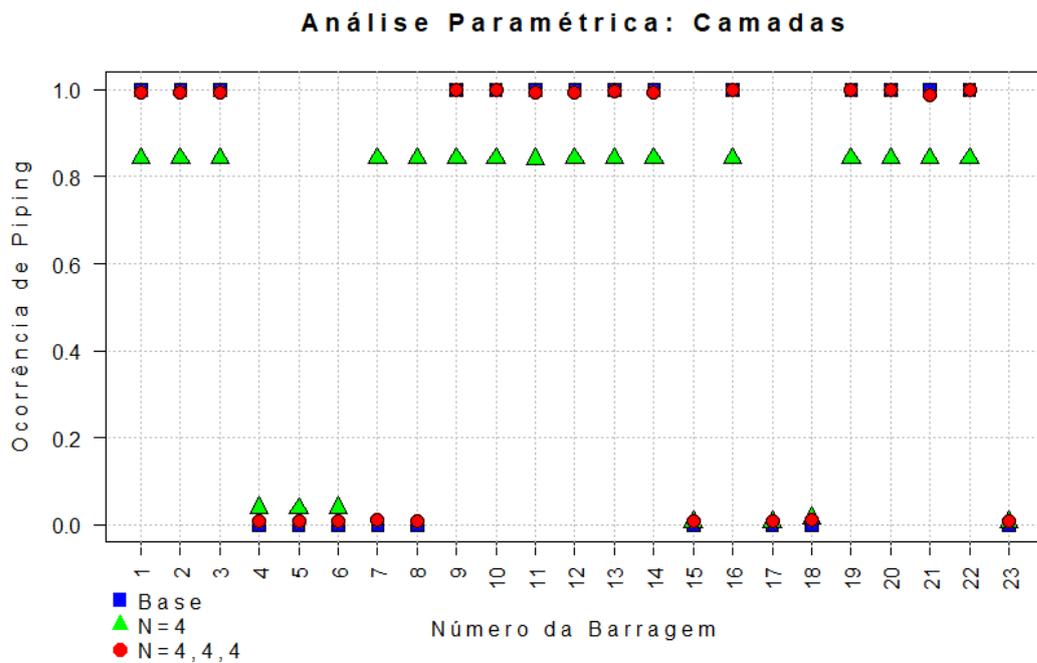


Figura 5.11 – Resultados com mais de uma camada oculta (N = 4 e N = 4,4,4 neurônios).

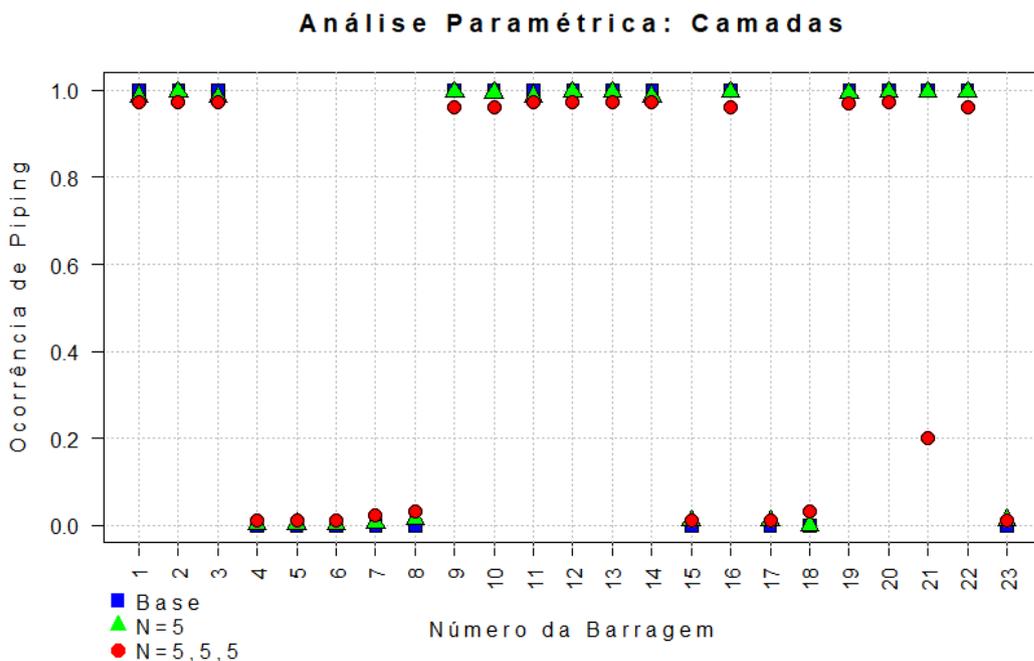


Figura 5.12 – Resultados com mais de uma camada oculta ($N = 5$ e $N = 5,5,5$ neurônios).

Conforme o exposto, nota-se que o aumento no número de camadas ocultas acarreta na maior acurácia de resultados, ao comparar o resultado esperado do banco de dados e o previsto pela RNA. Com uma e duas camadas ocultas de quatro neurônios cada, observa-se uma distribuição das previsões distante dos valores esperados, 0 para a não ocorrência de *piping* e 1 para a ocorrência. A partir de três camadas ocultas com quatro neurônios cada e a partir de uma camada oculta com cinco neurônios, os resultados previstos demonstram maior estabilidade, embora sejam constatadas previsões com pouca acurácia nos dados de validação (observações 20 a 23).

5.2.3. Análise Paramétrica – Algoritmos de Otimização

Para os algoritmos de otimização, foram adotadas duas opções, *backpropagation* (BP), com 100 iterações e taxa de aprendizado de 0,00001, e *resilient backpropagation* (RBP), com 1000 iterações. No algoritmo BP, segundo Tayfur *et al.* (2005), a taxa de aprendizado deve ser a menor possível para aumentar a acurácia de resultados. Todavia, ao diminuir a taxa de aprendizado, o tempo de cada iteração aumenta significativamente e exige maior esforço computacional.

A Figura 5.13 apresenta o resultado da simulação, em que foram consideradas três camadas ocultas com respectivamente oito, cinco e três neurônios. Essa arquitetura foi obtida a partir de diversos testes das etapas anteriores como a mais correta nas previsões.

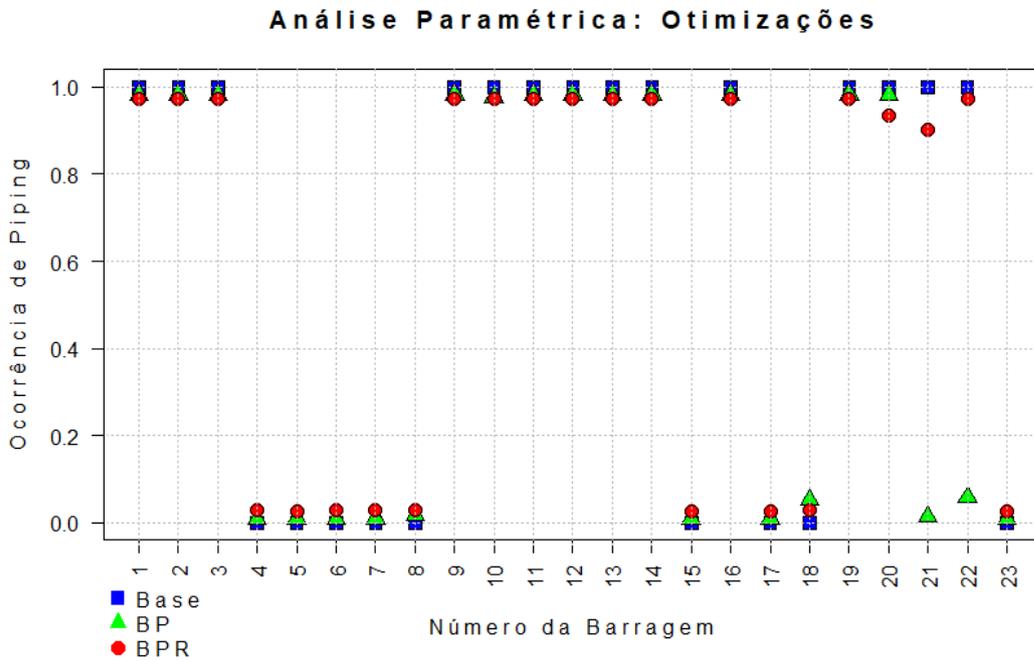


Figura 5.13 – Resultados dos Algoritmos de Otimização.

O algoritmo RBP apresentou maior acurácia de resultados. No caso do algoritmo BP, nota-se que nos dados de validação nas observações 21 e 22, as previsões da RNA contrastaram significativamente do esperado. Em modelos de precisão de *piping*, esse resultado é o menos desejado, visto que uma barragem já com processos de erosão interna foi classificada como não suscetível pelo modelo.

5.2.4. Análise Paramétrica – Funções de Ativação

Para as funções de ativação, foram testadas a tangente hiperbólica e a sigmoide. Com uma arquitetura de oito neurônios na primeira camada oculta, cinco na segunda camada oculta e três na terceira camada oculta, a simulação foi realizada com o algoritmo de otimização de pesos RBP com 100 repetições. Os dados de treinamento continuaram sendo as observações de 1 a 19 do banco de dados.

Pela Figura 5.14, nota-se que os dados de treinamento tiveram bons resultados das previsões para função sigmoide, enquanto houve discrepâncias nos dados de validação

(observações 20 e 21). A função tangente hiperbólica mostrou-se com a distribuição dispersa, em que muitas previsões ficaram entre valores intermediários e em alguns distantes do esperado (observações 7 e 18). Para a função sigmoide, na observações 20 e 21, nota-se uma situação similar à apresentada anteriormente, em que há a previsão de não ocorrência do fenômeno sendo que na realidade esse ocorre na barragem de estudo. Esse comportamento pode ser refinado com o ajuste nos demais parâmetros da RNA.

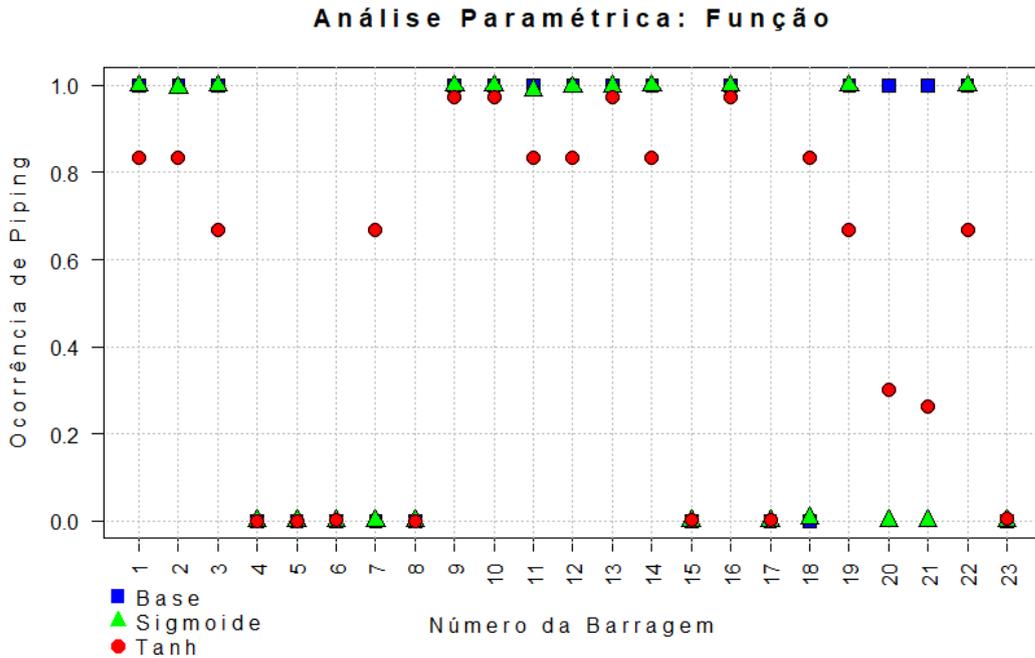


Figura 5.14 – Resultados das Funções de Ativação.

5.2.5. Retroanálise do Banco de Dados

Como mencionado no Capítulo 3 de Metodologia, a variável V8, correspondente ao diâmetro efetivo da partícula, foi retirada do banco de dados para o treinamento e validação visto a falta de disponibilidade de informações para a modelagem adequada da barragem hipotética no *Geostudio*. Assim, a partir da análise de parâmetros realizada, diferentes arquiteturas e parâmetros foram testados com o intuito de obter resultados com maior acurácia. Desta forma, a rede neural artificial (RNA) obtida com esse processo é ilustrada na Figura 5.15 e a Figura 5.16 apresenta os resultados das previsões.

A arquitetura com três camadas ocultas, compostas por respectivamente, 8, 5 e 3 neurônios foi adotada a partir de testes e verificações de acurácia quanto aos dados de treinamento. Foram simulados distintos modelos de RNA, começando com 1 neurônio até o valor de 10, em apenas uma camada oculta. Depois, o número de neurônios com a

melhor acurácia era mantido na primeira camada oculta, e assim, repetia-se o teste para a próxima camada oculta. Visto que a maneira como a arquitetura adotada foi atingida, houve um trabalho braçal que pode ser otimizado com algoritmos e *frameworks*, conforme será abordado nos tópicos de sugestões.

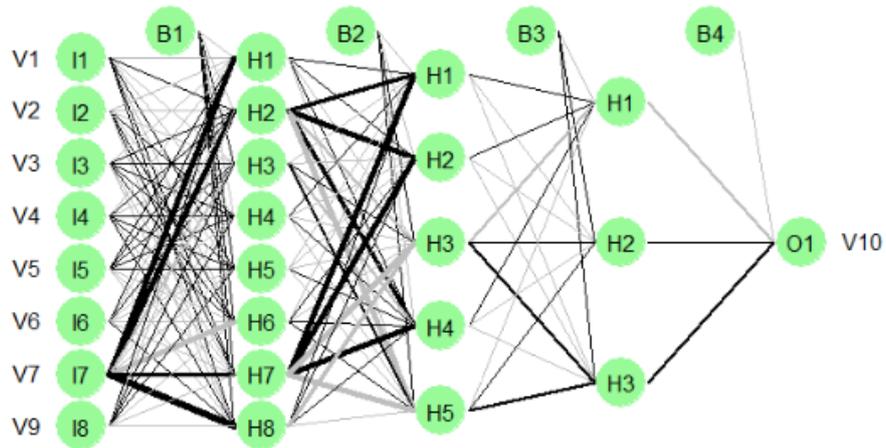


Figura 5.15 – Arquitetura da RNA após Treinamento.

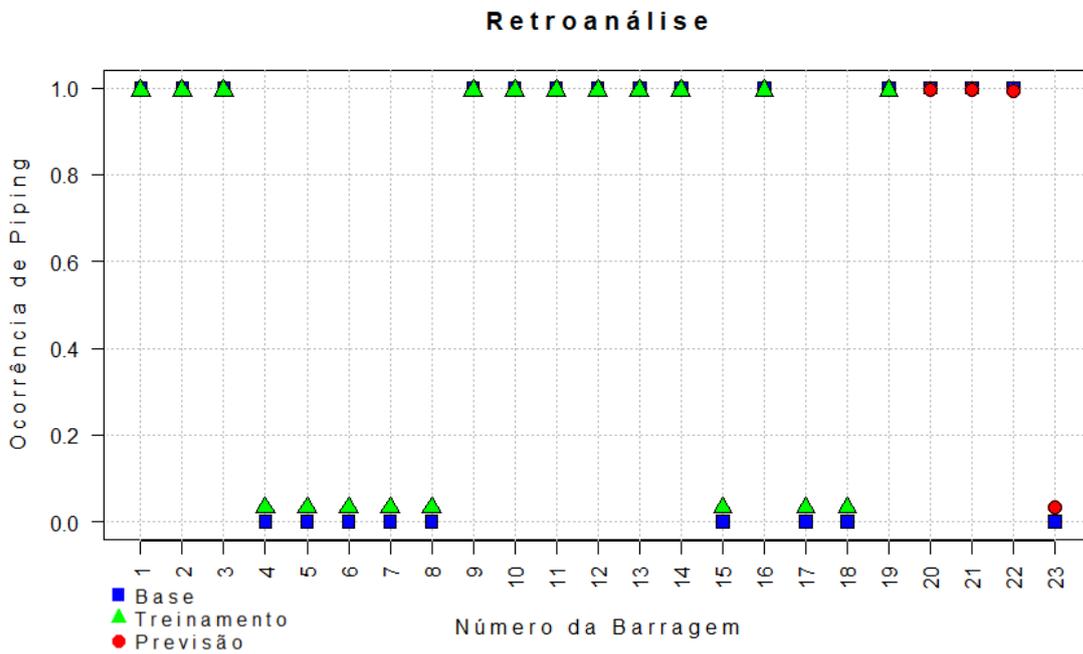


Figura 5.16 – Resultados da Previsão da RNA após Treinamento.

Para o treinamento da RNA, foi utilizada a função de ativação sigmoide e o algoritmo de otimização dos pesos RBP, *resilient backpropagation*, com 10.000 repetições para o treinamento. A Tabela 5.2 aponta os resultados da base de dados e os previstos pela RNA

para a variável V10, ocorrência ou não de *piping*, em cada uma das observações. Os maiores erros associados foram na previsão das observações sem a ocorrência de *piping*, conforme a tabela, em torno de 3%.

Tabela 5.2 – Comparação de Resultados da Base de Dados com os Previstos.

No.	V10	RNA	Diferença	No.	V10	RNA	Diferença
1	1	0,994598770	0,005401230	13	1	0,994598730	0,005401270
2	1	0,994860180	0,005139820	14	1	0,994598770	0,005401230
3	1	0,994919750	0,005080250	15	0	0,032346680	-0,032346680
4	0	0,032355840	-0,032355840	16	1	0,994882950	0,005117050
5	0	0,032355840	-0,032355840	17	0	0,032346690	-0,032346690
6	0	0,032355900	-0,032355900	18	0	0,032357720	-0,032357720
7	0	0,032374780	-0,032374780	19	1	0,994582670	0,005417330
8	0	0,032523000	-0,032523000	20	1	0,994693060	0,005306940
9	1	0,994882940	0,005117060	21	1	0,994789380	0,005210620
10	1	0,994882940	0,005117060	22	1	0,994056590	0,005943410
11	1	0,994598760	0,005401240	23	0	0,032346710	-0,032346710
12	1	0,994598290	0,005401710				

5.2.6. Validação da RNA

Neste item é proposto um exercício para validação da metodologia para a previsão da ocorrência de *piping* em barragens de terra homogêneas hipotéticas para o mecanismo de erosão regressiva. O intuito é consolidar a metodologia de análise dos fenômenos de erosão interna por meio de técnicas de aprendizado de máquina, para que dessa forma, essa possa ser utilizada em casos reais.

Na Figura 5.17 e a na Figura 5.18 são exibidos os resultados da simulação, respectivamente, nos módulos SLOPE/W e SEEP/W, em que os eixos de elevação e distância estão em metros. A barragem hipotética, definida de acordo com os critérios detalhados no Capítulo 3, foi submetida aos esforços decorrente da operação com um nível de água constante.

Pela análise de estabilidade, o valor obtido para o fator de segurança mínimo para o talude jusante foi de 3,214, o que está condizente com as recomendações de projeto que estabelecem um valor mínimo de 1,5 (DAS & SOBHAN, 2014).

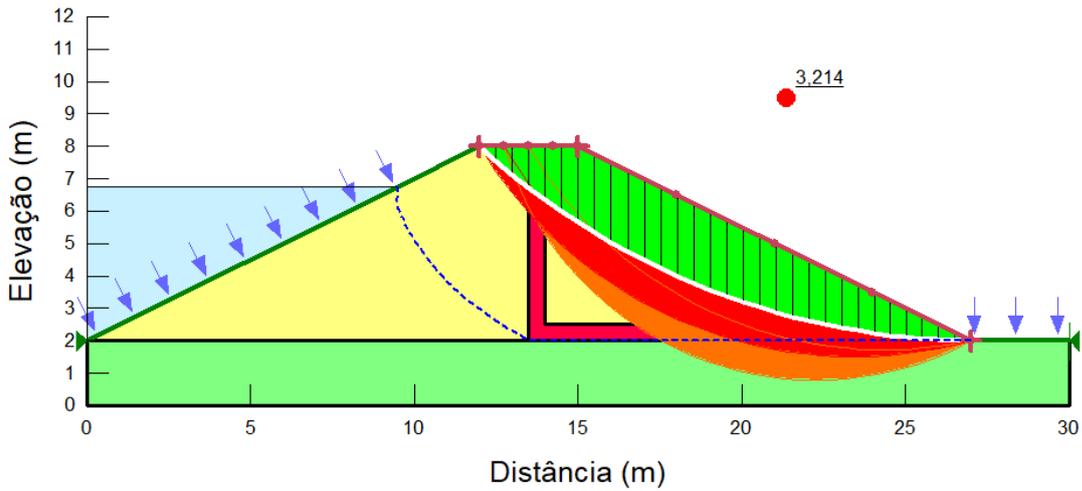


Figura 5.17 – Resultados da Simulação no SLOPE/W.

Pela simulação realizada no SEEP/W, foi identificado o gradiente máximo da seção da barragem, conforme a Figura 5.18. A partir da Equação (2.4), determina-se o gradiente crítico ascendente para o modelo da barragem. Como a seguir, caso $\gamma_{sat} = 20,2 \text{ kN/m}^3$:

$$i_{cr} = \frac{\gamma_{sub}}{\gamma_w} = \frac{20,2 - 10}{10} = 1,02 \quad (5.1)$$

Neste caso, o gradiente máximo ascendente observado pela seção da Figura 5.18 nos pontos de análise foi de 1,44, valor superior ao gradiente crítico de 1,02, o que indica a susceptibilidade aos fenômenos de erosões internas. Portanto, considera-se a ocorrência de *piping* (valor unitário).

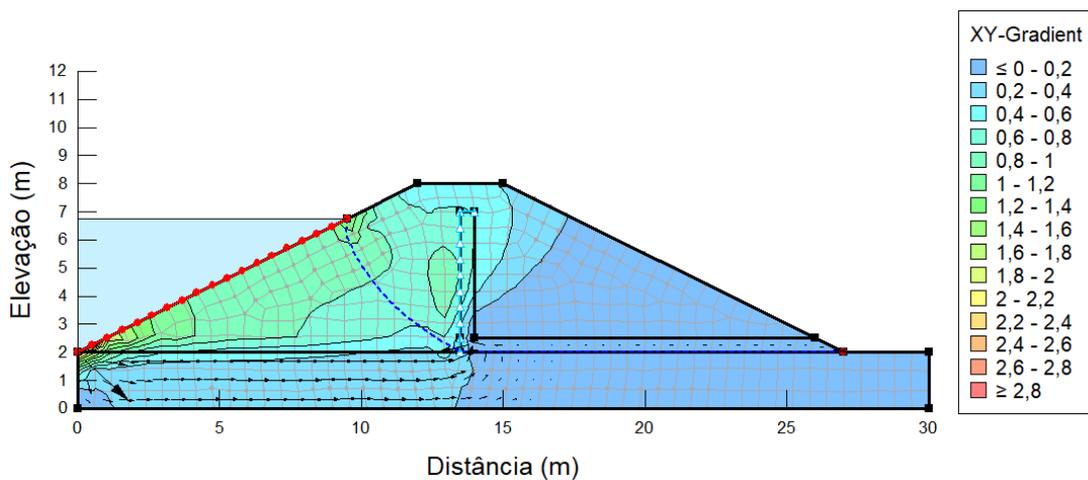


Figura 5.18 – Resultados da Simulação no SEEP/W.

O mesmo procedimento foi adotado para cada uma das barragens hipotéticas deste trabalho. Foram adotados ângulos de inclinação do dreno vertical de 90° e 45° sendo a seção tipo apresentada pela, respectivamente, Figura 5.19 e Figura 5.20.

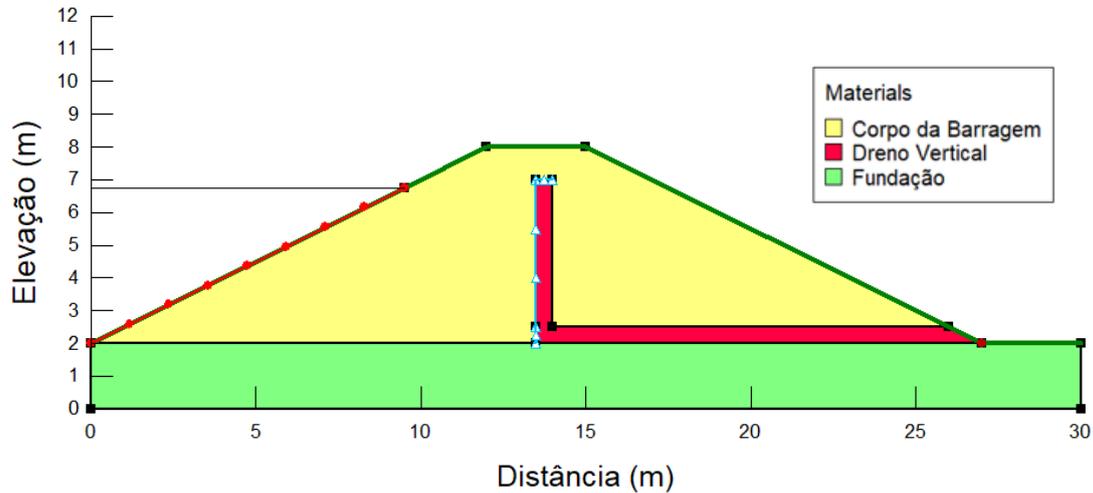


Figura 5.19 – Barragem Hipotética com Dreno Vertical a 90°.

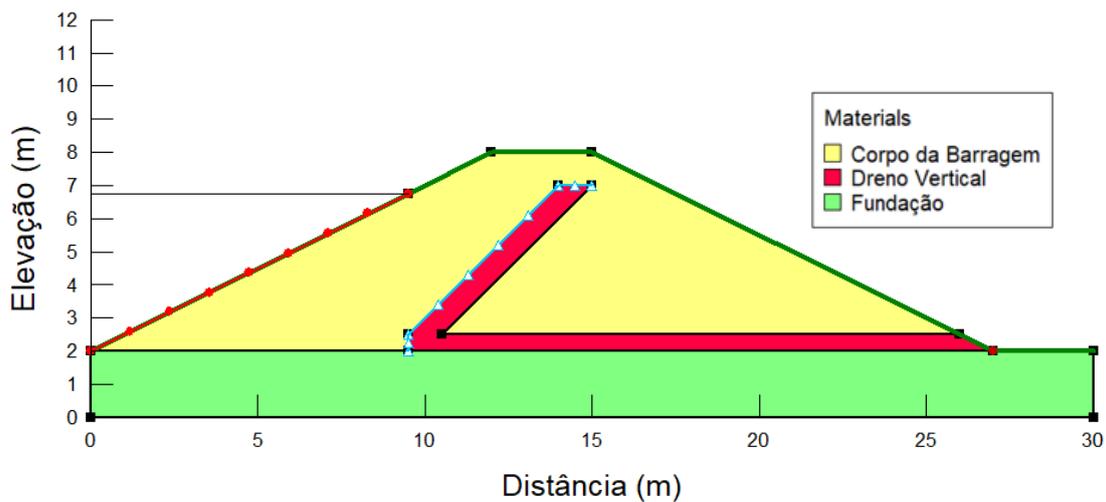


Figura 5.20 – Barragem Hipotética com Dreno Vertical a 45°.

Dessa forma, a partir da simulação de 30 barragens hipotéticas, conforme a Tabela 5.3 e a Tabela 5.4, foram obtidos os resultados da Tabela 5.5 e da Figura 5.21. Em que FS_{Mon} e FS_{Jus} são os fatores de segurança de, respectivamente, montante e jusante da barragem, calculados pelo método de Morgenstern-Price (1965), i_{cr} é o gradiente crítico calculado pela Equação (2.4), $i_{máx}$ é o gradiente máximo ascendente observado na modelagem no SEEP/W e FS_{Piping} é o fator de segurança para a ocorrência de *piping* de acordo com a Equação (2.8).

Tabela 5.3 – Barragens Hipotéticas da Validação.

No.	V1	V2	V3	V4	V5	V6	V7	V9	V10
	H (m)	H _w (m)	λ	c' (kPa)	ϕ' (°)	γ_{sat} (kN/m ³)	k (m/s)	δ (°)	Atual
1	6,00	4,75	0,500	20	22,0	18,0	1,00E-07	90	1
2	6,00	4,75	0,500	20	22,0	18,0	1,00E-07	45	1
3	6,00	4,75	0,500	51	38,5	20,0	7,00E-05	90	1
4	6,00	4,75	0,500	51	38,5	20,0	7,00E-05	45	1
5	6,00	4,75	0,500	32	27,2	20,2	5,50E-10	90	1
6	6,00	4,75	0,500	32	27,2	20,2	5,50E-10	45	1
7	6,00	4,75	0,500	20	30,0	19,5	3,80E-07	90	1
8	6,00	4,75	0,500	20	30,0	19,5	3,80E-07	45	1
9	6,00	4,75	0,500	40	27,0	21,3	3,00E-09	90	1
10	6,00	4,75	0,500	40	27,0	21,3	3,00E-09	45	1
11	6,00	4,75	0,500	109	21,2	20,7	5,10E-08	90	1
12	6,00	4,75	0,500	109	21,2	20,7	5,10E-08	45	1
13	50,00	45,00	0,455	42	34,5	20,2	1,00E-07	90	1
14	50,00	45,00	0,455	42	34,5	20,2	1,00E-07	45	1
15	50,00	45,00	0,455	100	19,3	21,0	2,90E-08	90	1
16	50,00	45,00	0,455	100	19,3	21,0	2,90E-08	45	1
17	50,00	45,00	0,455	30	29,0	19,5	2,00E-09	90	1
18	50,00	45,00	0,455	30	29,0	19,5	2,00E-09	45	1
19	50,00	45,00	0,455	40	27,0	21,3	3,00E-09	90	1
20	50,00	45,00	0,455	40	27,0	21,3	3,00E-09	45	1
21	30,00	28,00	0,250	20	26,7	20,8	4,00E-06	90	0
22	30,00	28,00	0,250	20	26,7	20,8	4,00E-06	45	1
23	30,00	28,00	0,250	15	26,7	20,1	7,80E-07	90	1
24	30,00	28,00	0,250	15	26,7	20,1	7,80E-07	45	1
25	30,00	28,00	0,250	30	31,6	20,8	2,00E-07	90	1
26	30,00	28,00	0,250	30	31,6	20,8	2,00E-07	45	1
27	30,00	28,00	0,250	153	24,8	21,2	4,80E-10	90	1
28	30,00	28,00	0,250	153	24,8	21,2	4,80E-10	45	1
29	30,00	28,00	0,250	26	27,8	20,8	2,00E-07	90	1
30	30,00	28,00	0,250	26	27,8	20,8	2,00E-07	45	1

Tabela 5.4 – Dimensões de Elementos das Barragens Hipotéticas.

Tipo	H (m)	L _{Crista} (m)	H _{Fundação} (m)	L _{Filtro} (m)
1	6,00	3,00	2,00	1,00
2	30,00	10,00	20,00	1,00
3	50,00	12,00	20,00	1,00

Tabela 5.5 – Resultados do SEEP/W e do SLOPE/W para as Barragens Hipotéticas.

No.	FS _{Mon}	FS _{Jus}	i _{cr}	i _{máx}	FS _{Piping}
1	4,41	2,30	0,80	1,28	0,63
2	4,43	2,24	0,80	1,51	0,53
3	5,85	3,66	1,00	1,10	0,91
4	5,47	3,63	1,00	1,39	0,72
5	5,61	3,21	1,02	1,40	0,73
6	4,95	3,67	1,02	1,90	0,54
7	4,56	2,63	0,95	1,19	0,80
8	4,98	2,43	0,95	1,48	0,64
9	5,48	3,28	1,13	2,82	0,40
10	4,95	3,07	1,13	1,87	0,60
11	7,39	7,40	1,07	1,34	0,80
12	5,71	4,38	1,07	1,52	0,70
13	3,61	2,09	1,02	1,06	0,96
14	4,07	2,05	1,02	1,23	0,83
15	3,30	1,91	1,10	1,29	0,85
16	3,58	1,89	1,10	1,72	0,64
17	3,29	1,71	0,95	1,65	0,58
18	3,57	1,69	0,95	2,11	0,45
19	3,03	1,71	1,13	1,63	0,69
20	3,26	1,67	1,13	2,09	0,54
21	3,42	2,68	1,08	0,98	1,10
22	3,74	2,80	1,08	1,31	0,82
23	3,45	2,61	1,01	1,05	0,96
24	3,78	2,74	1,01	1,28	0,79
25	4,91	3,14	1,08	1,10	0,98
26	5,31	3,13	1,08	1,28	0,84
27	6,14	3,36	1,12	1,79	0,63
28	6,49	3,38	1,12	2,11	0,53
29	4,40	2,89	1,08	1,10	0,98
30	4,80	2,95	1,08	1,28	0,84

A partir desses resultados, as barragens hipotéticas foram avaliadas pela RNA treinada com o banco de dados da pesquisa, em que obtiveram-se previsões conforme a Figura 5.21 e a Tabela 5.6. Nota-se que 73% das previsões realizadas foram acuradas, sendo que em oito observações houve discrepâncias. Para as observações 4, 8, 10, 22, 24, 26 e 30, a RNA previu a não ocorrência de *piping*, em contraste com os resultados do *Geostudio* para as barragens hipotéticas. Para a observação 21, a RNA previu a ocorrência do fenômeno, sendo que não foi o constado no *Geostudio*.

No caso da observação 4, os valores de coesão efetiva, do ângulo de atrito efetivo, do peso específico saturado e da condutividade hidráulica foram retirados da observação 17 do banco de dados original, em que não foi constatado a ocorrência de *piping*. Pelo ângulo de inclinação do filtro serem similares, a RNA interpretou com a não ocorrência do fenômeno de acordo com o treinamento. No geral, nas observações em que a RNA não previu de maneira correta, nota-se que as simulações são de barragens com ângulos de inclinação do dreno vertical de 45°. Conforme discutido em tópicos anteriores, o ângulo de inclinação do dreno vertical é uma das variáveis mais importantes do banco de dados, em que quanto maior o seu valor, maior a ocorrência de *piping*. Dessa forma, a RNA mais uma vez analisou de maneira coerente com os dados de entrada.

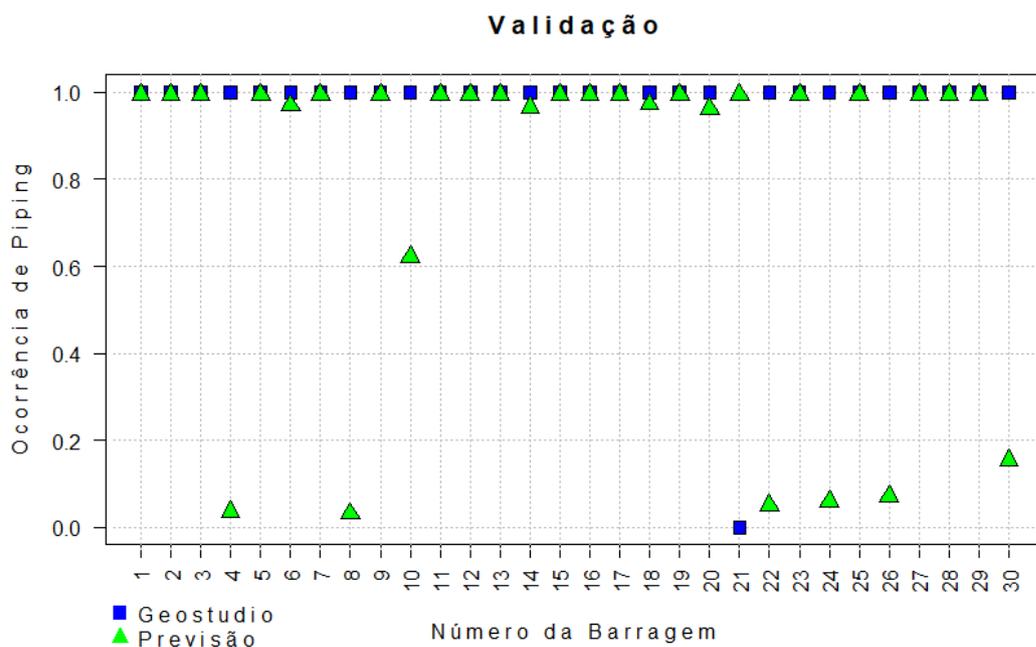


Figura 5.21 – Validação da RNA com Resultados do *Geostudio*.

Tabela 5.6 – Comparação de Resultados da Base de Dados com os Previstos.

No.	V10	RNA	Diferença	No.	V10	RNA	Diferença
1	1	0,994883	0,005117	16	1	0,994568	0,005432
2	1	0,994818	0,005182	17	1	0,994915	0,005085
3	1	0,994883	0,005117	18	1	0,972844	0,027156
4	1	0,037440	0,962560	19	1	0,994947	0,005053
5	1	0,994883	0,005117	20	1	0,961731	0,038269
6	1	0,969619	0,030381	21	0	0,994883	-0,994883
7	1	0,994884	0,005116	22	1	0,051235	0,948765
8	1	0,032541	0,967459	23	1	0,994939	0,005061
9	1	0,994883	0,005117	24	1	0,060812	0,939188
10	1	0,622403	0,377597	25	1	0,994883	0,005117
11	1	0,994883	0,005117	26	1	0,072318	0,927682
12	1	0,994816	0,005184	27	1	0,994883	0,005117
13	1	0,994604	0,005396	28	1	0,994553	0,005447
14	1	0,965403	0,034597	29	1	0,994883	0,005117
15	1	0,994883	0,005117	30	1	0,155021	0,844979

Após a validação da RNA, verifica-se que o modelo foi capaz de prever com 73% de acerto a ocorrência do fenômeno para o mecanismo de erosão regressiva. Enfatiza-se que os critérios adotados pela modelagem no *Geostudio* influenciam no gradiente máximo obtido, sendo assim, a validação foi realizada como um exercício no caso da aplicação da RNA em barragens de terra homogêneas hipotéticas.

6. CONCLUSÕES E SUGESTÕES

Neste capítulo, serão elencadas as principais conclusões acerca da pesquisa realizada, em que será dado destaque para os fatores limitantes e sugestões para trabalhos futuros.

A partir dos resultados apresentados, o objetivo geral deste trabalho foi atingido com êxito em que o algoritmo Boruta e os teste t de hipóteses de Welch foram utilizados para identificar os parâmetros chave de ocorrência de *piping* em barragens de terra homogêneas. Além disso, a rede neural artificial (RNA) foi elaborada com o intuito de prever a ocorrência do fenômeno, de acordo com banco de dados.

A consulta na literatura técnica da área permitiu demonstrar a relevância do *piping* em rupturas de barragens de terra homogêneas, conforme estudos anteriores de análises de séries históricas. Os parâmetros físicos que influenciam na ocorrência do fenômeno foram documentados e um conjunto de dados foi escolhido para o desenvolvimento deste trabalho.

O estudo contribui para a melhoria das ferramentas de previsão de *piping* e uma melhor compreensão do fenômeno em barragens de terra homogêneas. Os resultados podem ser aplicados a modelos de análises de *piping*, pois indicam as variáveis mais importantes para este evento: o ângulo de inclinação do dreno vertical da barragem e o diâmetro máximo efetivo das partículas de solo. Assim, a identificação das variáveis relevantes permite o desenvolvimento de modelos com menor exigência de dados de entrada. Para o modelo de rede neural artificial (RNA) proposto, foi realizada o treinamento com o banco de dados disponível e a validação com modelos de barragens hipotéticas no *Geostudio*. Os métodos propostos podem ser utilizados em estudos de outras falhas em diferentes tipologias de barragens, como as de rejeitos e as barragens zonadas, a partir do levantamento de dados e parâmetros necessários para as análises.

6.1. ETAPA 1 – IDENTIFICAÇÃO DAS VARIÁVEIS RELEVANTES

Este estudo foi capaz de aplicar dois métodos diferentes para análises estatísticas relacionadas com a previsão de *piping* em barragens de terra homogêneas. O conjunto de dados foi composto por nove variáveis associadas à estrutura da barragem, ao tipo de

solo, e ao reservatório. Estas variáveis foram analisadas para determinar a importância de cada uma para a ocorrência do fenômeno estudado. A ocorrência do *piping* foi introduzida nas análises como uma variável binária que indica se a barragem sofreu ou não do fenômeno.

Para a base de dados utilizada, o algoritmo de aprendizagem de máquina Boruta indicou que a variável mais importante é o ângulo de inclinação do dreno vertical, seguido pela condutividade hidráulica, o diâmetro máximo efetivo das partículas de solo, e a inclinação do talude a jusante.

O teste t de Welch, aplicado à mesma base de dados, confirmou a importância do ângulo de inclinação do dreno como o parâmetro principal para a ocorrência de *piping*. Contudo, o teste t demonstrou também que o diâmetro máximo efetivo das partículas de solo, a altura da barragem e o nível de água a montante resultaram em uma maior significância entre as variáveis testadas. Os resultados da análise estão diretamente relacionados com a base de dados e consideraram que os valores e parâmetros catalogados nos estudos anteriores são verdadeiros.

Os resultados obtidos pelos dois métodos, com o uso do tradicional teste t e o algoritmo moderno de aprendizagem de máquina foram distintos, o que reforça a discussão do uso de diferentes abordagens em análises de problemas complexos. Por um lado, o teste t é reconhecido na estatística, amplamente utilizado em vários setores, e pode ser programado manualmente, com maior controle das suas funções e das etapas de operação.

Por outro lado, o algoritmo de inteligência artificial pode ser adaptado para diversos bancos de dados por ter maior facilidade em se trabalhar com grandes volumes de dados (*big data*) e possuem maior facilidade de operacionalização. Além disso, a versatilidade dos algoritmos de aprendizagem de máquina permite com que o operador possa selecionar o algoritmo mais adequado para o tipo de dado e de resultado que se deseja gerar. Em análises estatísticas, recomenda-se a utilização de mais do que uma técnica para verificar a correlação entre os dados e a importância das variáveis de maneira complementar, sendo possível desta forma, uma melhor tomada de decisão.

6.2. ETAPA 2 – SIMULAÇÃO DA REDE NEURAL ARTIFICIAL (RNA)

Com a aplicação da técnica de redes neurais artificiais (RNA), foram obtidos resultados satisfatórios para a previsão da ocorrência de *piping* em barragens de terra homogêneas. Pela análise paramétrica, distintos cenários foram propostos com o intuito de permitir uma melhor compreensão do comportamento do modelo frente aos dados de entrada quanto ao número de neurônios por camada, número de camadas ocultas, funções de ativação e algoritmos de otimização de pesos.

A partir das análises, pode-se concluir que não necessariamente um número maior de neurônios por camada e o número de camadas ocultas pode trazer resultados mais acurados, sendo necessário uma rotina de testes com o intuito de identificar a arquitetura e os parâmetros que possuem as melhores previsões.

Com a RNA após o treinamento, a retroanálise dos resultados obtidos por Xue *et al.* (2014) foi realizada. Verificou-se que o algoritmo de otimização de pesos *resilient backpropagation* (RBP) e a função de ativação sigmoide, juntamente com a arquitetura da rede proposta, trouxeram previsões acuradas de acordo com a base de dados utilizada. Foram observados erros da ordem de 3%.

As simulações de barragens hipotéticas no *Geostudio*, com os módulos SEEP/W e SLOPE/W, permitiram avaliar a aplicação do modelo em casos fora do banco de dados de treinamento da RNA para o mecanismo de erosão regressiva. Sendo assim, por meio do método de fatias de Morgenstern-Price (1965), o fator de segurança foi calculado para os taludes das barragens hipotéticas, indicando a sua estabilidade, em que não foram constatados casos do fator de segurança inferiores a 1,5 nos taludes de montante e jusante.

Com o SEEP/W, foram utilizados critérios de projeto com base em estudos anteriores para as fundações e o dreno vertical, sendo estes mantidos entre as barragens hipotéticas para trazer consistência para as simulações. A avaliação de resultados permitiu avaliar que 73% das previsões feitas pela RNA foram satisfatórias. As discrepâncias ocorreram de maneira esperada por conta dos dados de treinamento da rede, o que indica mais uma vez a importância de uma maior quantidade de dados para que os resultados sejam mais acurados.

6.3. SUGESTÕES PARA TRABALHOS FUTUROS

Para sugestões em trabalhos futuros na continuação da linha de pesquisa sugere-se a busca por bancos de dados com um maior número de observações, para que o resultado obtido seja estatisticamente significativo. Pela falta de disponibilidade de dados confiáveis, este estudo pôde aplicar as técnicas de inteligência artificial em 23 casos reais. A identificação das variáveis relevantes, desta forma, é coerente apenas para essas vinte e três barragens analisadas.

Em relação ao treinamento da RNA, sugere-se a utilização de distintos algoritmos de otimização de pesos e funções de ativação para identificar os mais adequados ao considerar bancos de dados com um número maior de observações. Além disso, propõe-se um estudo de algoritmos de otimização de arquitetura de redes neurais ou de *frameworks*, para que o número de neurônios e o número de camadas ocultas possam ser escolhidos em múltiplas iterações de maneira automatizada.

Para a validação do modelo de RNA a partir do treinamento, sugere-se a modelagem de barragens reais em *softwares* de simulação de percolação e estabilidade, para que os resultados sejam condizentes com a realidade. O modelo de barragem deve incluir os elementos típicos encontrados nas seções dos barramentos, como *rip-raps*, muros de contenção, drenos de pé, tapetes drenantes, núcleos impermeáveis e *cut-offs*.

Por fim, frente à metodologia proposta de previsão de *piping* em barragens de terra homogêneas, recomenda-se a aplicação considerando distintas tipologias de barramentos, dentre as de concreto armado, zonadas, com face de concreto, com núcleo impermeável, dentre outras. Além do estudo de outras patologias que contribuem para as falhas de barragens, como galgamentos, instabilidades de taludes e liquefação. Em estudos de inteligência artificial, a veracidade e a consistência dos dados de entrada são fundamentais para que os resultados obtidos sejam coerentes. Com uma base de dados estruturada, é possível aplicar diferentes algoritmos para contribuir em tomadas de decisão com critérios objetivos e quantitativos.

REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA BRASIL (2019). **Mariana vê melhora econômica e aguarda Samarco após 4 anos da tragédia**. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2019-11/mariana-ve-melhora-economica-e-aguarda-samarco-apos-4-anos-da-tragedia>. Acesso em 02 de setembro de 2020.

AGÊNCIA NACIONAL DE ÁGUAS (2020). **Relatório de Segurança de Barragens de 2011 a 2019**. Brasília, DF, Brasil.

ASSIS, A.P., HERNANDEZ, H.M., & COLMANETTI, J.P (2003). **APOSTILA DE BARRAGENS**. Publicação G.AP-AA006/02. Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, Brasil.

BARROS, M. L., BARAN, C.T., & CAVALCANTE, A.L.B. (2020). **STATISTICAL ANALYSIS OF PIPING IN HOMOGENOUS EARTH DAMS**. Fourth International Dam World Conference. Laboratório Nacional de Engenharia Civil, Lisboa, Portugal.

BRASIL. Lei nº 12.334, de 20 de setembro de 2010. **Lei de Segurança de Barragens**. Brasília: Casa Civil. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Lei/L12334.htm> . Acesso em 4 de novembro de 2019.

BREIMAN, L. (2001). **Random Forests**. *Machine Learning*. Vol. 45. p. 5-32. 10.1023/A:1010950718922.

CASAGRANDE, A. (1937). **Characteristics of cohesionless soils affecting the stability of slopes and earth fills**. *J. Boston Society of Civil Engineers*, 23(1), p. 13-32.

CHEN, J., LI, K., TANG, Z., BILAL, K., YU, S., WENG, C., & LI, K. (2016). **A parallel random forest algorithm for big data in a spark cloud computing environment**. *IEEE Transactions on Parallel and Distributed Systems*, 28(4), p. 919-933.

CHEN, S. S., ZHONG, Q. M., & SHEN, G. Z. (2019). **Numerical modeling of earthen dam breach due to piping failure**. *Water Science and Engineering*, 12(3), p. 169-178.

CIELEN, D., MEYSMAN, A. D., & B. ALI, M. (2016). **Introducing Data Science: Big Data, Machine Learning, and more using Python Tools**. 1. pub. Shelter Island - Estados Unidos: Editor Manning Publication.

DALGAARD, P. (2008). **Introductory Statistics with R**. DOI: 10.1007/978-0-387-79054-1. Springer Science & Business Media.

DARCY, H. (1856). **Les Fontaines Publiques de la Ville de Dijon**. Dalmont, Paris, França.

DAS, B. M. (2019). **Advanced Soil Mechanics**. In Advanced Soil Mechanics. 5th ed. In CRC Press. <https://doi.org/10.1201/9781351215183>.

DAS, B.M., & SOBHAN, K. (2014). **Principles of Geotechnical Engineering**. 8th ed. In Cengage Learning. ISBN-13: 978-1-133-10867-2.

EMPRESA DE PESQUISA ENERGÉTICA (2019). **Balanco Energético Nacional de 2019. Relatório síntese / Ano Base 2018**. Ministério de Minas e Energia. Rio de Janeiro – RJ, Brasil.

ERTEL, W. (2017). **Introduction to Artificial Intelligence**. 2nd ed. Springer Science & Business Media.

FELL, R., MACGREGOR, P., STAPLEDON, D., BELL, G., & FOSTER, M. (2015). **Geotechnical Engineering of Dams**. 2nd Ed. In CRC Press.

FERDOS, F. (2016). **Internal Erosion Phenomena in Embankment Dams**. Tese (Doutorado em Engenharia Civil) – School of Architecture and Built Environment, KTH Royal Institute of Technology, Stockholm, Sweden, 74 p.

FOSTER, M., FELL, R., & SPANNAGLE, M. (2000). **The statistics of embankment dam failures and accidents**. Canadian Geotechnical Journal, 2000. 25: p. 1000-1024.

FOSTER, M., & FELL, R. (1999). **A framework for estimating the probability of failure of embankment dams by internal erosion and piping using event tree methods**. UNICIV Report No. R-377, School of Civil and Environmental Engineering, Univ. New South Wales, Sydney, Austrália.

FREDLUND, D. G., & MORGENSTERN, N. R. (1976). **Constitutive Relations for Volume Change in Unsaturated Soils**. Canadian Geotechnical Journal, 13(3), p. 261–276. <https://doi.org/10.1139/t76-029>

G1 (2019). **Barragem da Vale se rompe em Brumadinho, MG**. Disponível em: <<https://g1.globo.com/mg/minas-gerais/noticia/2019/01/25/bombeiros-e-defesa-civil-sao-mobilizados-para-chamada-de-rompimento-de-barragem-em-brumadinho-na-grande-bh.ghtml>>. Acesso em 16 de novembro de 2020.

G1 (2020). **Desastre Ambiental em Mariana, MG**. Disponível em: <<https://g1.globo.com/mg/minas-gerais/desastre-ambiental-em-mariana/>>. Acesso em 16 de novembro de 2020.

GAMELEIRA, E.L. & AMARAL, F.N.S. (2018). **Análise do fluxo de água no maciço de uma barragem de terra com filtro**. TCC (Graduação em Engenharia Civil) – Centro Universitário CESMAC, Maceió – AL, Brasil.

GEEKSFORGEEEKS (2019). **Effect of Bias in Neural Network**. Disponível em: <<https://www.geeksforgeeks.org/effect-of-bias-in-neural-network/>>. Acesso em 29 de novembro de 2019.

GEOSTUDIO (2012). Seepage Modeling with SEEP/W. An Engineering Methodology. Calgary, Alberta, Canadá.

GEOSTUDIO (2012). Stability Modeling with SLOPE/W. An Engineering Methodology. Calgary, Alberta, Canadá.

GÉRON, A. (2017). **Hands-On Machine Learning with Scikit-Learn**. 1st ed. O’Reilly Media. ISBN: 9781491962299.

GM METROLOGIA (2020). **Precisão x Exatidão**. Disponível em: <<https://gmmetrologia.com/precisao-x-exatidao/>>. Acesso em 29 de novembro de 2020.

GOLDEMBERG, J. (2000). **Pesquisa e desenvolvimento na área de energia**. São Paulo em Perspectiva, v. 14, n. 3, p.91-97, jul. 2000. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s0102-88392000000300014>.

GÜNTHER, F., & FRITSCH, S. (2010). **Neuralnet: Training of neural networks**. R Journal, 2(1), p. 30–38. <https://doi.org/10.32614/rj-2010-006>.

HARZA, L. F. (1935). **Uplift and seepage under dams in sand**. Trans. ASCE, 100(1), p. 1352–1385.

INTERNATIONAL COMMISSION ON LARGE DAMS (2019). **General Synthesis**. Disponível em: <https://www.icold-cigb.org/GB/world_register/general_synthesis.asp>. Acesso em 23 de novembro de 2019.

INTERNATIONAL COMMISSION ON LARGE DAMS (2019). **Dams' safety is at the very origin of the foundation of ICOLD**. Disponível em: <https://www.icold-cigb.org/GB/dams/dams_safety.asp>. Acesso em 23 de novembro de 2019.

JAMEL, A. A. J. (2018). **Investigation and Estimation of Seepage Discharge Through Homogenous Earth Dam with Core by Using SEEP/W Model and Artificial Neural Network**. Diyala Journal of Engineering Sciences, 11(3), p. 54-61.

KOENDERS, M. A. & SELLMEIJER, J. B. (1992). **Mathematical model for piping**. Journal of geotechnical engineering, 118(6), p. 943-946.

KURSA, M. B., & RUDNICKI, W. R. (2010). **Feature selection with the boruta package**. Journal of Statistical Software, 36(11), p. 1–13. <https://doi.org/10.18637/jss.v036.i11>.

MCCULLOCH, W. S., & PITTS, W. H. (1943). **A Logical Calculus of the Ideas Immanent in Nervous Activity**. Bulletin of Mathematical Biophysics, 5, p. 115–133.

MEIRELLES, F.S.C (2019). **Curso Segurança de Barragens, Módulo 1, Unidade 9**. Disponível:<https://capacitacao.ead.unesp.br/dspace/bitstream/ana/110/20/Unidade_9-modulo1.pdf> Acesso em: 29 de novembro de 2019.

MESSERKLINGER, S. (2013). **The Design of Filter Materials and their Importance in Geotechnical Engineering**. Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering, Paris, França.

MINISTÉRIO DE MINAS E ENERGIA (2019). **MME estabelece eixos estruturantes para o setor mineral**. Disponível em: <http://www.mme.gov.br/web/guest/pagina-inicial/outras-noticias/-/asset_publisher/32hLrOzMKwWb/content/mme-estabelece-eixos-estruturantes-para-o-setor-mineral>. Acesso em: 28 de outubro de 2019.

MOLINDER, G. (2016). **Internal erosion in the pervious foundation of an embankment dam: A case study on the Lossen dam**. Disponível em: <<http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A944358&dswid=1271>>. Acesso em 8 de dezembro de 2020.

MORGENSTERN, N.R., & PRICE, V.E., 1965. **The Analysis of the Stability of General Slip Surfaces**. Geotechnique, vol. 15, p. 79-93.

MUNDO EDUCAÇÃO (2019). **Neurônio: o que é, função, tipos, estrutura**. Disponível em: <<https://mundoeducacao.bol.uol.com.br/biologia/neuronios.htm>>. Acesso em 25 de novembro de 2019.

OKEKE, A. C. U., & WANG, F. (2016). **Hydromechanical constraints on piping failure of landslide dams: an experimental investigation**. Geoenvironmental Disasters, 3(1). <https://doi.org/10.1186/s40677-016-0038-9>.

PATHAK, M. (2018). **Feature Selection in R with the Boruta R Package**. Disponível em: <<https://www.datacamp.com/community/tutorials/feature-selection-R-boruta>>. Acesso em 6 de dezembro de 2020.

POLAMURI, S. (2017). **How the random forest algorithm works in machine learning**. Disponível em: <<https://dataaspirant.com/random-forest-algorithm-machine-learning/>> Acesso em 5 de dezembro de 2020.

RICHARDS, L.A. (1931). **Capillary conduction of liquids through porous mediums**. *Physics* 1 (5), p. 318–333.

ROSENBLATT, F. (1957). **The Perceptron - A Perceiving and Recognizing Automaton**. In Report 85, Cornell Aeronautical Laboratory. Cornell Aeronautical Laboratory.

ROTUNNO, A. F., CALLARI, C., & FROIIO, F. (2018). **A numerical approach for the analysis of piping erosion in hydraulic works**. In European Working Group on Internal Erosion. p. 159-167. Springer.

SELLMEIJER, J.B. 2006. **Numerical Computation of Seepage Erosion below Dams (Piping)**. Proceedings Third International Conference on Scour and Erosion, Amsterdam, The Netherlands, p. 596-601.

TAYFUR, G., SWIATEK, D., WITA, A., & SINGH, V. P. (2005). **Case study: Finite element method and artificial neural network models for flow through Jeziorsko earthfill dam in Poland**. *Journal of Hydraulic Engineering*, 131(6), p. 431–440. [https://doi.org/10.1061/\(ASCE\)0733-9429\(2005\)131:6\(431\)](https://doi.org/10.1061/(ASCE)0733-9429(2005)131:6(431))

TERZAGHI, K. (1931). **Underground erosion and the Corpus Christi Dam failure**. *Engineering News-Record*, 107, p. 90-92.

THE CONSTRUCTOR (2019). **Common Causes of Failures of Dam Structures**. Disponível em: <<https://theconstructor.org/structures/common-causes-failures-dam-structures/20858/>>. Acesso em 30 de novembro de 2019.

TOWARDS DATA SCIENCE (2019). **Everything you need to know about Neural Networks and Backpropagation — Machine Learning Easy and Fun**. Disponível em: <<https://towardsdatascience.com/everything-you-need-to-know-about-neural-networks->

and-backpropagation-machine-learning-made-easy-e5285bc2be3a>. Acesso em 20 de novembro de 2019.

VAN GENUCHTEN, M. T. (1980). **A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils**. Soil Science Society of America Journal, 44(5), p. 892–898. <https://doi.org/10.2136/sssaj1980.03615995004400050002x>.

WASSERMAN, L. (2004). **All of Statistics. A Concise Course in Statistical Inference**. In Design. <https://doi.org/10.1016/j.peva.2007.06.006>. Springer Science & Business Media.

WELCH, B. L. (1947). **The Generalization Of “Student”S’ Problem when several different population variances are involved**. Biometrika, 34(1-2), p. 28–35. DOI:10.1093/biomet/34.1-2.28.

XUE, X., YANG, X., & CHEN, X. (2014). **Estimating piping potential in earth dams and levees using generalized neural networks**. Acta geotechnica Slovenica. 11. p. 58-69.

YONGBIAO, L. (2012). **Prediction methods to determine stability of dam if there is piping**. Ieri Procedia, 1, p. 131-137.

ZHANG, L.M., XU, Y., & JIA, J.S. (2009). **Analysis of earth dam failures: A database approach**. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 3:3, 184-189, DOI: 10.1080/17499510902831759.

ZHANG, L. M., PENG, M., CHANG, D., & XU, Y. (2016). **Dam failure mechanisms and risk assessment**. John Wiley & Sons.