



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA CIVIL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

MARIA JÚLIA DE OLIVEIRA HOLANDA

**SOLOS COLAPSÍVEIS E EXPANSIVOS NO BRASIL: classificação da suscetibilidade
de ocorrência aplicando redes neurais artificiais**

Recife
2022

MARIA JÚLIA DE OLIVEIRA HOLANDA

SOLOS COLAPSÍVEIS E EXPANSIVOS NO BRASIL: classificação da suscetibilidade de ocorrência aplicando redes neurais artificiais

Tese apresentada ao Programa de Pós-graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Doutora em engenharia civil.

Área de concentração: Geotecnia.

Orientador: Prof. Dr. Silvio Romero de Melo Ferreira.

Coorientador: Prof. Dr. Samuel Franca Amorim.

Recife
2022

Catálogo na fonte
Bibliotecário Gabriel Luz, CRB-4 / 2222

H722s Holanda, Maria Júlia de Oliveira.
Solos colapsíveis e expansivos no Brasil: classificação da suscetibilidade de ocorrência aplicando redes neurais artificiais / Maria Júlia de Oliveira Holanda. 2022.
159 f: il.

Orientador: Prof. Dr. Silvio Romero de Melo Ferreira.
Coorientador: Prof. Dr. Samuel Franca Amorim.
Tese (Doutorado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Civil, Recife, 2022.
Inclui referências e apêndices.

1. Engenharia civil. 2. Redes neurais artificiais. 3. Colapso. 4. Expansão. I. Ferreira, Silvio Romero de Melo (Orientador). II. Amorim, Samuel Franca (Coorientador). III. Título.

UFPE

624 CDD (22. ed.) BCTG / 2023 - 30

MARIA JÚLIA DE OLIVEIRA HOLANDA

SOLOS COLAPSÍVEIS E EXPANSIVOS NO BRASIL: classificação da suscetibilidade de ocorrência aplicando redes neurais artificiais

Tese apresentada ao Programa de Pós-graduação em Engenharia Civil da Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, como requisito parcial para a obtenção do título de Doutora em engenharia civil. Área de Concentração Geotecnia.

Aprovada em: 18/11/2022

Orientador: Prof. Dr. Silvio Romero de Melo Ferreira - UFPE
Coorientador: Prof. Dr. Samuel Franca Amorim - UFPE

BANCA EXAMINADORA

participação por videoconferência
Prof.^a Dr.^a Maria Odete Holanda Mariano (examinadora interna)
Universidade Federal de Pernambuco

participação por videoconferência
Prof.^a Dr.^a Kátia Vanessa Bicalho (examinadora externa)
Universidade Federal do Espírito Santo

participação por videoconferência
Prof.^a Dr.^a Ana Patrícia Nunes Bandeira (examinadora externa)
Universidade Federal do Cariri

participação por videoconferência
Prof. Dr. Fernando Artur Nogueira Silva (examinador externo)
Universidade Católica de Pernambuco

participação por videoconferência
Prof. Dr. Joaquim Teodoro Romão de Oliveira (examinador externo)
Universidade Católica de Pernambuco

Dedico este trabalho a minha mãe Jandira e ao meu irmão Jefferson.

AGRADECIMENTOS

Á Deus, pela graça de ter concluído esse projeto, por ter me sustentado até aqui com saúde, coragem e disciplina.

Agradeço aos meus pais: Edinaldo Alves de Holanda e Jandira Costa de Oliveira, por serem meus maiores incentivadores na vida educacional, por não medirem esforços para me apoiar em cada novo desafio. Agradeço também ao meu irmão, Jefferson Henrique, por ser um exemplo de amor e dedicação a profissão, e sobretudo por sua garra diante da luta por sua saúde.

Agradeço ao meu noivo, Júlio Vinicius, pela torcida, incentivo, paciência, por acompanhar o passo a passo em revisões das ideias e do texto da tese, por sempre trazer soluções nos momentos em que me vi sem saída.

Agradeço imensamente aos meus orientadores, Professor Sílvio e Professor Samuel, definitivamente, esse trabalho é nosso! Cada um a seu modo e com sua especialidade na geotecnia não mediram esforços, técnicos e emocionais, para que eu mesma acreditasse que seria possível alcançar nossos objetivos. A escolha de um orientador é, sobretudo, um ato de confiança e entrega àqueles ideais.

Agradeço à Universidade Federal de Pernambuco – UFPE e todo seu corpo docente e administrativo, por todos os aprendizados e oportunidades, e por ter me possibilitado a chance de contribuir com a evolução da Engenharia.

Agradeço a banca avaliadora, que desde a qualificação vem contribuindo de forma brilhante com os resultados desta tese, sempre com cordialidade, gentileza e sabedoria.

Agradeço aos meus amigos de doutorado, em especial ao Grupo de Pesquisa de Solos Não Saturados – GNSat: Jesce, Larissa e Raiza, pelo trabalho em equipe na composição do banco de dados, pela parceria no desenvolvimento dos artigos, por me ouvirem e aconselharem nos momentos de angústia e pela torcida e vibração nas conclusões de cada etapa.

Agradeço a Dimas pela colaboração na codificação do programa e pela paciência diante de todos os ajustes e modificações solicitadas até a conclusão da tese.

Enfim, pela grande oportunidade de convívio, mesmo que virtual, com todos que acima citei e aqueles que de alguma forma contribuíram... meus sinceros agradecimentos.

Apesar dos avanços da ciência
E de suas inúmeras criações
Nada disso seria, enfim, palpável
Sem a mente da gente em incursões
É preciso gozar das descobertas
Com afinho traçar melhores metas
Pra uma vida mais plena, em unidade
Da ciência provém o aprendizado
E é por isso que deve o seu legado
Ser por todos os cantos espalhado.
(MONIQUE D'ÂNGELO LOPES, 2023)

RESUMO

Solos colapsíveis e expansivos são solos problemáticos na Engenharia Civil. Identificar, classificar e compreender o comportamento hidro geomecânico requer procedimentos de ensaios de laboratório e campo e a utilização de modelos computacionais que levam em consideração o estado tensional em que o solo se encontra e a que será submetido, relacionando a variação de volume devido à mudança de umidade. Redes Neurais Artificiais (RNA) constituem uma ferramenta importante para essa finalidade, através da correlação entre os preditores e das propriedades a serem estimadas. A pesquisa tem por objetivo identificar o grau de probabilidade e classificar a suscetibilidade de ocorrência de solos colapsíveis e expansivos no Brasil baseadas em variáveis geotécnicas e variáveis pedológicas, geológicas e climatológicas. A partir das informações do Banco de Dados de Solos Especiais (BANDASE) do Grupo de Pesquisa de Solos Não Saturados (GNSat) da UFPE, utilizando RNA através do *Neural Design*, as redes são desenvolvidas e tem seu desempenho analisado segundo testes de classificação binária (acurácia, taxa de erro, previsão, sensibilidade e especificidade). Três redes foram elaboradas. A primeira rede (PE04) criada a partir de 87 amostras de PE (53 treinamento, 17 seleção e 17 teste), conta com 4 variáveis de entrada (% de areia, % de argila, índices de plasticidade e atividade) e tem acurácia de classificação de 76,5%. A segunda rede (PE07) considera as mesmas 87 amostras com 7 variáveis de entrada (as 4 variáveis da rede PE04, além de clima, pedologia e geologia) e obtém uma acurácia de 88,2%. A terceira rede (BR03), desenvolvida com 393 amostras (237 treinamento, 78 seleção e 78 teste), utiliza 3 variáveis entrada (clima, pedologia e geologia) e tem acurácia de 89,7%. A generalização dos padrões de previsão das redes apresenta taxas de acurácia de 91,11% e 81,95% para as redes PE04 e BR03 e são aplicadas interpolando amostras dentro do mesmo domínio que foram desenvolvidas; enquanto a rede PE07, validada extrapolando o domínio com amostras do Nordeste e do Brasil, teve um decréscimo na taxa de acurácia: 65,5% e 56,7% respectivamente. Com a melhor rede (BR03) foi desenvolvido um programa para classificação de solos e elaborado um mapa probabilístico de ocorrência dos solos colapsíveis e expansivos, “disponibilizados gratuitamente” em ambiente *web*. As RNA PE04 e PE07 possibilitaram a produção de estimativas de identificação e classificação de solos colapsíveis e expansivos de Pernambuco com acurácia satisfatória, estabelecendo uma boa correlação entre as variáveis e a rede BR03 e ratificou a importância das variáveis de origem e formação do solo para classificar locais de propensão a esses fenômenos.

Palavras-chave: redes neurais artificiais; colapso; expansão.

ABSTRACT

Collapsible and expansive soils are problematic soils in Civil Engineering. Identifying, classifying and understanding the hydro-geomechanical behavior requires laboratory and field testing procedures and the use of computer models that take into account the tensional state in which the soil is and to which it will be subjected, relating the volume variation due to the change in moisture. Artificial Neural Networks (ANN) constitute an important tool for this purpose, through the correlation between the predictors and the properties to be estimated. The research aims to identify the degree of probability and classify the susceptibility of occurrence of collapsible and expansive soils in Brazil based on geotechnical variables and pedological, geological and climatological variables. Based on information from the Special Soils Database (BANDASE) of the Research Group on Unsaturated Soils (GNSat) at UFPE, using ANN through Neural Design, the networks are developed and their performance is analyzed according to binary classification tests (accuracy, error rate, prediction, sensitivity and specificity). Three networks were elaborated. The first network (PE04) created from 87 PE samples (53 training, 17 selection and 17 testing), has 4 input variables (% sand, % clay, plasticity and activity indices) and has classification accuracy of 76.5%. The second network (PE07) considers the same 87 samples with 7 input variables (the 4 variables of the PE04 network, in addition to climate, pedology and geology) and obtains an accuracy of 88.2%. The third network (BR03) developed with 393 samples (237 training, 78 selection and 78 testing), uses 3 input variables (climate, pedology and geology) and has an accuracy of 89.7%. The generalization of the prediction patterns of the networks presents accuracy rates of 91.11% and 81.95% for the PE04 and BR03 networks and are applied by interpolating samples within the same domain that they were developed; while the PE07 network, was validated with samples from the Northeast and Brazil, had a decrease in the accuracy rate: 65.5% and 56.7% respectively. With the best network (BR03) a soil classification program was developed and a probabilistic map of the occurrence of collapsible and expansive soils, “available for free” in a web environment. The ANN PE04 and PE07 allowed the production of estimates for the identification and classification of collapsible and expansive soils in Pernambuco with satisfactory accuracy, establishing a good correlation between the variables and the BR03 network and ratifying the importance of the variables of origin and soil formation for classify places prone to these phenomena.

Keywords: artificial neural networks; collapse; expansion.

LISTA DE FIGURAS

Figura 1 –	Representações Gráficas das estruturas de argilas silicatadas do tipo 1:1 (caulinita) e do tipo 2:1, não expansivas (ilita) e expansivas (vermiculita e montmorilonita).	31
Figura 2 –	Suscetibilidade de ocorrência de solos colapsíveis (a) e expansivos (b) no Estado de Pernambuco (carta interpretativa).	39
Figura 3 –	Indicação dos solos colapsíveis e expansivos estudados pelo mundo.	41
Figura 4 –	Registros de Solos Colapsíveis e Expansivos no Brasil.	43
Figura 5 –	Etapas metodológicas.	52
Figura 6 –	Proporção de ocorrências no BANDASE para 180 amostras do Nordeste.	53
Figura 7 –	Proporção de ocorrências no BANDASE para 526 amostras do Brasil.	54
Figura 8 –	Localização das amostras usadas para o desenvolvimento das RNA.	54
Figura 9 –	Análise de teste de RNA por Regressão (a) e Classificação (b).	55
Figura 10 –	Cartas de Atividade (Skempton, 1953) (a) e Plasticidade dos Solos Colapsíveis (b), Critério de Identificação de Solos Colapsíveis Segundo Reginatto e Ferreiro (1973) (c), Cartas de Atividade (Skempton, 1953) (d) e Plasticidade dos Solos Expansivos (e) e Critério de Identificação de Solos Expansivos Segundo Jimenes e Salas (1980) (f).	59
Figura 11 –	Histograma das variáveis escalares, segundo as 87 amostras da rede PE04: % de Argila (a), % de Areia (b), Índice de Plasticidade - IP (c) e Índice de Atividade - Ia (d).	64
Figura 12 –	Matriz de correlação entre as variáveis principais pelo <i>Neural Designer</i> .	66
Figura 13 –	Principais correlações entre as variáveis de entrada e saída.	66
Figura 14 –	Caracterização das amostras da rede PE07, segundo as variáveis categóricas: Climatologia, Geologia e Pedologia.	67
Figura 15 –	Principais correlações entre as variáveis de entrada e saída.	69
Figura 16 –	Correlação entre as variáveis categóricas da Rede BR03. Climatologia (a), Geologia (b) e Pedologia (c).	71

Figura 17 – Principais correlações entre as variáveis de entrada e saída.	71
Figura 18 – Processo de Elaboração das Cartas de Suscetibilidade.	78
Figura 19 – Gráfico do histórico dos erros do método <i>Quase-Newton</i> .	77
Figura 20 – Representação da melhor arquitetura da rede neural.	78
Figura 21 – Gráfico ROC.	79
Figura 22 – Gráfico de Ganho Acumulado.	80
Figura 23 – Valores Críticos da Variável Argila (a), Areia (b), Índice de Plasticidade -IP (c) e Índice de Atividade - Ia (d) para as Amostras de Solos Expansivos.	81
Figura 24 – Localização das 45 amostras de solos colapsíveis e expansivos do Nordeste (exceto Pernambuco, Alagoas e Maranhão) implementadas a rede PE04.	82
Figura 25 – Histograma de População.	84
Figura 26 – Histórico dos erros do método Quase-Newton.	85
Figura 27 – Representação da melhor arquitetura da rede neural.	85
Figura 28 – Gráfico ROC.	87
Figura 29 – Gráfico de Ganho Acumulado.	88
Figura 30 – Valores Críticos da Variável Argila (a), Areia (b), IP (c) e Ia (d) para as Amostras de Solos Expansivos.	89
Figura 31 – Localização das 284 amostras do Nordeste brasileiro validadas às cegas com a rede PE07.	90
Figura 32 – Histograma de População.	91
Figura 33 – Localização das 463 amostras do Brasil validadas às cegas com a rede PE07.	92
Figura 34 – Histograma de População.	93
Figura 35 – Histórico dos erros do método Quase-Newton.	94
Figura 36 – Representação gráfica da melhor arquitetura da rede neural.	95
Figura 37 – Gráfico ROC.	96
Figura 38 – Gráfico de Ganho Acumulado.	97
Figura 39 – Localização das 133 amostras do Brasil implementadas a rede BR03.	98
Figura 40 – Histograma de População.	99

Figura 41 – Indicadores de Desempenho das Redes nas Etapas de Testes (a) e Validação (b).	102
Figura 42 – Classificação de Amostra Expansiva para previsão de expansão pelo SECBRAApp.	104
Figura 43 – Classificação de Amostra Expansiva para previsão de colapso pelo SECBRAApp.	104
Figura 44 – Classificação de Amostra Colapsível para previsão de colapso pelo SECBRAApp.	105
Figura 45 – Classificação de Amostra 'Não-classificável' para previsão de colapso pelo SECBRAApp.	105
Figura 46 – Percentuais de ocorrência de solo colapsível no Brasil.	106
Figura 47 – Percentuais de ocorrência de solo expansivo no Brasil.	107
Figura 48 – Carta derivada de suscetibilidade de ocorrência de solos colapsíveis no Brasil.	108
Figura 49 – Carta derivada de suscetibilidade de ocorrência de solos expansivos no Brasil.	109

LISTA DE QUADROS

Quadro 1 –	Métodos indiretos e diretos de identificação de solos colapsíveis.	28
Quadro 2 –	Formação dos solos expansivos.	33
Quadro 3 –	Métodos de identificação dos solos expansivos.	35
Quadro 4 –	Características das Principais RNA Desenvolvidas.	62

LISTA DE TABELAS

Tabela 1 – Estatística das variáveis que compõem as amostras para o total desenvolvimento das Redes Neurais Artificiais. Rede PE04 (a), Rede PE07 (b) e Rede BR03 (c).	63
Tabela 2 – Índice de Classificação dos Solos pelo programa.	75
Tabela 3 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).	79
Tabela 4 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).	83
Tabela 5 – Matriz de confusão (a) e Métricas da Avaliação Binária (b).	86
Tabela 6 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).	90
Tabela 7 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).	93
Tabela 8 – Matriz de confusão (a) e Métricas da Avaliação Binária (b).	96
Tabela 9 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).	99
Tabela 10 – Síntese de Desempenho das RNA.	101

SUMÁRIO

1	INTRODUÇÃO	17
1.1.	Importância e Justificativa	18
1.2	Objetivos	21
1.2.1	Objetivo Geral	21
1.2.2	Objetivos Específicos	22
1.3	Hipótese	22
1.4	Delimitações da Tese	22
1.5	Estrutura da Tese	22
2	REFERENCIAL TEÓRICO	23
2.1	Solos Colapsíveis e Expansivos	23
2.1.1	Solos Colapsíveis	24
2.1.1.1	<i>Mecanismo de Colapso</i>	25
2.1.1.2	<i>Fatores de Ocorrência</i>	26
2.1.1.3	<i>Identificação Geotécnica de Solos Colapsíveis</i>	27
2.1.2	Solos Expansivos	29
2.1.2.1	<i>Mecanismo de Expansão</i>	30
2.1.2.2	<i>Fatores de Ocorrência</i>	31
2.1.2.3	<i>Identificação Geotécnica de Solos Expansivos</i>	33
2.1.3	Cartografia Geotécnica de Solos Colapsíveis e Expansivos	36
2.1.4	Ocorrência de Solos Colapsíveis e Expansivos no Mundo e no Brasil	40
2.2	Inteligência Artificial	44
2.3	Redes Neurais Artificiais	45
2.3.1	Áreas de Aplicação da Redes Neurais Artificiais em Engenharia Civil	46
2.3.2	Redes Neurais Artificiais Aplicadas a Geotecnia	47
2.3.3	Áreas de Aplicação na Colapsividade e Expansividade dos Solos	50
3	MATERIAIS E MÉTODOS	52
3.1	Apresentação do Banco de Dados	53
3.2	Desenvolvimento das Redes Neurais	55
3.2.1	Programa suporte – Neural Designer	55
3.2.2	Escolha de Variáveis e Arquiteturas das Redes Neurais Artificiais	56

3.2.3	Características e Correlações dos Dados de Entrada das Redes Neurais Artificiais	64
3.3	Estratégias de Treinamento das RNA	72
3.4	Aplicação dos Padrões das RNA	74
3.5	Elaboração de Programa de Classificação de Solos	74
3.6	Carta de Classificação de Solos Colapsíveis e Expansivos do Brasil	76
4	RESULTADOS	77
4.1	Rede Neural PE04	77
4.1.1	Composição da RNA PE04	77
4.1.2	Teste – Análise do Desempenho	78
4.1.3	Aplicação dos Padrões da Rede PE04 para Solos dos Estados do Nordeste	82
4.2	Rede Neural PE07	84
4.2.1	Composição da RNA PE07	84
4.2.2	Teste – Análise do Desempenho	86
4.2.3	Aplicação dos Padrões da Rede PE07 para Solos dos Estados do Nordeste	89
4.2.4	Aplicação dos Padrões da Rede PE07 para Solos do Brasil	91
4.3	Rede Neural BR03	94
4.3.1	Composição da RNA BR03	94
4.3.2	Teste – Análise do Desempenho	95
4.3.3	Aplicação dos Padrões da Rede BR03 para Solos do Brasil	97
4.4	Elaboração de Programa de Classificação dos Solos – Aplicação da equação da RNA BR03	103
4.5	Carta de Classificação de Solos Colapsíveis e Expansivos do Brasil	105
5	CONCLUSÕES E SUGESTÕES	110
5.1	Conclusões Finais	110
5.2	Sugestões Para Futuras Pesquisas	112
	REFERÊNCIAS	113
	APÊNDICE A – Redes Neurais Artificiais	128
	A.1. – Neurônio Artificial	129
	A.2. - Funções de Ativação	131
	A.3. - Escolha e Preparação das Variáveis	134

A.4. - Arquitetura da Rede	135
A.5. - Tipo de Aprendizado	138
A.6. - Modelagem da Rede Supervisionada Tipo MLP com BP	140
A.7. - Testes de Desempenho	143
APÊNDICE B - Resumo do Desenvolvimento das RNA	150
APÊNDICE C – Equações Matemáticas das RNA	152
APÊNDICE D – Interface do Programa SECBRApp	157

1 INTRODUÇÃO

Muitos edifícios, estruturas leves, rodovias, ferrovias, canais, tubulações, muros de contenção de terra, represas e pontes são levados a patologias ou ruínas devido ao comportamento inesperado dos solos colapsíveis e expansivos (SIMÕES *et al.*, 2006; NAKASHIMA *et al.*, 2014). Nesses casos, tensões maiores podem ser criadas quando a variação de volume de um material é impedida. O valor dessas tensões pode diminuir ou aumentar rapidamente quando alterações de volume são parcialmente permitidas. Tendo em vista o uso e ocupação do solo por obras civis, é recomendado que o solo seja avaliado quanto ao colapso e expansão sempre que esse seja considerado não saturado.

Alterações no nível da água subterrânea e infiltração de água ou umidade da superfície como resultado de precipitação sazonal ou absorção de água nas camadas mais secas ou ainda por águas servidas e fluídos contaminantes, podem alterar o conteúdo natural da água na camada do solo. Alguns solos argilosos mostram mudanças de volume (colapso e/ou expansão) nessas condições, assim como solos arenosos e siltosos sofrem colapso. Portanto, solos expansivos submetidos a baixos níveis de tensões podem apresentar uma elevação diferencial significativa e, solos colapsíveis submetidos a tensões mais elevadas, com baixo teor de umidade, podem causar sérios danos nas estruturas devido ao comportamento de colapso (ASHAYERI & YASREBI, 2009).

Os progressos na compreensão da natureza dos solos colapsíveis e expansivos, nos últimos 40 anos, trouxeram contribuições consideráveis que podem ser separadas em duas categorias: A primeira enfatiza o estudo teórico e a segunda compreende as análises da mineralogia e estrutura do solo, além de ensaios de laboratório. Estudos mais recentes como os de Barbosa (2019) e Magalhães e Soares (2021), com base na sucção e pressão osmótica do solo, apresentam propriedades de solos expansivos pouco conhecidas. Procedimentos mais modernos para a realização desses ensaios possibilitam o controle da sucção, permitindo uma descrição mais detalhada da ocorrência do fenômeno de expansão, proporcionando, por exemplo, técnicas de translação de eixos e de transferência de vapor aplicada aos solos. A segunda categoria de progressos, refere-se à performance dos solos expansivos com ênfase no critério de projeto e precauções construtivas para estruturas apoiadas neles (ENGEGRAUT, 2015).

No caso de solos colapsíveis, os resultados de ensaios edométricos ou de campo (como cone ou expansocolapsômetro) são usados para estimar o colapso admitindo a hipótese de inundação extrema, sendo a maioria dessas estimativas realizadas com a implementação de modelos constitutivos em programas de cálculo (ALONSO *et al.*, 1990; FUTAI *et al.*, 2001; SILVA

FILHO *et al.*, 2010). Entretanto, um procedimento simplificador também é usual, no qual o deslocamento vertical é estimado através da multiplicação das deformações de colapso obtidas em ensaios edométricos pela espessura da camada envolvida no processo de colapso, determinando-se as tensões no solo sob a fundação por meio de equações da teoria da elasticidade.

Estimar o comportamento de tais solos é muito complexo e ter a capacidade de prever a expansão e o colapso de um solo, durante a etapa de investigação, contribuiria substancialmente na elaboração do projeto de fundações, a fim de diminuir a possibilidade de ocorrência das deformações associadas ao umedecimento dos solos. Nessa perspectiva, várias abordagens são desenvolvidas, além dos ensaios laboratoriais, com métodos de identificação diretos e indiretos (ABIODUN E NALBANTOGLU, 2015; TEIXEIRA, 2019), têm-se também os métodos estatísticos (SAVINO, 2017), e de inteligência artificial com a aplicação de Redes Neurais Artificiais (RNA) interpretando o colapso e a expansão dos solos.

A pesquisa apresenta parte do conhecimento de redes neurais para identificação de solos colapsíveis e expansivos, baseadas em variáveis de entrada geotécnicas, ampliando essa identificação para um grau de probabilidade de classificação de solos colapsíveis e expansivos do Brasil, com base em características de origem e formação do solo contribuindo no avanço de metodologias a partir de RNA, segundo a compreensão dos modelos de rede desenvolvidos.

1.1 Importância e Justificativa

Para evidenciar a importância dos solos expansivos na engenharia geotécnica, Murthy (2003) posiciona os danos provocados por estes solos, em terceiro lugar dentre as seis catástrofes naturais de maior risco do mundo, sendo elas: I) terremotos; II) escorregamentos; III) solos expansivos; IV) ciclone; V) furacão e VI) enchentes. Para Jones e Holtz (1974), o dano causado pelos solos expansivos em edificações e pavimentos é maior que os danos causados por outros desastres naturais incluindo terremotos e inundações. De acordo com NYS DHSES (2014) estima-se que os EUA tenham um custo de aproximadamente US\$ 300 milhões/ano devido a danos causados pela expansão dos solos. Entretanto, para o Brasil não se tem dados registrados que permitam, em uma análise semelhante, contabilizar os custos - compondo um banco de dados unificado nacionalmente, tampouco um tratamento estatístico desses dados, de modo a gerar conhecimento acerca dos prejuízos, como os citados em outros países, ou a posição na classificação de riscos geológicos.

Nos solos colapsíveis além das alterações na própria estrutura do solo, os impactos oriundos do processo de colapso manifestam-se, de forma mais clara, nas obras civis apoiadas sobre estas camadas, especialmente em prédios, e secundariamente, tubulações, canais e pavimentos. Os projetos de fundações, em solos potencialmente colapsíveis, devem adotar técnicas de melhoramento, além de medidas preventivas quanto à possibilidade de vazamento de esgotos e infiltração de águas pluviais.

A maior parte dos impactos diretos que ocorrem como consequências de processos de colapso de solos são perdas econômicas, visto não se ter registro, no Brasil e no exterior, de danos provocados por colapso que tenham ocasionado perda de vidas. Os custos mais comuns, relativos à recuperação das obras danificadas, em consequência da ocorrência de colapso, são os danos em alvenaria e dispositivos de fechamento, além dos reparos necessários nas redes de utilidade, especialmente água e esgoto (LOLLO, 2008).

Pesquisas têm sido realizadas por Vargas *et al.* (1989), Ferreira e Teixeira (1989), Ferreira (1995), Ferreira *et al.* (1998a), Ferreira *et al.* (1998b), Justino da Silva (2001, 2005) e Burgos e Campos (2002) tanto nas microrregiões do semiárido nordestino, onde as precipitações pluviométricas são baixíssimas na maioria dos meses do ano e elevadas nos demais, como também em solos de áreas litorâneas, onde as precipitações são, em geral, elevadas (BASTOS, 1994; JUSTINO DA SILVA, 2001; CAVALCANTE *et al.*, 2006). Essas pesquisas embasam e enriquecem o banco de dados de solos colapsíveis e expansivos possibilitando uma melhor caracterização das amostras de solos pelas mais variadas categorias, e sob quais os principais aspectos são determinados os padrões de comportamento dessas amostras.

Na identificação de solos colapsíveis e expansivos, alguns cuidados na fase de elaboração do projeto podem diminuir a possibilidade de ocorrência das deformações associadas ao umedecimento. Como esses solos estão presentes em praticamente todo o mundo e têm gerado grandes prejuízos, principalmente em regiões onde a evapotranspiração anual excede a precipitação, que são as zonas áridas, há um crescente interesse em aplicar técnicas que minimizem a diferença entre a previsão das medições oriundas de ensaios de campo e os resultados de modelos numéricos.

Um dos principais objetivos de uma caracterização geotécnica do solo é classificar várias camadas de solo para interpretação do subsolo. A investigação do local gera séries de dados estratigráficos, em intervalos contínuos, que são agrupados e então classificados com base na experiência e nas relações empíricas, respectivamente. A classificação do solo pode ser

correlacionada com seu comportamento e tem sido utilizada por diversos estudos para prever as propriedades de engenharia.

Nesse aspecto, as Redes Neurais Artificiais vêm apresentando resultados promissores na classificação de padrões comportamentais de solos, possibilitando conjecturar que o reconhecimento dessas características intrínsecas pode ser obtido de forma competente, a partir de variáveis geotécnicas que todo programa de sondagem traz, de baixo custo e de informações disponíveis gratuitamente para todo o Brasil dos condicionantes pedogenéticos.

As RNA, em seu sistema de aprendizado supervisionado, abordam dois conjuntos de dados previamente conhecidos, de variáveis de entrada e de saída correlatas para determinar a estrutura e os parâmetros desse modelo, a partir dos quais pode ser feito o treinamento. De acordo com Leite (2019) dentro dos seus benefícios estão flexibilidade, adaptabilidade, capacidade de generalizar e aprender um problema por meio dos estímulos externos e internos e capacidade de produzir uma resposta para problemas lineares e não-lineares. Outras características importantes são: a capacidade de auto-organização e de processamento temporal.

Segundo as pesquisas de Ikizler *et al.* (2010) e Merouane e Mamoune (2018) além de oferecerem uma excelente capacidade de modelar problemas não lineares com múltiplas variáveis, as RNA produziram previsões com grau de correlações em seus resultados acima de 90% e sobretudo, apresentaram a vantagem de poder ser atualizadas imediatamente à medida que novos resultados são acumulados, conforme os dados tornem-se disponíveis, com novos modelos de treinamento, os resultados de previsão e a correlação desses tendem a melhorar. Esses fatores se combinam para tornar as RNA uma poderosa ferramenta de modelagem de engenharia geotécnica.

A solução de problemas através de RNA tem grande relevância para a academia e para a indústria, tendo em vista a possibilidade de extrair correlações dos dados de entrada e saída das redes neurais sem subjetividade de especialista, possibilidade de trabalho com múltiplas variáveis - gerando funções e correlacionando a bom termo e capacidade de generalizar a informação aprendida para novos casos, cujo resultado se desconhece previamente, é sua principal função, e um dos atrativos principais desta ferramenta.

Assim, a identificação e classificação da suscetibilidade de ocorrência de solos colapsíveis e expansivos utilizando RNA, constitui uma contribuição técnico-científica relevante para o planejamento do uso e ocupação do solo.

1.2 Objetivos

1.2.1 Objetivo Geral

Classificar a suscetibilidade através do valor da probabilidade de ocorrência de um solo ser colapsível ou expansivo no Brasil, baseadas em variáveis geotécnicas, climatológicas, geológicas e pedológicas, utilizando RNA.

1.2.2 Objetivos Específicos

- a) Contribuir com a ampliação de informações geotécnicas de solos colapsíveis e expansivos, do Banco de Dados de Solos Especiais (BANDASE) desenvolvido pelo Grupo de Pesquisa de Solos Não Saturados da UFPE - GNsat;
- b) Identificar através de análise multivariada de componentes principais e desempenho das Redes Neurais Artificiais elaboradas a partir de parâmetros geotécnicos, climatológicos, geológicos e pedológicos responsáveis pelo comportamento de colapso e expansão dos solos;
- c) Definir um modelo de redes neurais de classificação, objetivo e suficientemente geral que torne possível aplicação de seus padrões nos âmbitos estaduais (PE), regionais (NE) e nacionais (BR), levando em consideração os diferentes ambientes geológico/pedológicos/geotécnicos e os dados normalmente obtidos na caracterização do solo;
- d) Disponibilizar em programa e mapa em portal da internet, para a comunidade técnica, o modelo matemático de Inteligência Artificial – Redes Neurais, como ferramenta de auxílio no estudo geotécnico, bem como o mapa probabilístico de ocorrência desses fenômenos.

1.3 Hipótese

É possível identificar e classificar o potencial de suscetibilidade de ocorrência de solos colapsíveis e expansivos, utilizando condicionantes climatológicos, geológicos, pedológicos e geotécnicos por meio de Redes Neurais Artificiais.

1.4 Delimitações da Tese

O desenvolvimento da presente pesquisa conta apenas com amostras de solos colapsíveis e expansivos para a elaboração das Redes Neurais Artificiais, bem como para a compreensão das ocorrências desses fenômenos nos solos do Brasil.

Não foi obtido resultados quanto ao percentual de colapso, percentual de expansão livre ou a pressão de expansão.

A classificação de suscetibilidade a ocorrência de colapso ou expansão não está relacionada ao grau de colapso ou expansão, e sim a probabilidade de ocorrência desses fenômenos.

As redes classificatórias não foram treinadas com amostras de solos não colapsíveis ou não expansivos.

Não houve amostras de outros países além do Brasil durante o desenvolvimento das redes, ou na etapa de aplicação dos parâmetros.

1.5 Estrutura da Tese

Esta tese está estruturada em cinco capítulos. O primeiro faz as considerações iniciais sobre o tema, define o problema, apresenta a justificativa, os objetivos da pesquisa e o arranjo geral da tese. O Capítulo 2 proporciona uma breve revisão bibliográfica sobre os solos expansivos e colapsíveis, inteligência artificial e sua aplicação em geotecnia. O Capítulo 3 descreve os procedimentos utilizados na simulação numérica. São apresentadas as características do banco de dados, bem como o trabalho estatístico utilizado, uma breve descrição sobre o programa utilizado para desenvolver as redes neurais, *Neural Designer*; e as modelagens e iterações feitas para prever e determinar a característica de colapsividade e expansividade dos solos. E os resultados obtidos utilizando as RNA são apresentados no Capítulo 4. Por fim, as conclusões e as sugestões para futuras pesquisas são apresentadas no Capítulo 5.

2 REFERENCIAL TEÓRICO

São abordados aqui os principais temas relacionados aos solos colapsíveis e expansivos como: Conceituação, classificação, mecanismos de colapso e expansão, fatores internos e externos de ocorrência, identificação por métodos diretos e indiretos e cartografia desses solos, bem como temas relacionados a Inteligência Artificial como: Conceituação de uma rede neural artificial, evolução histórica, processos de treinamento, seleção e teste de desempenho do modelo, bem como aplicações dessa tecnologia à engenharia civil, à geotecnia e à expansividade de solos.

2.1 Solos Colapsíveis e Expansivos

O solo é um dos materiais de construção mais utilizados pela humanidade, entretanto o comportamento deste material varia de acordo com a origem, tipologia da construção, clima e meio ambiente. Segundo Latifi *et al.* (2017) em algumas regiões do mundo ocorre escassez de solos adequados para obras de engenharia, seja como material de construção, seja como material de fundação para estruturas; em virtude disso, frequentemente é necessário melhorar as características do solo existente ou aumentar a capacidade de reconhecimento de suas características a fim de contribuir para os projetos de Engenharia Geotécnica.

Solos geotecnicamente problemáticos estão sendo ocupados, ultimamente, com mais frequência em áreas urbanas, as suas principais características são a baixa ou alta permeabilidade, baixa resistência ao cisalhamento e alta compressibilidade. Além disso, solos especiais em geral, podem apresentar uma variação espacial considerável resultante do processo de formação, fazendo com que ocorram mudanças de comportamento geomecânico em um mesmo depósito (HALLAL, 2003). O problema das obras situadas em solos colapsíveis e expansivos é recorrente em obras que cruzam diversos terrenos (lineares) e/ou hidráulicas (pois carregam água, elemento deflagrador de suas variações volumétricas), como canais de transposição, dutovias, caixas d'água, ou obras que podem sofrer grandes prejuízos como estradas, aeroportos e grandes conjuntos habitacionais, refinarias ou estaleiros.

Alguns solos não saturados, quando aumentam a quantidade de água em seus vazios, ou quando impostos a um carregamento e posteriormente umedecidos, apresentam uma variação volumétrica. Por isso, seu comportamento depende das variações de umidade sazonais. Podendo apresentar características de compressibilidade, de expansão e de plasticidade em decorrência do aumento da umidade; e contração com ressecamento e redução de umidade.

Com isso, as edificações de engenharia em solos colapsíveis e/ou expansivos podem manifestar distorções provocadas por estes solos que vão desde fissuras, que determinam a paralização no todo ou em parte do uso da construção, levando ao Estado Limite de Serviço – ELS; até mesmo a ruína da obra, com indícios de comprometimento da durabilidade da estrutura, não respeitando as condições especificadas para o uso normal da construção, que por sua ocorrência, repetição ou duração, atingem o Estado Limite Último - ELU.

2.1.1 Solos Colapsíveis

Os solos colapsíveis são geralmente compostos de minerais estáveis em presença de água como caulinita e quartzo. Tal fenômeno é tradicionalmente explicado pela destruição parcial do agente cimentante do solo, tendo como causas primárias a molhagem e a existência de uma pressão-limite. Todo solo suscetível ao colapso possui uma estrutura aberta ou porosa (DUDLEY, 1970; VARGAS 1973), potencialmente instável e parcialmente saturado (CLEMENCE & FINBARR, 1981), com baixa coesão e granulometria bem distribuída.

O colapso estrutural do solo se caracteriza por ser rápido e capaz de provocar sérios danos em construções, especialmente as de pequeno porte, em que as tensões de trabalho nas fundações variam entre 100 e 200 kPa, por essas terem suas fundações apoiadas nas camadas mais rasas dos solos - onde se encontra a camada ativa. A situação de risco mais comum é aquela em que as fundações são inundadas, seja por vazamento da rede de esgotos, seja pela percolação de águas de chuvas em casas desprovidas de passeios laterais, levando o solo a um grau de saturação crítico (superior a 60% até provavelmente 85%) e, em seguida, a uma drástica diminuição de volume (TEIXEIRA, 2019).

Entretanto, Teixeira (2019) alerta para a existência das tensões-limites, inferior e superior, abaixo e acima das quais a inundação do solo não produz deslocamentos verticais por colapso. Embora se tenha conhecimento de ocorrência de colapso em aterros de barragens, a compactação é, em princípio, um procedimento eficaz na sua prevenção, posto que diminui o índice de vazios e elimina a causa de metaestabilidade do solo, antecipando-se à destruição dos vínculos cimentícios dos grãos sólidos.

2.1.1.1. Mecanismo de Colapso

O mecanismo de colapso difere completamente do processo clássico de adensamento. Para ocorrer o fenômeno de adensamento em solos, o acréscimo de pressão neutra, por ação de sobrecarga nos interstícios do arcabouço sólido do solo, é dissipado por ação da drenagem. Já o fenômeno do colapso pode se manifestar por ação simples ou combinada de dois mecanismos:

- (1) Aumento do grau de saturação ao exceder o valor crítico inferior e,
- (2) Tensões que excedem o valor limite (sobrecarga ou peso próprio).

O aumento do teor de água de um solo pode dar origem, em condições de tensão total constante, a deformações volumétricas compressivas conhecidas como deformações por colapso. O mecanismo responsável por estas deformações é a redução das forças capilares inter partículas associadas ao acréscimo do grau de saturação (redução da sucção matricial). As deformações por colapso podem ocorrer em qualquer tipo de estrutura geotécnica (fundações, aterros, escavações, barragens de aterro, túneis) e provocar alterações significativas no campo de tensões e/ou deformações. Erroneamente, este fenômeno, descrito na literatura como hidro compactação de solos colapsíveis, é frequentemente atribuído à perda da coesão por dissolução das ligações (cimentação) entre partículas (WALTHAM, 1994).

A pressão isotrópica presente na água intersticial, que é consequência das condições físico-químicas, responsável pelo sistema água-solo absorver ou perder água, é denominada de sucção. Ela pode ser dividida em duas parcelas: matricial e osmótica (MARINHO, 1997).

A sucção matricial tem sido reconhecida como a mais importante fonte de sucção a interferir no comportamento mecânico de solos não saturados. Dessa forma, é necessário que se considere a sucção como uma variável independente, suficientemente capaz de governar a permeabilidade, a variação de volume e a resistência ao cisalhamento dos solos (LODI E LOLLO, 2008). Já a sucção osmótica é frequentemente desprezada sob o argumento de que suas variações são pequenas e o impacto dessas variações em problemas típicos de engenharia não é relevante. Mas, qualquer problema que envolva alteração da concentração de sais no solo deve ser examinado, considerando com atenção as sucções osmóticas.

A relação existente entre a umidade e a sucção nos solos, é denominada de curva característica de sucção ou curva de retenção de água. Fredlund e Rahardjo (1993) explicam que, à medida que decresce o grau de saturação do solo, a distribuição das fases sólida, líquida e de ar altera-

se em conjunto com o estado de tensões no solo. Para cada solo existe uma relação entre a taxa na qual ocorrem as variações nas tensões com a área de contato água-partícula. Fredlund e Xing (1994) ressaltam que existem dois pontos fundamentais da curva característica solo/água que são relevantes: a) o valor de entrada de ar, sucção em que o solo começa a dessaturar e o ar começa a entrar nos maiores poros; e b) a sucção residual, a partir da qual é necessário um grande incremento na sucção para que seja possível retirar água do solo. Geralmente, a tendência é que o valor de entrada de ar aumente quanto maior o teor de finos, por possuir poros menores.

De acordo com Reichardt (1985), o principal fator que influencia na curva de retenção dos solos é a textura, pois ela indica a área de contato entre a água e as partículas sólidas e, além disso, influencia na organização das partículas e na distribuição dos poros. As forças iterativas entre as partículas sólidas e o fluido percolante dependem, dentre outros fatores, da superfície específica do argilomineral presente no solo. Quanto maior a superfície específica maior será a área que o solo disponibilizará para interagir com o fluido percolante.

2.1.1.2. Fatores de Ocorrência

De acordo com Ferreira *et al.* (2008), os solos de comportamento colapsível podem ser encontrados em:

- solos não saturados;
- áreas onde a evapotranspiração excede à precipitação;
- solos de estrutura porosa;
- solos com partículas interligadas por argila, óxido de ferro, alumínio ou carbonatos;
- solos de origem recente e de fácil drenagem;
- regiões tropicais úmidas com lixiviação dos horizontes superficiais;
- regiões de alternância de estações secas e chuvas intensas e concentradas.

Um solo estará propenso à ocorrência de colapso desde que ele possua uma estrutura macro porosa, com equilíbrio metaestável entre as partículas maiores. A estrutura deste solo, em particular, é mantida pela presença de algum vínculo capaz de proporcionar uma resistência adicional temporária.

Quanto ao tipo de material, não há uma faixa granulométrica específica que enquadre os solos colapsíveis. Na maioria dos casos, os solos são caracterizados por estruturas fofas, com granulometria variando de silte a areia fina, geralmente uma mistura de areia fina, silte e argila, com predominância do primeiro. Todavia, há exemplos na literatura de solos reconhecidamente colapsíveis predominantemente argilosos.

Os solos de comportamento colapsível variam tanto em textura quanto em estrutura. As maiores constatações são nos depósitos eólicos, aluviais, coluvionares, solos vulcânicos, solos compactados no ramo seco e solos residuais, sendo a maior incidência nos depósitos eólicos, em especial os *loess*. Dentro deste universo, os solos colapsíveis têm sido encontrados em quase todas as regiões do mundo.

2.1.1.3. Identificação Geotécnica de Solos Colapsíveis

Os ensaios e critérios para identificação do colapso e das características de um solo colapsível em laboratório possibilitam estudar parâmetros específicos (como a composição química e o arranjo estrutural dos solos) e a suscetibilidade ao colapso, de modo a auxiliar na caracterização do colapso com ensaios que retratam condições ideais. Já a aplicação dos ensaios e dos critérios de campo, direta ou indiretamente, buscam a estimativa do potencial de colapso em escala real, através da análise geral do comportamento do solo na umidade natural e quando submetido aos efeitos da saturação.

A depender do critério de identificação, os métodos podem ser classificados em dois grupos: métodos diretos e métodos indiretos. Os métodos indiretos são aqueles que se utilizam os índices físicos e limites de consistência, ou parâmetros ligados à textura de simples obtenção em ensaios de laboratório e de campo, para identificar a potencialidade ao colapso estrutural, sendo de informação orientativa. Os métodos diretos baseiam-se na medida do potencial de colapso do solo. O Quadro 1 resume alguns critérios de identificação, de acordo com a classificação de Ferreira (1995).

Quadro 1 - Métodos indiretos e diretos de identificação de solos colapsíveis

MÉTODOS	Subdivisão	BASE PARA DEFINIÇÃO DO CRITÉRIO	REFERÊNCIAS BIBLIOGRÁFICAS
INDIRETO	Indicativos	Microscopia eletrônica de varredura	Collins e McCown (1974), Wolle <i>et al.</i> (1978), Derbyshire e Mellors (1988).
	Orientativos	Pedologia Ensaio expeditos	Ferreira (1990), Ferreira (1993), Arman e Thornton (1972), Jennings e Knight (1975).
	Qualitativos	Índices físicos	Ensaio de campo – cone
Ensaio SPT-T			Décourt e Quaresma Filho (1944)
DIRETO	Avaliativos	Ensaio edométrico duplo	Reginato e Ferrero (1973)
	Quantitativos	Ensaio edométrico simples	Bally <i>et al.</i> (1973), Jennings e Knight (1975), Vargas (1978), Lutenecker e Saber (1988)
		Ensaio de campo - Expansocolapsômetro	

Fonte: Adaptado por Mendonça Neto (2011) e Ferreira (2008).

Os métodos indiretos, baseados em ensaios de laboratório, buscam relacionar o índice de vazios na umidade, equivalente ao limite de liquidez, com o índice de vazios inicial do solo; ou utilizam o índice de consistência. Outros relacionam características granulométricas do solo, diretamente ou associadas a outros índices físicos. Praticamente todas as propostas desses métodos são qualitativas, baseadas nos índices físicos ou nos limites de Atterberg, tornando o uso desses métodos restrito a solos particulares (FUTAI, 2000).

Já para os métodos indiretos, baseados em ensaios de campo, uma característica observada em várias formações das regiões sudeste e centro-oeste são os baixos valores de N_{SPT} , entretanto nenhum autor assume algum valor de N_{SPT} como um critério de identificação geral, tornando esse parâmetro restrito a algumas formações. Com isso, não existe um critério consolidado para identificação de solos colapsíveis por métodos indiretos de ensaios de campo, a não ser as peculiaridades já reconhecidas em solos colapsíveis de cada região.

A avaliação da colapsibilidade de um solo pelos métodos diretos, por meio de ensaios edométricos e do expansocolapsômetro, tem a vantagem de levar em consideração as tensões atuantes e quantificar o potencial de colapso, sendo extensível a qualquer formação e tipo de solo. Tal caracterização é possível a partir da realização de ensaios de adensamento para observar o comportamento do solo, além de ferramentas complementares de análise, como o ensaio de sedimentação com e sem o uso de defloculante, o ensaio de desagregação, a curva característica solo-água e a distribuição de poros.

2.1.2 Solos Expansivos

Os solos expansivos são solos especiais, cujo comportamento não segue os padrões da geotecnia clássica, como solos colapsíveis e por isso são problemáticos para a Engenharia Geotécnica. É importante evitá-los quando da construção de obras de engenharia, como estruturas de fundações e pavimentos, embora seja possível, tratar, conviver e diferenciar os enfoques antes e após o projeto.

O fato de se conhecer o potencial expansivo do solo previamente ao projeto tem se mostrado fundamental para evitar prejuízos financeiros, por isso tem-se estudado o comportamento desses solos e a proposição das mais variadas técnicas para identificá-lo, além da estimativa de seus parâmetros de caracterização, segundo modelos estatísticos, como por exemplo, a variação volumétrica, que pode chegar a valores superiores a 30%, quando se altera o teor de água no solo (QI; VANAPALLI, 2015). Os solos expansivos apresentam elevada superfície específica, característica que condiz com a alta capacidade de trocas catiônicas (CTC) e com baixa permeabilidade deste material, essa característica é devida a natureza dos argilominerais componentes dos solos expansivos (CHRISTOPHER; CHIMOBİ, 2019). As mudanças de volume em nível microestrutural, ocorridas nos argilominerais, acarretam mudança de volume e reorganização do esqueleto do solo em escala macroestrutural, produzindo o efeito da expansão (MONROY; ZDRAVKOVIC; RIDLEY, 2015).

Nos períodos de estiagem, se encontram geralmente com uma sucção elevada, a qual confere uma alta resistência, dificultando possíveis trabalhos de escavação. Porém, com o aumento da umidade esses solos apresentam, muitas vezes, expressivos valores de expansão (CAVALCANTE *et al.*, 2006).

2.1.2.1 Mecanismo de Expansão

A expansibilidade é a variação volumétrica dos minerais argilosos devido à variação do teor em água. Os mecanismos de expansão podem ser mecânicos ou físico-químicos. A combinação desses mecanismos interfere diretamente na mudança de volume dos solos potencialmente expansivos (FERREIRA, 1995).

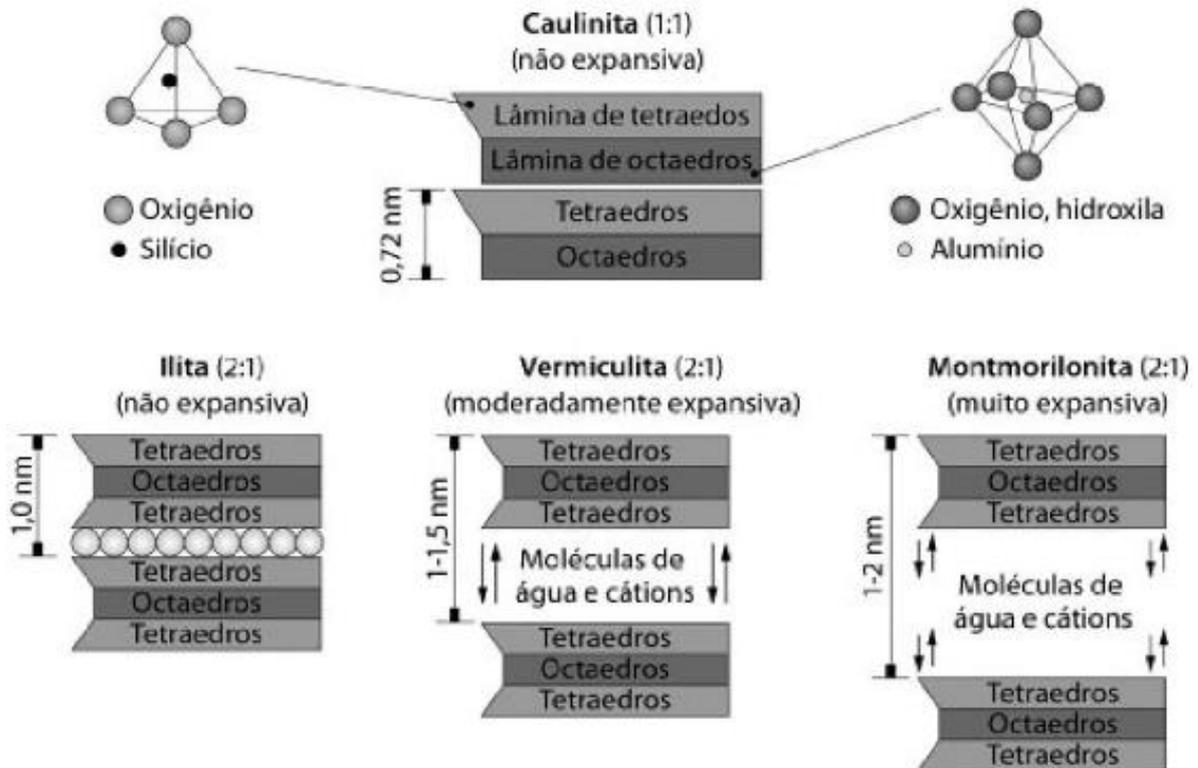
A expansibilidade intracristalina resulta na adsorção de duas ou mais camadas moleculares de água nas superfícies internas dos minerais argilosos. Devido à substituição isomórfica, as partículas de argila consistem em cargas negativas superficiais (SCHMITZ, 2006; SRIDHARAN E CHOUDHURY, 2008). Como resultado, existem forças eletrostáticas entre a superfície negativa da argila e cátions trocáveis no meio fluido dos poros da argila cuja força depende da química do cátion trocável (SCHMITZ, 2006).

A expansibilidade intercristalina acontece quando é possível a absorção de água no interior dos cristais, ou entre as camadas estruturais, ou seja, entre as superfícies externas dos cristais dos minerais argilosos, e dos vazios entre esses cristais (BARBOSA, 2013). Segundo Neves (1993), a absorção da água sobre as partículas dos minerais argilosos, afasta as unidades estruturais e as próprias partículas, podendo chegar ao ponto de separar por completo as lâminas de argilominerais, como no caso das montmorilonitas.

A expansibilidade osmótica é admitida em um processo de duas fases: na primeira, as moléculas de água penetram entre as camadas estruturais das argilas, tendendo a separá-las; na segunda fase, as superfícies das partículas estão já separadas por distâncias superiores a quatro camadas monomoleculares de água, sendo por isso a energia de adsorção relativamente reduzida, o que permite que a energia de repulsão da dupla camada elétrica se sobreponha. Essa expansibilidade é responsável por variações volumétricas observadas que, a depender da condição, pode levar a grande pressão osmótica (OLPHEN, 1963).

A capacidade de troca catiônica é a quantidade de cátions necessária para a manutenção da neutralidade na superfície da argila. Para manter tal neutralidade no meio fluido dos poros da argila, existe uma afinidade natural para que os íons sejam atraídos para a superfície das partículas de argila, diminuindo assim sua concentração com a distância da superfície da argila. Essa mudança na concentração produz a propriedade eletrostática da superfície conhecida como camada dupla difusa (YADAV E TIWARI, 2017), Figura 1.

Figura 1 –Representações Gráficas das estruturas de argilas silicatadas do tipo 1:1 (caulinita) e do tipo 2:1, não expansivas (ilita) e expansivas (vermiculita e montmorilonita)



Fonte: LEPSCH (2011).

A dupla camada, conseqüentemente, causa uma separação entre os minerais e as partículas, acompanhada pelo comportamento de dilatação para minerais argilosos expansíveis, como a montmorilonita (SCHMITZ, 2006) e influencia significativamente todas as propriedades de engenharia do solo argiloso, especialmente a condutividade hidráulica (BESQ *et al.*, 2003; SRIDHARAN E NAGARAJ, 2005; SCHMITZ, 2006) que reduz com o comportamento de dilatação e o inverso quando a dupla camada diminui (SRIDHARAN E CHOUDHURY, 2008). Em essência, isso fornece a estrutura o comportamento de contração ou de inchamento de solos expansivos com redução ou aumento de umidade, respectivamente. Esse comportamento também é influenciado pela morfologia do mineral de argila expansível.

2.1.2.2 Fatores de Ocorrência

A expansão de um solo pode ser provocada por diversos fatores, mas está relacionada principalmente à dois requisitos básicos, **um intrínseco e outro extrínseco**, sendo necessário um ou outro para um solo exibir expansividade, ou ainda ambos.

O requisito intrínseco está relacionado com a composição mineralógica, textura e estrutura. Os argilominerais são componentes chaves para um solo ser expansivo, entretanto nem todos os minerais argilosos experimentam modificações volumétricas, com variação de umidade. Argilas como a vermiculita, e em especial a montmorilonita, apresentam uma importante instabilidade; enquanto argilominerais como a illita e a clorita, podem ou não ser expansivos. Os minerais de argila interestratificados, contendo camadas de minerais expansivos, também podem apresentar expansão em função do mineral e do número de camadas presentes no interestratificado (PAIVA, 2016).

Argilas expansivas podem controlar o comportamento de praticamente qualquer tipo de solo se a porcentagem de argila é mais que 50%, do seu peso (PAIVA, 2016). Contudo há outras características que influenciam na ocorrência da expansão como: A porosidade, perfil estratigráfico, espessura da camada e presença ou não de cimentação (CAMAPUM DE CARVALHO *et al.*, 2015; ESTABRAGH; PARSAEI; JAVADI, 2015).

A expansão de um solo pode acontecer devido a alterações estruturais ou devido à sua composição mineralógica. A primeira podendo ocorrer devido a um alívio de carga (energia externa) ou pelo aumento de umidade (energia interna), tanto em solos ricos em minerais não expansivos como a caulinita ou em solos ricos em minerais expansivos como é o caso da montmorilonita. Já a expansão mineralógica ocorre, principalmente, pela composição química e arranjo estrutural que caracteriza cada mineral de argila expansivo após a entrada de água entre as camadas do mineral, influenciadas tanto pela variação de energia externa como interna (CAMPOS *et al.*, 2008).

O requisito extrínseco está relacionado com a climatologia, a hidrogeologia, a vegetação e a ocupação antrópica que são capazes de transferir a umidade de um ponto a outro do solo. Os requisitos relacionados as mudanças sazonais estão vinculadas as variações climáticas ao longo do ano e sofrem influência antrópica. Além disso, se, em um determinado período, a precipitação exceder a evapotranspiração e a umidade do solo superar a capacidade de campo, produz-se, por drenagem, a eliminação do excesso de água infiltrada. Essa água, dependendo da estratigrafia da área, pode dar lugar a um nível freático ou alimentar um já existente. Nessas condições, as alterações sazonais de umidade são pouco frequentes (FERREIRA *et al.*, 2009).

Outros fatores são responsáveis pelo potencial expansivo do solo, como: distribuição das partículas de argila, porosidade, orientação mineralógica, cimentação, espessura do solo, perfil estratigráfico e outros; o Quadro 2 traz esses fatores segundo Santos (2017).

Quadro 2 – Fatores de ocorrências da variação volumétrica dos solos expansivos

		FATOR	DESCRIÇÃO
Fatores externos	Condições das tensões		Como as mudanças nas tensões normais ou nas de carregamento, podem conduzir a resultados diferentes, em termos de percentagem de expansão.
	Profundidade da zona ativa		Região próxima da superfície do terreno em que os solos experimentam variações no teor de água.
	Gênese e posição topográfica		Determinam a mineralogia secundária do solo e fornece ainda informações sobre os materiais originais de que derivam os solos.
	Teor em água inicial e sua variação		Influencia a densidade, a consistência e a resistência de um solo.
	Variações climáticas		Determinam a presença da água necessária à expansão ou às condições de evaporação que induzem a retração.
Fatores internos	Minerais argilosos expansivos	Expansão inter e intracristalina	Quando a absorção de água é restrita às superfícies cristalinas e às camadas moleculares individuais, respectivamente.
		Área da superfície específica	A superfície específica é inversamente proporcional à dimensão das partículas, mas depende também da respectiva forma.
	Outras variações volumétricas	Arranjo dos grãos	Argilas de estrutura orientada tendem a expandir mais rapidamente e têm menor expansão secundária do que aquelas com uma estrutura aleatória.
		Sulfato de cálcio	A conversão da anidrita em gesso, reação da esmectita e oxidação na pirite e marcassite.
		Mecanismos internos dos minerais argilosos	Absorção capilar, osmótica, hidratação de cátions permutáveis e forças de Van der Waals.

Fonte: Adaptado de Santos (2017).

2.1.2.3 Identificação Geotécnica de Solos Expansivos

A identificação dos solos expansivos é essencial para o sucesso de projetos de engenharia. Segundo Ferreira (1999), a identificação visual desses solos não é fácil e os índices tradicionalmente utilizados para a caracterização dos solos, como SPT, granulometria e índices de consistência, não são suficientes para determinar o grau de expansão de um solo.

Ayala *et al.* (1986) e Carvalho e Andrade (2015) apontam algumas características do solo expansivo que podem auxiliar na identificação no campo:

- Ausência, em geral, de árvores e arbustos caducifólios;

- Cores primárias do solo cinzentas, esverdeadas ou azuladas;
- Fendas poligonais amplas, profundas e, por vezes, de bordas irregulares, na superfície do terreno, durante a estação de seca;
- Fácil moldagem do solo quando umedecido, ou seja, alta plasticidade;
- Alta pegajosidade, ou seja, alta aderência ao maquinário;
- Torrões do solo e deformações produzidas no terreno molhado são muito persistentes quando secos;
- Quando seco, o solo é resistente, sendo difícil rompê-lo com instrumentos manuais;
- Cortes efetuados no solo úmido adquirem, ao secar, o aspecto das superfícies de ruptura produzidas em estado seco;
- Os taludes se apresentam bem desagregados com aberturas de junções prismáticas (estrutura do tipo prismática, com grau de desenvolvimento forte) na superfície, com queda de pequenos blocos e torrões de argila ao pé do talude;
- Os vertissolos podem apresentar um micro relevo característico devido aos movimentos verticais de solo ocasionados pelos ciclos de umedecimento e secagem das argilas.

A identificação da expansividade de um solo pode ser avaliada de duas formas: a primeira é uma pesquisa no local, verificando vários fatores como topografia, vegetação, sinais de contração no solo, fissuras nas edificações e dados com os moradores da região. A segunda forma é interpretar os ensaios de laboratório dos métodos de identificação que pode ser na forma de ensaios mineralógicos, métodos diretos e indiretos.

Métodos diretos são aqueles que disponibilizam resultados quantitativos, fundamentados na expansão induzida do solo ou na tensão necessária para impedir que o solo se expanda. E métodos indiretos são fundamentados na análise mineralógica e propriedades dos materiais argilosos. Como métodos diretos tem-se o ensaio de expansão “livre” e tensão de expansão, ensaios edométricos duplos e simples. E como métodos indiretos tem-se a granulometria, limites de Atterberg, difração de raio X, análise térmica diferencial, análise térmica gravimétrica, capacidade de troca de cátions, entre outros, como pode ser visto no Quadro 3.

Quadro 3 – Métodos de identificação dos solos expansivos

Métodos	Subdivisão	BASE PARA DEFINIÇÃO DO CRITÉRIO	REFERÊNCIAS BIBLIOGRÁFICAS
DIRETOS	Avaliativos	Ensaio de expansão de Lambe Índices Edométricos	Lambe (1960).
	Quantitativos	Ensaio de expansão livre e pressão de expansão. Ensaio edométricos simples e duplos. Ensaio edométricos com sucção controlada	Seed <i>et al.</i> (1960); Chen (1965) Vijayverviya e Ghazzaly, (1973); Rodriguez Ortiz (1975); Cuellar (1978) Jimenez Salas (1980). Escario (1967 e 1969); Aitchison <i>et al.</i> (1974); Johnson (1978); Mc Keen, (1980).
INDIRETOS	Identificativos	Difração de raio-X. Microscopia Eletrônica de Varredura. Análise termodiferencial. Físico-Químicos. Capacidade de troca de cátions (CTC). Análise por sedimentação coloidal.	Carcedo <i>et al.</i> (1986). Fink <i>et al.</i> (1971).
	Qualitativos	Granulometria Geomorfologia Pedologia Identificação visual	Priklonskij (1952); Skempton (1953); Seed <i>et al.</i> (1962); Van Der Merwe (1964); Chen (1965); USBR (1953); Holtz e Gibbs (1956); USBR – HOLTZ (1953); Cuellar (1978); Vijayverviya e Ghazzaly (1973); Rodriguez Ortiz (1975); Daksanamurthy e Raman (1973); Ranganathan e Satyma (1965); Williams & Donaldson (1980).
	Orientativos	Geologia, Pedologia; Geomorfologia, Identificação visual. Baseado no estado do solo; (Umidade natural e Saturação). Baseado nas características organolépticas (Aspectos e Estruturas)	Patrick e Snethen (1976); Carcedo <i>et al.</i> (1986); Ferreira (1990 e 1993).

Fonte: Modificado de Ferreira (1995) e Ferreira *et al.* (2008).

Pesquisas constantes sobre o comportamento expansivo do solo vem sendo desenvolvidas desde a década de 1950. Alguns dos métodos existentes incluem o método do potencial aumento vertical (PVR) (MCDOWELL, 1956), métodos de sucção (MC KEEN, 1992; LYTTON, 1994; PTI, 1996), teste de Clod (MILLER *et al.*, 1995) e método de teste de redução de conteúdo de

água (BRIAUD *et al.*, 2003). Cada um desses métodos tem vantagens e desvantagens, muitos dos quais são discutidos por Briaud *et al.* (2003).

2.1.3 Cartografia Geotécnica de Solos Colapsíveis e Expansivos

Os documentos gráficos, que deram início à cartografia geotécnica no mundo foram apresentados inicialmente por Langen em 1913, a partir do qual as cartas foram surgindo com maiores informações, análises diferenciadas e detalhes em função das escalas, nos quais se classificavam os atributos do meio físico para diferentes usos. No Brasil esses trabalhos tiveram início nos anos de 1965-1966 pelo professor Haberlehner, na Universidade Federal do Rio de Janeiro (UFRJ).

De acordo com Franco *et al.* (2010) a crescente expansão do uso da cartografia geotécnica e do seu reconhecimento como ferramenta de contribuição aos diversos ramos da ciência, a exemplo das áreas de ciências da terra e das engenharias, tem levado ao seu emprego desde o planejamento de obras até o diagnóstico do meio físico.

A cartografia geotécnica, ao disponibilizar informações referentes aos fenômenos associados à superfície terrestre, constitui-se em importante instrumento de planejamento e implantação de obras civis, principalmente na fase de pré-projeto.

Segundo Gomes, Marques e Franco (2017) é a partir do desenvolvimento da cartografia digital que a cartografia geotécnica ganha impulso como um dos instrumentos para planejamento e gestão, constituindo uma representação gráfica das limitações e potencialidades do meio físico.

Os termos cartas e mapas, são usados muitas vezes, como sinônimos, o que deve ser encarado com certo cuidado. Por causa de suas próprias características, a terminologia de mapa ou carta é utilizada diferentemente, de acordo com o país e o idioma correspondente. No caso do Brasil, a Associação Brasileira de Normas Técnicas (ABNT – NBR 13.133/94) confere às seguintes definições:

- Mapa: representação gráfica, em geral numa superfície plana e numa determinada escala, com a representação de acidentes físicos e culturais da superfície da Terra, ou de um planeta ou satélite.
- Carta: representação dos aspectos naturais e artificiais da Terra, destinada a fins práticos da atividade humana, permitindo a avaliação precisa de distâncias, direções e a localização

plana, geralmente em média ou grande escala, de uma superfície da Terra, subdividida em folhas, de forma sistemática, obedecendo a um plano nacional ou internacional.

De acordo com o trabalho de Zuquette e Nakazawa (1998) os produtos geotécnicos (com denominações diversas), podem ser agrupados nas seguintes categorias:

- Carta geotécnica: Propriamente dita, quando expõem limitações e potencialidades dos terrenos e definem diretrizes de ocupação para um ou vários tipos de uso do solo;
- Carta de risco: Quando prepondera a avaliação de dano potencial à ocupação ante uma ou várias características ou fenômenos naturais ou induzidos pela ocupação;
- Carta de suscetibilidade: Quando apresentam gradações de probabilidade de desencadeamento de um ou mais fenômenos naturais ou induzidos pela ocupação;
- Carta de atributo (s) ou de parâmetro (s): Quando se limita a apresentar a distribuição espacial de característica (s) – geotécnicas, geológicas etc. – do terreno.

A investigação direta ou de documentação cartográfica e bibliográfica existente auxilia no reconhecimento da distribuição espacial dos materiais inconsolidados. Pois o mapeamento desses materiais permite, dentre outros, a observação e a realização de correlações com as forças de percolação, com as condições de fluxo superficial e subsuperficial, com a profundidade do nível freático e com a espessura dos horizontes não saturados (zona ativa) (ROMÃO *et al.*, 2015).

Ayala *et al.* (1986) elaboraram um mapa de previsão de risco a expansividade de argilas na Espanha, na escala de 1:1.000.000, classificando a suscetibilidade em quatro categorias: muito baixa, baixa a moderada, moderada a alta e de alta a muito alta suscetibilidade, considerando características litológicas, climatológicas e geotécnicas. Pereira e Pejon (1999) elaboraram uma carta de predisposição à ocorrência de problemas expansivos, para solos localizados a noroeste da Região Metropolitana de Curitiba/PR, utilizando tensões de expansão e expansão livre. Oliveira (2002) elaborou uma carta de zoneamento de risco de solos colapsíveis na área urbana de Ilha Solteira/SP com base em informações de obras de infraestrutura urbana de rede de água e esgoto.

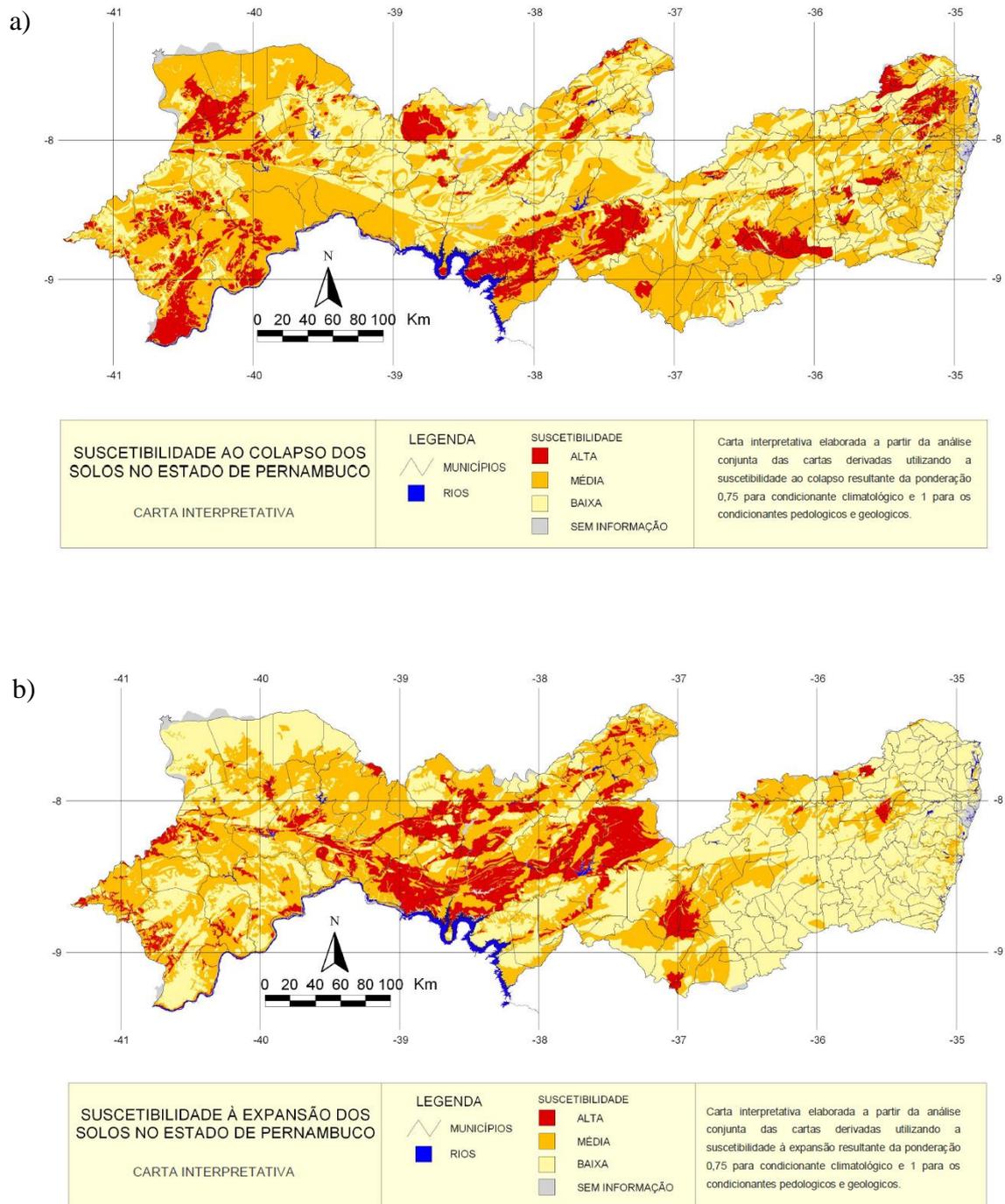
Mais recentemente, Aquino e Ferreira (2022) elaboraram cartas interpretativas de suscetibilidade à ocorrência de solos colapsíveis e expansivos de Teresina/PI, por meio de levantamentos climatológicos, geológicos e pedológicos; Tavares *et al.* (2022) elaboraram cartas de suscetibilidade a colapso e expansão dos solos do estado do Ceará, a partir das características das classes pedológicas, classificando-as em alta, média ou baixa suscetibilidade

e Santos *et al.* (2022) apresentaram mapas derivados da pedologia, a partir do “Mapa Exploratório – Reconhecimento dos Solos do Estado de Sergipe, 1973” para classificar os solos de Sergipe quanto a suscetibilidade (alta, média e baixa) à ocorrência de colapso e expansão.

Para o estado de Pernambuco, em especial, foram criados mapas, através de levantamentos climatológicos, pedológicos e geológicos. Tais mapas são denominados “Mapas Fundamentais”, sendo esses, segundo Zuquette e Nakazawa, (1998), documentos gráficos que registram as informações sobre os diferentes componentes do meio físico.

Silva e Ferreira (2003) elaboram cartas de suscetibilidade ao colapso e à expansão, devido à inundação em solos, do município de Petrolina/PE e Amorim (2004), a partir dos Mapas Fundamentais - climatológicos, pedológicos e geológicos, elaborou as “Cartas Derivadas” de colapso e expansão de Pernambuco. Posteriormente, para a elaboração das “Cartas Interpretativas”, e diante da inviabilidade de realização de baterias de ensaios por todo o território de Pernambuco, optou-se pelo Método Orientativo (baseado em opinião experta) a partir dos Mapas Fundamentais disponíveis de clima, geologia e pedologia e de atributos; resultando nas Cartas Interpretativas de suscetibilidade ao colapso e a expansão, em três diferentes níveis (alto, médio e baixo), observadas na Figura 2.

Figura 2 – Suscetibilidade de ocorrência de solos colapsíveis (a) e expansivos (b) no Estado de Pernambuco (carta interpretativa)



Fonte: Amorim (2004).

As escalas de trabalho que têm sido utilizadas nas diferentes cartas elaboradas (derivadas e interpretativas) respeitaram aquelas disponíveis nos respectivos mapeamentos (fundamentais). O mapa climatológico de VAREJÃO-SILVA (2003), por exemplo, foi disponibilizado na escala 1:600.000; os mapas pedológicos e geológicos, oriundos de trabalhos recentes da

EMBRAPA (2001) e CPRM (2001), estão nas escalas de 1:100.000 e 1:500.000, respectivamente. A escala adotada para a carta interpretativa foi, conseqüentemente, de 1:600.000.

O Sistema de Informações Geográficas dos Solos Expansivos e Colapsíveis do Estado de Pernambuco (SIGSEC-PE) indicou uma suscetibilidade de ocorrência alta em 12,5% da área do Estado, média em 38,7% e baixa em 45,2% (AMORIM, 2004).

A análise do comportamento destes solos, *a priori*, condicionados tanto aos aspectos geológicos, pedológicos e climatológicos, como ao estado tensional em que se encontram e ao que serão submetidos, visa orientar o uso e ocupação do solo e formular diretrizes para planejamento urbano e construção civil, e permitir uma rápida visualização e interpretação dos dados, facilitando a análise (AMORIM, 2004). Entretanto, as Cartas Interpretativas ou quaisquer outros métodos de identificação do comportamento dos solos não substituem as investigações detalhadas dos locais, com ensaios em campo e análise de amostras em laboratório, na fase de exploração de projetos. Seu objetivo é auxiliar no planejamento do uso e ocupação do solo, podendo ser utilizados nas diferentes fases do projeto e implantação de obras de engenharia civil, nos processos de planejamento territorial, urbano e ambiental, assim como no desenvolvimento e na conservação do meio ambiente.

2.1.4 Ocorrência de Solos Colapsíveis e Expansivos no Mundo e no Brasil

Na natureza, solos não saturados são encontrados em diversas condições, assim, considerando a ampla variedade de tais ocorrências no mundo, nas últimas décadas, tem-se constatado grandes esforços no sentido de se compreender o comportamento desses solos com o intuito de melhorar e aperfeiçoar projetos e construções nessas áreas.

Os solos expansivos são encontrados principalmente em regiões áridas e semiáridas de clima tropical e temperado, onde a evapotranspiração excede a precipitação, tais como: os da região do Mediterrâneo (Espanha, Itália, Turquia, Israel, Marrocos, Tunísia etc.), da África do Sul, da região sul da América do Norte, do Norte do México e do Norte da América do Sul (Colômbia, Equador, Peru, Brasil). Estes solos podem também ser predominantes em países onde existiu deposição generalizada de cinzas vulcânicas. Em ambientes vulcânicos tropicais, a cinza vulcânica, rica em alumínio, é depositada geralmente em áreas amplas, em algumas situações encontram-se concentradas em depressões ou áreas baixas, que são geralmente inundadas ou

saturadas com água (MORALES & MORALES, 2004). Entretanto, há registros de ocorrências desses solos em todo o mundo (em mais de 60 países, como pode ser visto na Figura 3), exceto nas regiões árticas (STEINBERG, 2000).

Grande parte da literatura existente, associada a estudos sobre solos expansivos, são de fontes oriundas dos Estados Unidos ou da África do Sul, onde esse tipo de solo é identificado em regiões de clima seco e desértico e a ocorrência de solos expansivos, em regiões de clima tropical úmido, tem sido erroneamente desconsiderada ou ignorada, frequentemente com resultados desastrosos.

Os solos colapsíveis podem ocorrer em várias situações genéticas: depósitos eólicos, aluvionares, coluvionares, solos residuais, materiais de escoamento e até de aterros compactados. A formação destes depósitos depende das variações climáticas e do ambiente geológico.

A presença de solos colapsíveis é mais frequente em regiões de clima quente, todavia, sabe-se da existência destes solos em várias partes do mundo e com climas diversos, Figura 3. Portanto, pode se afirmar que os climas quentes e áridos, embora predominantes, não representam condições essenciais para o desenvolvimento de solos colapsíveis.

Figura 3 - Indicação dos solos colapsíveis e expansivos estudados pelo mundo



Fonte: Adaptado de Paiva *et al.* (2016) e Freitas (2017).

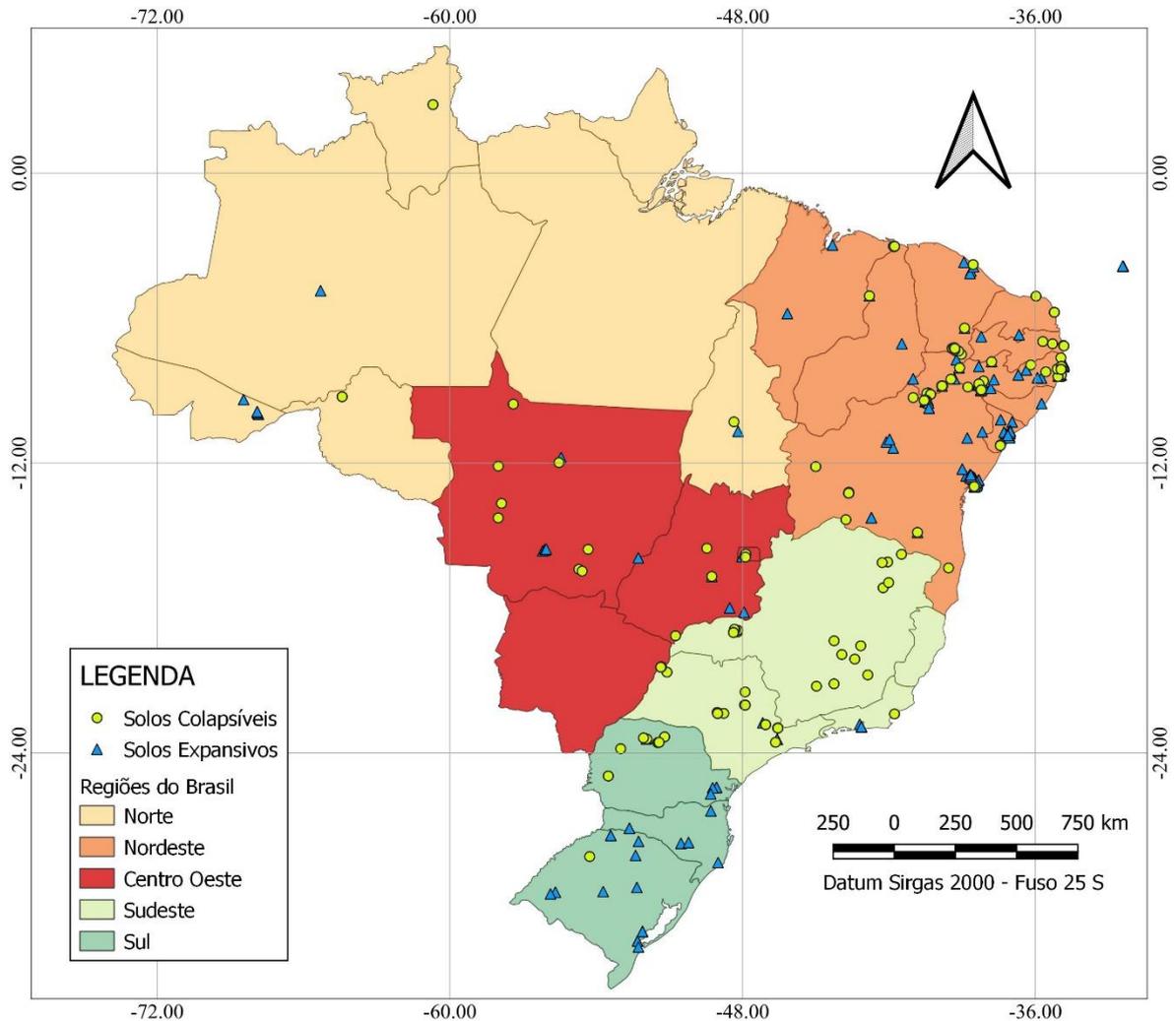
Benatti (2010) esclarece que os solos tropicais apresentam peculiaridades diferentes dos solos de clima temperado. As condições climáticas das regiões tropicais levam à formação de solos lateríticos, dando origem a um horizonte superficial poroso, permanecendo os minerais mais estáveis – quartzo, magnetita, ilmenita e caulinita onde é comum a agregação das partículas finas pela ação dos óxidos e hidróxidos de ferro e alumínio com características e comportamentos mecânicos e hidráulicos não condizentes com sua textura.

No Brasil, são identificados solos expansivos em todos os estados da região Nordeste (Maranhão, Bahia, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco (incluindo o arquipélago de Fernando de Noronha), Alagoas e Sergipe), bem como nos da região Sul (Paraná, Santa Catarina e Rio Grande do Sul), no Centro-Oeste (Mato Grosso e Goiás), no Sudeste (São Paulo e Rio de Janeiro) e no Norte (Amazonas, Tocantins e Acre) do país (BANDASE, 2022), Figura 4.

A maior quantidade de registros de solos colapsíveis no Brasil está localizada, principalmente, nas regiões do Centro-sul (Mato Grosso, Mato Grosso do Sul, Goiás, Santa Catarina e Rio Grande do Sul), Nordeste (Maranhão, Bahia, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas e Sergipe) e grande parte do Sudeste (São Paulo, Rio de Janeiro e Minas Gerais), onde existe o maior número de estudos científicos como o de Mendonça, (1990), Ferreira *et al.* (1991), Andrade (2005) e Futai e Suzuki (2010); proveniente das primeiras investigações geotécnicas relacionadas ao comportamento desses solos, que remontam, principalmente, à década de 60, quando durante a construção das grandes barragens na região Centro-Sul, foram identificados solos porosos e sujeitos a recalques repentinos, durante a fase de enchimento dos reservatórios. Entretanto também são registradas ocorrências no Norte do país (Roraima, Rondônia e Tocantins) (BANDASE, 2022), Figura 4.

Como pode ser observado na Figura 4, as ocorrências de solos colapsíveis e expansivos mais conhecidas são a Formação Aliança e solos aluviais nas bordas do Rio São Francisco no estado de Pernambuco, as integrantes das Bacias Sedimentares do Recôncavo Baiano, do Paraná e do Rio Grande do Sul. Assim, identificar previamente esses solos é de grande importância, principalmente porque atinge regiões economicamente ativas.

Figura 4 - Registros de Solos Colapsíveis e Expansivos no Brasil



Fonte: A Autora (2022).

A Amazônia apresenta unidades geotécnicas de solo, cujo comportamento, por vezes, difere dos solos descritos nas bibliografias usuais de geotecnia clássica. A maioria dessas unidades geotécnicas, especialmente no estado do Acre, é formada por solos oriundos de rochas sedimentares, parcialmente saturados, com presença de argilominerais expansivos, um comportamento ainda pouco estudado, na região Norte, devido a sua posição e dimensões geográficas (BORGES, 2019).

Mediante tais características de fatores condicionantes e variabilidade dos aspectos geológico, pedológico e climatológico dos solos brasileiros, ratifica-se a necessidade de empenho em pesquisas que facilitem a identificação de solos expansivos e colapsíveis.

2.2 Inteligência Artificial

O conceito de Inteligência Artificial remonta à década de 1950 e foi nomeado na conferência DartMouth, o termo se refere à capacidade de "raciocínio" dos dispositivos que reflete a função do cérebro humano, a capacidade de aprender e de se adaptar a um ambiente desconhecido ou a uma situação nova (BAUCHSPIESS, 2002). Assim, duas abordagens teóricas regem a IA: a cognitiva/simbolista/simbólica e a conexionista.

A IA simbólica é referente ao modo do ser humano racionar. Nela, é preciso inserir no sistema, de forma manual, os dados específicos do problema para que ele possa raciocinar e tomar as decisões, como, por exemplo, os sistemas especialistas. Enquanto a IA conexionista, tem o objetivo tentar imitar a forma que o sistema nervoso processa informações. Nesse tipo de programa é possível deduzir os padrões de maneira automática a partir de dados existentes, o que deixa a IA simbólica um pouco obsoleta (GRANATYR, 2017).

A rede neural faz parte da IA conexionista e nela existem as Redes Neurais Artificiais, que permitem realizar tarefas relativas à “percepção”. Junto com plataformas de *hardware* de computação com *multicore* atuais, é possível interconectar muitas camadas neurais para fornecer uma habilidade maior em se concentrar nos aspectos essenciais, ignorando características menos importantes ou acidentais (PEREZ, 2017). As RNA são, portanto, adaptativas, generalizando a experiência e resolvendo situações que possuem variação.

Para construir programas utilizando a IA simbólica é necessário alimentar o sistema com os dados específicos do problema, já para a IA conexionista o programa consegue inferir os padrões automaticamente por meio dos dados existentes. Estes são os mais utilizados em resolução de problemas de engenharia, por sua capacidade de estabelecer padrões que correlacionem os dados de saída e entrada, através de exemplos, mesmo sem um conhecimento prévio das relações entre os dados, em contraste com os modelos tradicionais desenvolvidos a partir da natureza das relações físicas entre os dados de entrada e saída. Dentre as diversas técnicas de aprendizado de máquina existentes, duas das mais aplicadas na resolução de problemas não-lineares em engenharia são RNA e Algoritmos Genéticos (FATEHNIA; AMIRINIA, 2018).

2.3 Redes Neurais Artificiais

A homogeneidade e isotropia são duas propriedades levadas em consideração para a maioria dos materiais, como aço e concreto num projeto de Engenharia Civil. Por outro lado, para solos tem sido provado que, devido à complexidade da formação geológica, que causa processos físicos imprecisos, as propriedades de engenharia geotécnica mostram comportamento variado e incerto (JAKSA, 1995). Portanto, o desenvolvimento de modelos analíticos ou empíricos em algumas situações simplificadas é viável; no entanto, modelos mais práticos que os analíticos são de interesse comum (SHAHIN *et al.*, 2001; PARK 2011; SHAHRI, 2016).

Em situação de complexidade de fenômenos a serem analisados por modelos lineares, os quais não apresentem desempenho satisfatório, deve-se considerar o desenvolvimento de um modelo não linear. Nelles (2001) atenta para o fato de que ao mudar de um modelo linear para outro não linear pode ocorrer que este não seja suficientemente flexível, apresentando um rendimento pior, frente ao linear. Para evitar esse efeito indesejável pode-se utilizar um modelo não linear que contenha um modelo linear como caso especial. Como exemplos desses modelos têm os do tipo polinômios (que simplificam um modelo linear pelo grau) ou modelos do tipo neural ou *fuzzy* (que simplificam para o modelo linear quando o número de camadas).

A teoria das redes neurais apresenta-se como método de aproximação de sistemas complexos, em que uma grande quantidade de dados deve ser modelada e analisada em um contexto multivariado, envolvendo simultaneamente, aspectos estatísticos e computacionais, dinâmicos e de otimização, particularmente úteis quando esses sistemas são difíceis de modelar usando métodos estatísticos convencionais (CABRAL, 2017).

Assim as RNA são modelos computacionais de inteligência artificial, aprendizado de máquina, baseados no cérebro humano e na sua capacidade de previsão de resultados, a partir do reconhecimento de padrões proveniente da experiência por exemplos previamente estudados (RIPLEY, 1996). Associado ao poder computacional de grande quantidade de dados, o seu desenvolvimento matemático é baseado na idealização da forma, como os neurônios humanos funcionam e na capacidade de percepção e aprendizagem de fenômenos complexos, não lineares e multivariados (BISHOP, 1995; HAYKIN, 2008; SCHMIDHUBER, 2015; DANTAS NETO *et al.*, 2017).

Caso o leitor não esteja habituado a temática de Redes Neurais Artificiais, poderá encontrar no Apêndice A uma breve revisão sobre o tema. Caso contrário, a seguir são apresentadas aplicações na área da engenharia civil, geotecnia e colapsividade e expansividade dos solos.

2.3.1 Áreas de Aplicação da Redes Neurais Artificiais em Engenharia Civil

As RNA podem ser empregadas em diversos problemas relacionados às engenharias. Atualmente, o campo de aplicação é muito amplo, abrangendo tópicos como otimização de processos, simulação de processos construtivos, previsão de ameaça sísmica, classificação de sinais provenientes de testes não destrutivos, estimativa de custos de construção, engenharia de transporte, análise estrutural, tratamento de águas, recuperação de aquíferos, previsão de vazões, avaliação do potencial de liquefação de solos e muitas outras.

Os trabalhos sobre redes neurais em Engenharia Civil tiveram sua aplicação intensificada a partir do final da década de 80, até os anos 2000; a maior parte desses trabalhos esteve compreendida na área da engenharia estrutural em pesquisas de: Previsão de cargas e definição de seções de vigas de concreto, seleção dos elementos estruturais de edificações, problemas de mecânica de fraturas, avaliação de danos em elementos estruturais, previsão do comportamento de conexões estruturais, previsão da resistência de elementos estruturais, determinação da relação carga-deslocamento em elementos estruturais, otimização dos processos de execução de projetos estruturais, identificação de sistemas dinâmicos estruturais, criação eficiente e automática de malhas de elementos finitos para análise estrutural, previsão das propriedades mecânicas e térmicas de materiais compostos, modelagem constitutiva de concreto e materiais compostos em geral; avançando posteriormente com a estimativa de resistência a compressão do concreto (CAVALCANTI & SILVA, 2012) e o dimensionamento de pilares de concreto (NOGUEIRA SILVA *et al.*, 2016).

Estudos de estimativas de custos usando redes neurais também obtiveram resultados satisfatórios, como por exemplo: análises de custo de construção de rodovias por Sodikov (2005), em que o percentual de erro encontrado foi de 25%, já no estudo de Gardner, Gransberg e Rueda (2016), o erro médio percentual absoluto foi igual a 23%; e na análise de Barros, Marcy e Carvalho (2018), o percentual do erro em rodovias brasileiras foi de 1%. Outro exemplo é o estudo de Juszczak, Leśniak e Zima (2018), no qual foi avaliado os custos estimados de construção de quadras de esporte na Polônia, em que os autores obtiveram resultado satisfatório, com erro médio percentual absoluto igual a 9,97%. Barros (2019) discutiu a estimativa dos custos de projetos de construção de rodovias com maior precisão na fase do desenvolvimento do projeto, apresentando o menor erro dos conjuntos de testes igual a 9%. Tais estudos contribuem e apoiam as decisões de viabilidade econômica de construção de rodovias, uma vez que os parâmetros analisados são facilmente quantificáveis na etapa de anteprojeto.

2.3.2 Redes Neurais Artificiais Aplicadas a Geotecnia

Em Engenharia Geotécnica as aplicações mais comuns estão relacionadas com a caracterização e classificação de solos, fundações, recalques, escavações, estabilidade de taludes, reforço de solos, água subterrânea, dinâmica de solos e mineração.

Uma perspectiva particular a quase todos os estudos da Engenharia Geotécnica é a atribuição para resolver problemas complexos, quando o solo está sujeito às mais variadas condições de carregamento, as quais geram tensões, deformações (elásticas e plásticas), amolecimento, endurecimento e podem inclusive levar a estrutura ao colapso (por exemplo MAYORAZ *et al.*, 1966; ZHOU E WU, 1994; BASHEER *et al.*, 1996; FERNANDEZ-STEEGER *et al.*, 2002; GOH, 2002; BAZIAR E GHORBANI, 2005; KIM E KIM, 2006; HANNA *et al.*, 2007; SHAHIN *et al.*, 2008; SAYADI *et al.*, 2013); bem como estimar as propriedades geotécnicas do solo (por exemplo, GRIBB E GRIBB, 1994; CAL, 1995; YANG E ROSENBAUM, 2002; LEE *et al.*, 2003; CELIK E TAN, 2005; ERZIN, 2007; SINHA E WANG, 2008); além da estimativa da sensibilidade das argilas (por exemplo, ROSENQUIST, 1953; TORRANCE, 1983; RANKKA *et al.*, 2004; SOLHEIM *et al.*, 2005; NADIM *et al.*, 2008; SHAHRI, 2016; LUNDSTRÖM *et al.*, 2009).

Diversos trabalhos na área de Mecânica das Rochas foram desenvolvidos, utilizando as redes neurais como ferramenta importante e valiosa para a elaboração de modelos de previsão. Na maioria, contemplam a previsão de propriedades da rocha intacta, como: resistência à compressão uniaxial, que podem ser encontrados nos trabalhos de Meulenkamp e Grima (1999), Moshrefii *et al.* (2018); módulo de elasticidade, podendo ser encontrados nas pesquisas de Sonmez *et al.* (2016), Yilmaz e Yusek (2008) e Dehghan *et al.* (2010); e por fim, resistência à tração segundo Singh *et al.* (2001). No que se refere ao desenvolvimento de modelos de previsão das descontinuidades rochosas utilizando *perceptron* multicamadas (MLP), vale citar os trabalhos de Dantas Neto *et al.* (2016, 2017) e Leite (2019).

Na área de pavimentação, Zeghal e Khogali (2005) previram com sucesso o Módulo de Resiliência para materiais granulares, utilizando RNA, tendo como dados de entrada: A densidade, o estado de tensões e a umidade do material. E para Ribeiro *et al.* (2018) a existência de mapeamentos preditivos ou de reconhecimento com propriedades de interesse pôde maximizar a qualidade dos anteprojetos rodoviários, pois a carência desses mapeamentos dificulta o conhecimento prévio dos materiais, resultando em maior consumo de recursos

naturais, maior tempo de execução e conseqüentemente elevação dos custos financeiro e ambiental dos projetos.

Estudos dos Movimentos de Massa Gravitacionais com RNA buscaram avaliar a suscetibilidade de áreas sujeitas a tais movimentos, análise de eventos perigosos, reconhecimento dos movimentos a partir de imagens, para previsão e monitoramento, dentre outros objetivos, como no trabalho de Melo & Pedrollo (2015), que utilizaram as RNA para estimar os parâmetros de ajuste da curva característica solo-água. Neste trabalho os autores mostram que é possível treinar uma RNA, desde que se tenha uma base de dados abrangente. Outros trabalhos utilizam as RNA a partir parâmetros geológico e geotécnicos e fatores ambientais para determinar áreas instáveis em taludes (LIAN *et al.*, 2015; GORDAN *et al.*, 2016; CHAKRABORTY & GOSWAMI, 2017 e MOAYEDI *et al.*, 2019). Pereira *et al.* (2022) aplicaram RNA e geoprocessamento, para avaliação e zoneamento da suscetibilidade de deslizamentos superficiais em Recife/PE.

Tanto no cenário brasileiro quanto no internacional, a análise de alguns exemplos em que as RNA apresentam resultados mais satisfatórios, quando comparados com os métodos clássicos, corroboram a premissa que é possível a utilização das RNA para resolução, de forma eficiente, dos problemas que envolvem a engenharia geotécnica, a exemplo da determinação da capacidade de carga de estacas de concreto armado e do tipo hélice contínua. Avaliados os valores dos coeficientes de determinação, histogramas de erros e valores extremos dos erros (máximos positivos e negativos), sob todos os aspectos observados, é possível afirmar que os modelos obtidos se apresentam como potenciais ferramentas alternativas para a determinação da capacidade de carga de ponta, por atrito lateral e total (DELAZZERI *et al.*, 2017, DELAZZERI, 2018).

Para avaliar a resistência dos solos, Ribeiro *et al.* (2018) apresentou um modelo de RNA, baseado em características biofísicas presentes em bases cartográficas para previsão e mapeamento de valores de CBR-N e CBR-I para a Região Metropolitana de Fortaleza (RMF), alcançando uma precisão dos modelos em torno de 0,98. No trabalho de Alan *et al.* (2020), foi feita uma tentativa de desenvolver relações entre o CBR e várias propriedades do índice do solo, para solos aluviais em Bengala Ocidental, Índia, de modo a prever valores de CBR em solos saturados e não saturados por RNA e Método de Krigagem. Os resultados revelam que os métodos podem ser usados com sucesso, possibilitando uma ideia preliminar sobre as propriedades de resistência de sub-base do solo. O trabalho de Al-Busultan *et al.* (2020),

utilizando as RNA, traz uma correlação entre o CBR da sub-base e os índices do solo de Karbala, Iraque. Os resultados indicaram que o índice de sais solúveis foi o fator mais influente entre os parâmetros do solo, enquanto o Índice de Plasticidade foi o fator com menor influência.

Na pesquisa de Totola (2020) a curva de retenção de água no solo (CRAS), que apresenta papel crucial no desenvolvimento da teoria e na aplicação da Mecânica dos Solos não saturados, foi estimada por RNA e avaliada para solos brasileiros, no intervalo de sucção de 0 e 1500 kPa. O desempenho das Redes se mostrou promissor podendo ser uma ferramenta eficiente na estimativa indireta da curva de retenção de secagem para solos tropicais e subtropicais do Brasil. Dentre os diversos cenários analisados, o valor mínimo do RMSE foi de 0,035 para a abordagem pontual e de 0,029 para a abordagem pseudocontínua. Os resultados são similares aos apresentados na literatura, para a utilização de RNA para solos temperados. O trabalho de Tamiru *et al.* (2022) validou a precisão de zonas de potencial de água subterrânea, usando o modelo RNA, considerado uma ferramenta eficaz, com uma porcentagem de precisão medida a partir da curva ROC e AUC, que concordam em 96% das estimativas.

Na costa brasileira, foi avaliado por Oliveira Filho *et al.* (2020), o potencial das RNA para prever o Índice de Compressão (C_c) de solos moles, a partir de testes de consolidação padrão (edômetro) e dos índices de propriedades dos solos correspondentes. O melhor desempenho das correlações propostas sobre aquelas da literatura evidencia a influência da previsão como alternativa durante a investigação preliminar da adequação de um local de fundação durante as etapas de planejamento. Os valores da Taxa de Compressão (C_R) mostraram-se menos favoráveis que as estimativas de C_c . para o conjunto de todas as amostras, embora exista forte correlação entre C_c e C_R (OLIVEIRA FILHO, 2019).

A aplicação de RNA a análise do desempenho de fundações surge da necessidade de olhar além das equações empíricas, para capturar as complexas iterações solo-fundação e o comportamento não linear do subsolo, Mola-Abasi *et al.* (2020) previram o assentamento de fundações rasas em solos sem coesão, usando modelagem polinomial de leituras contínuas, medidas a partir do teste de penetração de cone (CPT). A análise de sensibilidade identificou que a resistência da ponta do cone tem a maior influência na previsão de assentamento. Wu *et al.* (2021) usaram RNA para desenvolver correlações entre o ensaio pressiométrico (PMT) e o teste do Piezocone (CPTU), para aplicação em fundações profundas, em mais de 150 conjuntos de dados, contendo informações de solos arenosos e argilosos. Os valores de recalques preditivos, foram validados usando recalques reais, medidos a partir dos testes das provas de carga.

2.3.3 Áreas de Aplicação na Colapsividade e Expansividade dos Solos

A abordagem de RNA na área de colapsividade e expansividade dos solos, tem apresentado vantagens, como a redução do número de ensaios diretos, durante a fase de reconhecimento do terreno, economizando tempo e custos financeiros; se comparado com métodos estatísticos.

O trabalho de Basma *et al.* (2003) implementou RNA sequenciais para modelar o comportamento de solos expansivos, de acordo com as variações climáticas, demonstrando que existe a possibilidade de desenvolver um modelo geral, que pode prever tal expansão, dependente do clima, com base nas propriedades básicas do solo, com precisão relativamente alta. Moosavi *et al.* (2006) também desenvolveram estudos de RNA eficientes, para modelar pressões, dependentes do clima, devido a expansão cíclica em rochas fracas, a fim de determinar o limite máximo de tais expansões, sob condições de secagem e umidificação cíclicas. E Doris *et al.* (2008) apresentou testes de RNA de retro propagação, investigados para mapear dados climáticos (temperatura e precipitação) para o movimento vertical da superfície do solo, em locais de campo no Texas e na Austrália, demonstrando um mapeamento bem-sucedido, dos dados de temperatura e precipitação, e a capacidade da RNA de capturar tendências no movimento vertical da superfície do solo.

Ashayeri & Yasrebi (2009) elaboraram modelos de RNA capazes de encontrar as funções de mapeamento, que introduzem os efeitos do estado de compactação e potencial composição no comportamento de expansão, de modo que, as forças de absorção de água e o potencial de substituição de ar e água, durante a saturação, podem explicar o comportamento de expansão livre das argilas compactadas.

A fim de lidar com as complicações na descrição do comportamento de expansão do solo, Ikizler *et al.* (2010) desenvolveram alternativas usando a abordagem de rede neural artificial. O modelo de previsão de pressão de expansão lateral transmitida e pressões de expansão vertical em uma estrutura de retenção foi desenvolvido para simular as condições de campo o mais próximo possível, reduzindo assim o número de testes de laboratório, para que o custo e o tempo fossem economizados.

Merouane & Mamoune (2018) trataram em sua pesquisa sobre a identificação e classificação do potencial de expansão de argilas na Argélia, segundo o desenvolvimento de dois modelos de redes neurais: uma rede para prever a pressão de expansão; e a outra para prever a amplitude de tal expansão, tendo como variáveis de entrada, os parâmetros de identificação geotécnica: profundidade, porcentagem da fração de argila, porcentagem de água, limite de liquidez, limite

de plasticidade, atividade e densidade seca. A eficácia de previsão dos modelos desenvolvidos pelas RNA, em comparação com um modelo de Regressão Linear Múltipla, destacou um aumento na qualidade da previsão da pressão em 51,10% e da amplitude em 33,47%.

O estudo de Salahudeen *et al.* (2020) aplicou as RNA para prever os valores de resistência a compressão não confinada (UCS) da argila expansiva nigeriana; utilizando oito dados de entrada. O desempenho da rede simulada foi satisfatório tendo valores que atenderam aos critérios mínimos recomendados para condição de forte correlação observada entre os valores experimentais de UCS, obtidos por procedimentos de testes laboratoriais e os valores previstos usando RNA.

Como exemplo de aplicação de modelos de RNA, para prever o potencial de colapso dos solos, cita-se o trabalho de Juang *et al.* (1999), que utilizaram sete parâmetros do solo como variáveis de entrada. A variável de saída é a colapsibilidade ou o potencial de colapso. A RNA com as técnicas de pré-processamento e pós-processamento propostas, se mostraram superior ao modelo de RNA convencional.

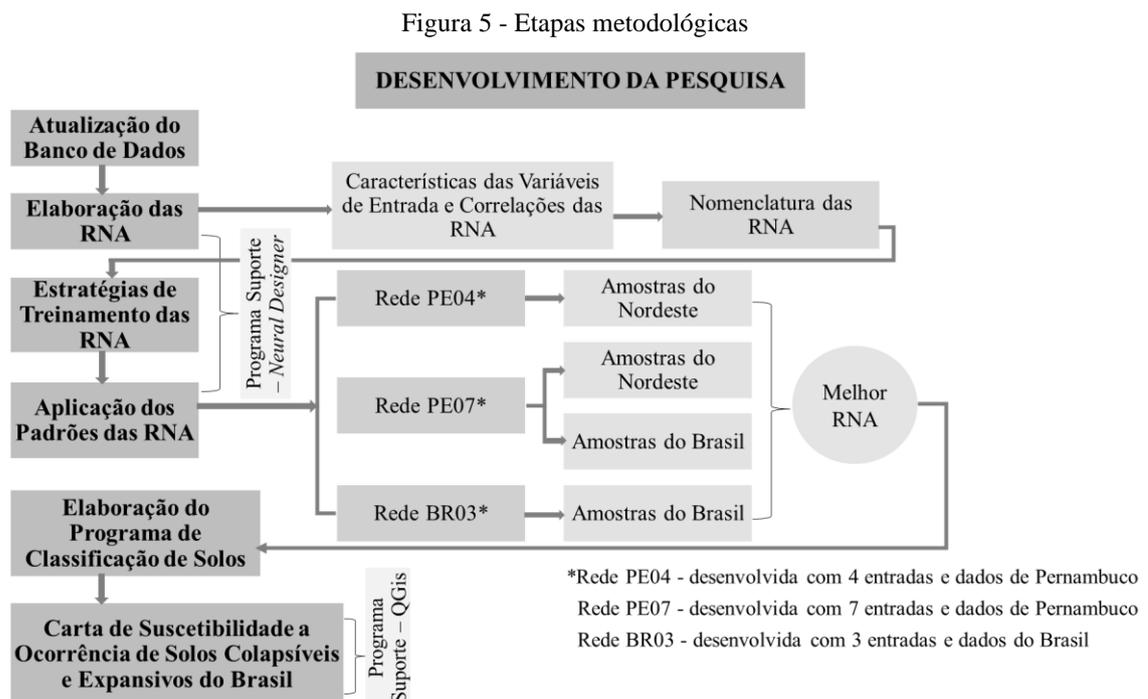
A pesquisa de Basma & Kallas (2004) envolveu a avaliação do potencial de colapso usando o aparelho oedômetro unidimensional. Para cobrir o escopo mais amplo possível de dados, um total de oito tipos de solos foram selecionados, cobrindo uma ampla faixa de gradação. O treinamento foi realizado até que o erro quadrático médio da soma (MSSE) fosse minimizado. Os resultados provaram que as redes neurais são muito eficientes na avaliação do comportamento complexo de solos colapsáveis, usando o mínimo de processamento de dados.

Salehi *et al.* (2015) propuseram três redes neurais para estimar o potencial de colapso dos solos do norte do Irã, com base nas propriedades básicas do índice. Várias arquiteturas de redes neurais e algoritmos de treinamento foram examinados e um estudo de comparação foi realizado para provar a eficiência de três tipos de neurais. Os resultados numéricos mostram que a melhor estimativa é obtida via rede MLP, com algoritmo de treinamento de retro propagação Levenberg Marquardt.

3 MATERIAIS E MÉTODOS

Os capítulos anteriores e o Apêndice A estabeleceram os termos e conceitos relacionados com a identificação e análise de solos colapsíveis e expansivos, que foram discutidos no âmbito deste estudo, cujos objetivos contemplam a identificação de solo colapsível e expansivo por meio de Redes Neurais Artificiais, com aplicação aos solos do Brasil. Este terceiro capítulo é metodológico e descreve, com maior ênfase, os materiais e métodos utilizados.

As etapas de desenvolvimento da pesquisa e as sub etapas que as compõem são apresentadas esquematicamente na Figura 5, e serão detalhadas nos itens seguintes. Inicialmente, tem-se a atualização e ampliação do Banco de Dados de Solos Especiais - BANDASE, (FERREIRA, 1990), seguida da etapa de elaboração das RNA, que ocorre com a definição das variáveis de entrada e saída e de suas correlações, determinando, posteriormente, as nomenclaturas das RNA. Na etapa seguinte, tem-se o desenvolvimento das RNA, a partir das estratégias de treinamento e seleção, tendo os desempenhos analisados, segundo testes de avaliação binária, e seguido da aplicação dos padrões de cada rede, às amostras de domínios diversos (amostras dos demais estados do Nordeste e/ou a amostras de todo o Brasil), de modo a definir a melhor rede, que é usada para o desenvolvimento do programa de identificação de solos colapsíveis e expansivos e para a elaboração de cartas de suscetibilidade de ocorrência de solos colapsíveis e expansivos.



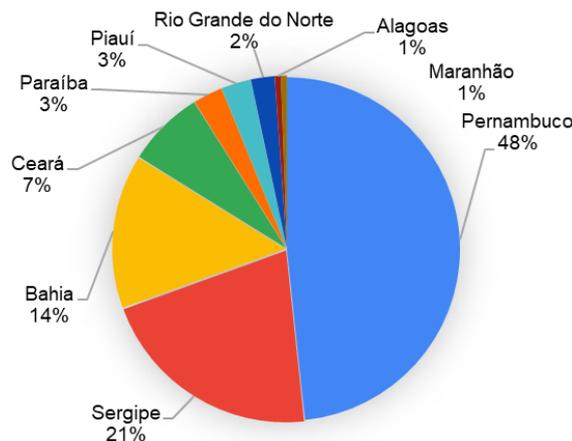
Fonte: A Autora (2022).

3.1 Apresentação do Banco de Dados

A partir do BANDASE (FERREIRA, 1990) do Grupo de Pesquisa de Solos Não-Saturados (GNSat) da Universidade Federal de Pernambuco (UFPE), e de pesquisas nas bases de dados, bancos de teses, anais de congressos e periódicos nacionais e internacionais, foi possível ampliar o banco de dados, de modo a reunir os atributos exigidos pela pesquisa. O BANDASE vem sendo construído desde 1990, sendo atualizado periodicamente, e conta com amostras georreferenciadas com dados da publicação, no qual se destacam a caracterização do solo, plasticidade e potenciais de colapso e expansão por métodos indiretos e diretos. Vale salientar que tais dados são determinados pelas normas brasileiras vigentes a época em que foram desenvolvidos os trabalhos, desse modo, os parâmetros de classificação das características desses solos atendem à época.

Em um primeiro momento (março de 2020), o banco de dados era formado apenas por amostras do Nordeste do Brasil (180 amostras), sendo 79 amostras de solos colapsíveis e 101 de solos expansivos. Desse total de amostras, 87 pertenciam a Pernambuco; 38 amostras de Sergipe, 26 da Bahia, 13 do Ceará, 5 da Paraíba, 5 do Piauí, 4 do Rio Grande do Norte, 1 de Alagoas e 1 do Maranhão, como poder ser observado na Figura 6.

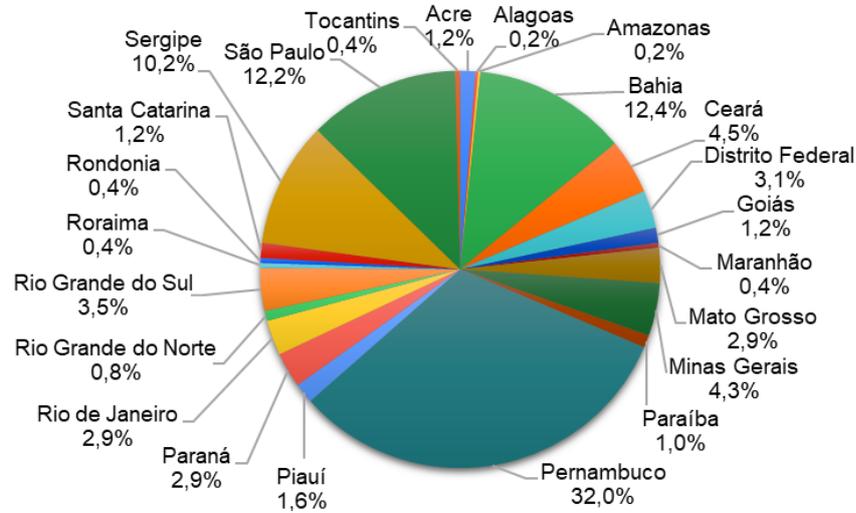
Figura 6 - Proporção de ocorrências no BANDASE para 180 amostras do Nordeste



Fonte: A Autora (2022).

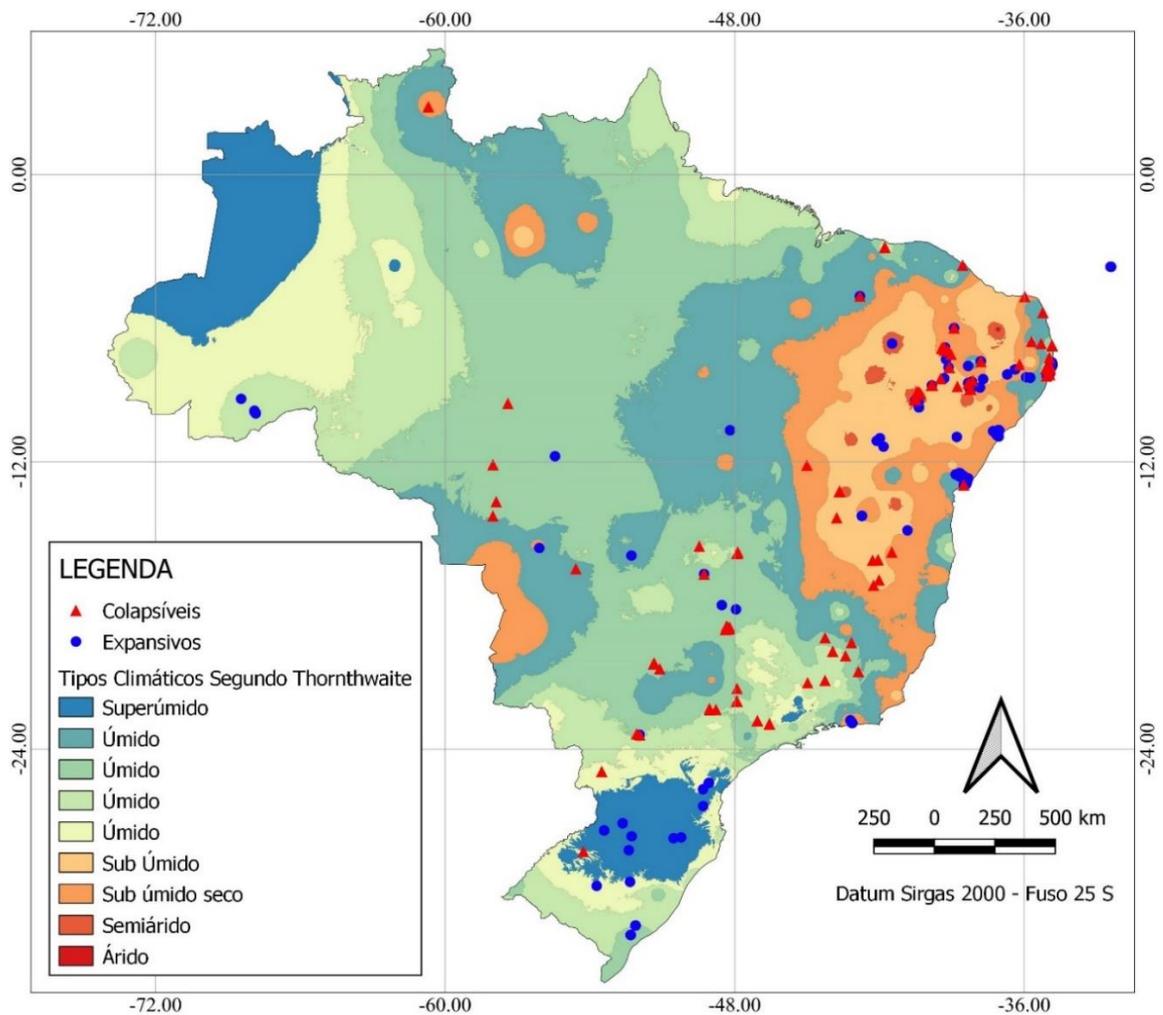
Como o avanço das pesquisas do GNSat, em março de 2022, esse banco de dados passa a contar com 526 amostras de solos colapsíveis e expansivos, das 5 regiões do Brasil (Figura 7), sendo 267 de solos colapsíveis e 259 amostras de solos expansivos. Com uma maior concentração de amostras nas regiões do Nordeste e Sudeste, onde ocorre uma maior quantidade de pesquisas em universidades, sobre esses comportamentos dos solos (Figura 8).

Figura 7 - Proporção de ocorrências no BANDASE para 526 amostras do Brasil



Fonte: A Autora (2022).

Figura 8 - Localização das amostras usadas para o desenvolvimento das RNA



Fonte: A Autora (2022).

3.2 Desenvolvimento das Redes Neurais

3.2.1 Programa suporte – *Neural Designer*

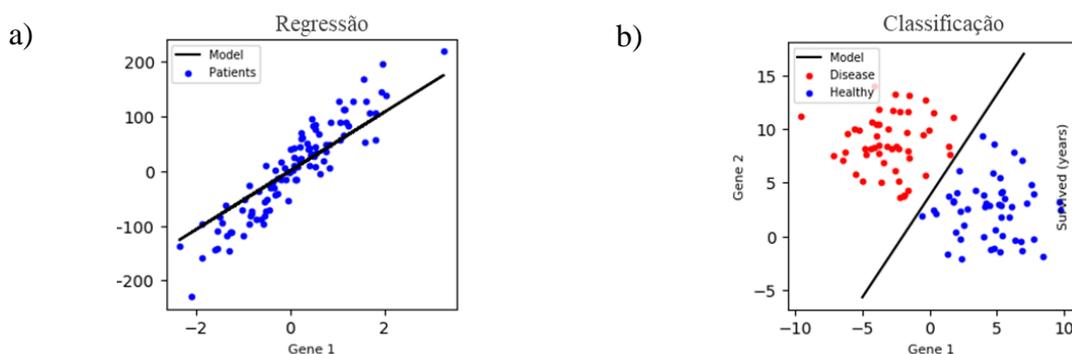
O *Neural Designer* é uma plataforma de prospecção de dados e aprendizado de máquina que ajuda a criar, treinar e implantar modelos de redes neurais, com o objetivo de estabelecer padrões consistentes, como regras de associação ou sequências temporais, para detectar relacionamentos sistemáticos entre variáveis, identificando assim, novos subconjuntos de dados.

A plataforma de ciência de dados e aprendizado de máquina foi criada pela *startup* espanhola *Artelnics* em 2014, e conta com técnicas avançadas para preparação de dados, identificação de padrões, tomada de decisões e implantação de modelos, aplicada com sucesso em vários projetos de diferentes setores.

Na versão gratuita 4.2.0 utilizada neste trabalho são oferecidos 2 tipos de análise de teste (que estão sujeitos ao tipo do projeto), sendo elas (Figura 9):

1. Aproximação (ou função de regressão): Onde o objetivo básico é ajustar uma função a partir de um conjunto de dados de entrada e de saída (Figura 9-a).
2. Classificação (ou reconhecimento de padrões): O objetivo básico é modelar as probabilidades de um indivíduo pertencer a um determinado grupo, condicionadas às variáveis de entrada (Figura 9-b). Pode-se distinguir entre dois tipos de modelos de classificação: classificação binária e classificação múltipla. Na classificação binária, a variável de destino geralmente é verdadeira ou falsa, enquanto na classificação múltipla, a variável de destino geralmente é nominal (classe_1, classe_2 ou classe_3).

Figura 9 – Análise de teste de RNA por Regressão (a) e Classificação (b).



Fonte: MC.AI (2020).

3.2.2 Escolha de Variáveis e Arquiteturas das Redes Neurais Artificiais

Definir as variáveis de entrada mais relevantes é um pré-requisito importante para aplicar com sucesso a RNA ou outro método de classificação ou reconhecimento de padrão, de modo que se possa realizar previsão sem perdas ou excesso de informações. A importância dessa definição fica evidente quando se tem em mente que uma RNA é treinada para identificar relações não lineares entre pares de dados de entrada e saída, e se as variáveis de entrada não forem suficientes para representar o problema, a capacidade de generalização da rede neural modelada estará comprometida. Além disso, se forem consideradas variáveis que não representam o problema, ou o excesso de variáveis, ainda que estejam correlacionadas ao problema, o tempo de treinamento pode ser estendido e a qualidade de previsão da rede reduzida.

A etapa inicial de elaboração da RNA consiste na preparação das variáveis de entrada e saída, contando previamente com uma análise criteriosa das amostras disponíveis para uso pelo BANDASE. Pela disponibilidade de amostras no banco de dados, as redes foram desenvolvidas com amostras pertencentes ao Estado de Pernambuco; tendo em vista que a etapa de desenvolvimento da rede não deve apresentar dados de variável de entrada ausentes em maior quantidade, comparados aos dados presentes, essa constatação foi possível a partir da observação do desempenho das redes geradas inicialmente, nas quais classificou-se amostras de solos colapsíveis e expansivos, mesmo com a ausência de informações de entrada, prejudicando o desempenho da classificação.

A primeira seleção de variáveis de entrada, que compõem as RNA, foi feita a partir da identificação em revisões bibliográficas, de quais variáveis geotécnicas contribuíam com melhores desempenhos em redes preditivas de características de colapsividade e expansividade, de modo que tais variáveis sintetizassem os custos e tempo de análise desses solos. As variáveis geotécnicas são de fácil obtenção, na maioria dos ensaios, e são utilizadas na identificação e classificação de solos colapsíveis e expansivos, como podem ser observados nos critérios de identificação de Skempton (1953). Para outros critérios de identificação de solos colapsíveis e expansivos, não foi possível aplicar as contribuições de forma efetiva na determinação das variáveis de entrada pela indisponibilidade de dados nas fontes pesquisadas, como é o caso dos gráficos de Reginatto e Ferrero (1973) e Jimenez Salas (1980) (Figura 10-c e Figura 10-f), os quais continham dados de menos de 15% e 11% das amostras de colapso e expansão, respectivamente.

Os índices físicos (variáveis geotécnicas) são provenientes da classificação convencional dos solos que, de modo geral, definem as camadas com base em mudanças de textura, cor e/ou resistência, obtidas facilmente em ensaios laboratoriais de baixo custo, em amostras deformadas e destorroadas, e através de correlações entre eles.

Inicialmente destacou-se 6 variáveis de entrada, que foram avaliadas estatisticamente e segundo a Análise de Componentes Principais (PCA), e puderam ser reduzidas a 4 variáveis (ver Apêndice B – Progressão das Redes Desenvolvidas), de acordo com a dispersão dos parâmetros que apresentaram maior coeficiente de variação, dentre as variáveis de entrada, e de acordo com os graus de correlação das variáveis de entrada entre si. Durante o andamento dessa pesquisa, diversas redes neurais foram desenvolvidas (Apêndice B), entretanto o desempenho de 3 delas é detalhado para uma melhor compreensão dos resultados alcançados, vale ressaltar que paralelo ao desenvolvimento das redes, também ocorreu o crescimento do banco de dados, por isso as primeiras redes foram criadas com uma quantidade menor de amostras.

A Análise de Componentes Principais, como uma maneira de identificar a correlação entre as variáveis de entrada e saída, faz-se bastante útil quando os vetores de características apresentam muitas dimensões, impossibilitando uma representação gráfica. A utilização desse tipo de análise reduz a quantidade de variáveis de entrada, sem interferir ou resultar em prejuízo significativo para o desempenho final da classificação.

Assim, para o desenvolvimento das RNA, quatro parâmetros de identificação geotécnica foram indicados após a PCA, são eles: porcentagem da fração de argila e de areia, índice de plasticidade (IP) e índice de atividade (Ia); e três variáveis qualitativas: pedologia, climatologia e geologia.

Das 180 amostras disponíveis pelo BANDASE no início da pesquisa, 87 pertenciam ao estado de Pernambuco, e 48 eram provenientes dos estados Alagoas, Maranhão e Sergipe e não apresentavam informações geotécnicas, apenas georreferências de seus locais de extração e por isso foram consideradas inadequadas à aplicação no desenvolvimento das redes neurais iniciais; as 45 amostras restantes estavam distribuídas entre os estados da Bahia, Ceará, Paraíba, Piauí, Rio Grande do Norte e Sergipe, e foram destinadas a validação às cegas dos parâmetros da rede neural, tendo em vista uma aplicação dos parâmetros da rede extrapolando o domínio com o qual a rede foi treinada (treinada apenas com dados de Pernambuco).

Como o Estado de Pernambuco tinha uma melhor caracterização das amostras, e uma maior quantidade, este foi escolhido para dar início ao desenvolvimento das redes neurais, com as redes PE04 e PE07. Para elaboração das duas RNA foram investigados os resultados dos testes de caracterização física e as correspondentes propriedades do índice dessas 87 amostras de solos de diferentes depósitos do estado, as quais apresentam características de baixa e alta plasticidade.

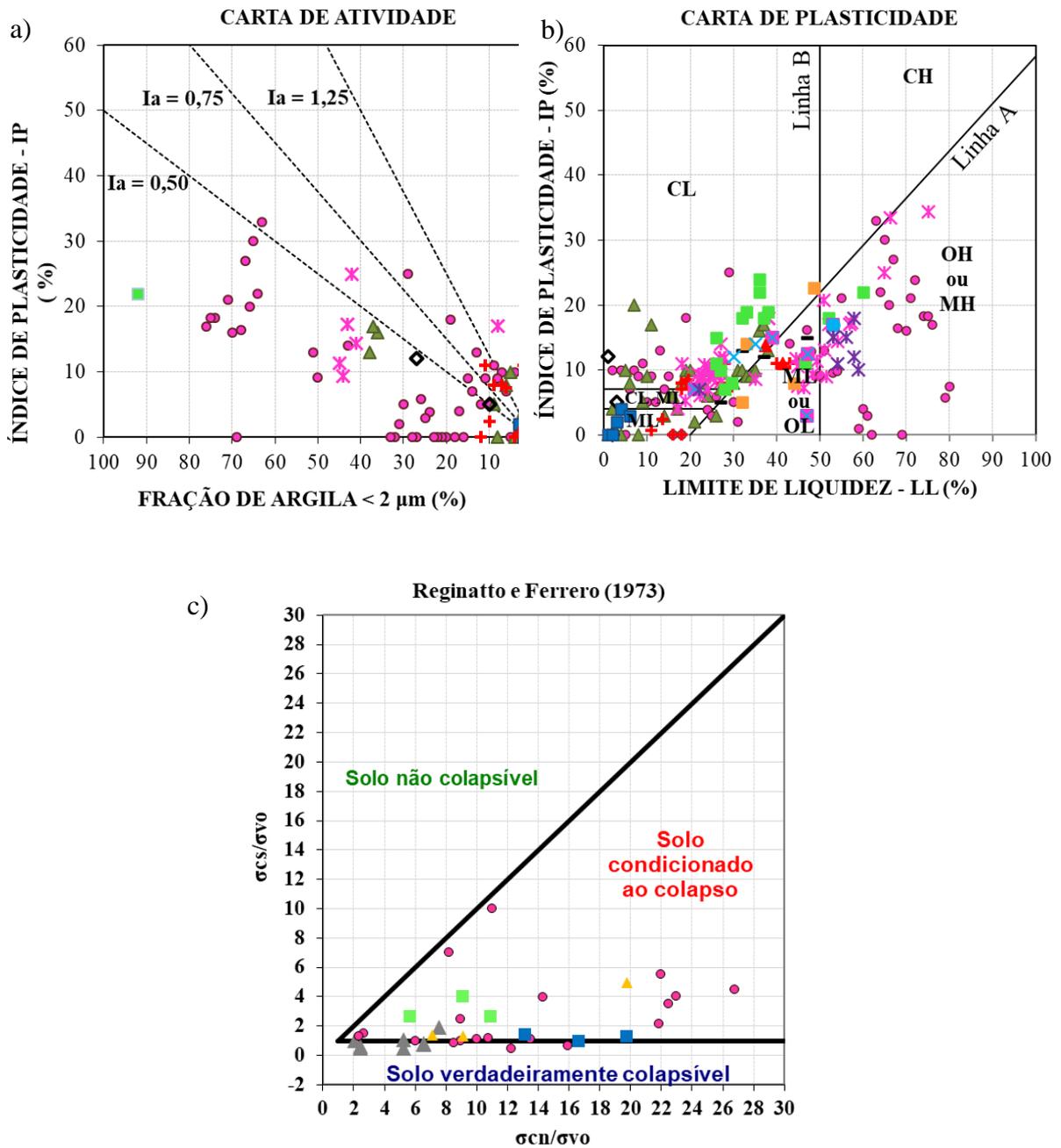
A partir da atualização do banco de dados, que passou para 526 amostras, foi possível o desenvolvimento da rede BR03, com 393 amostras de solos, que de acordo com a presença e/ou ausência de informações, apresentando condições pertinentes para a etapa de treinamento, seleção e teste de uma rede neural. As 133 amostras restantes desse banco de dados, foram utilizadas na etapa de validação, aplicando os padrões de classificação da rede ao mesmo domínio a qual foi treinada.

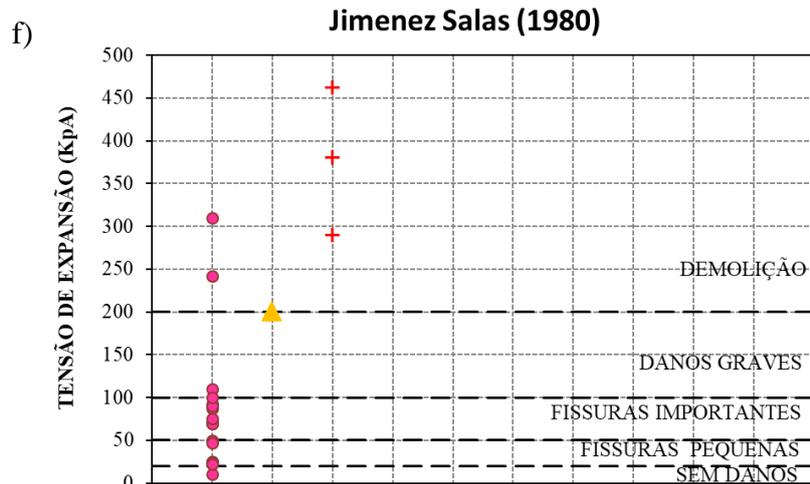
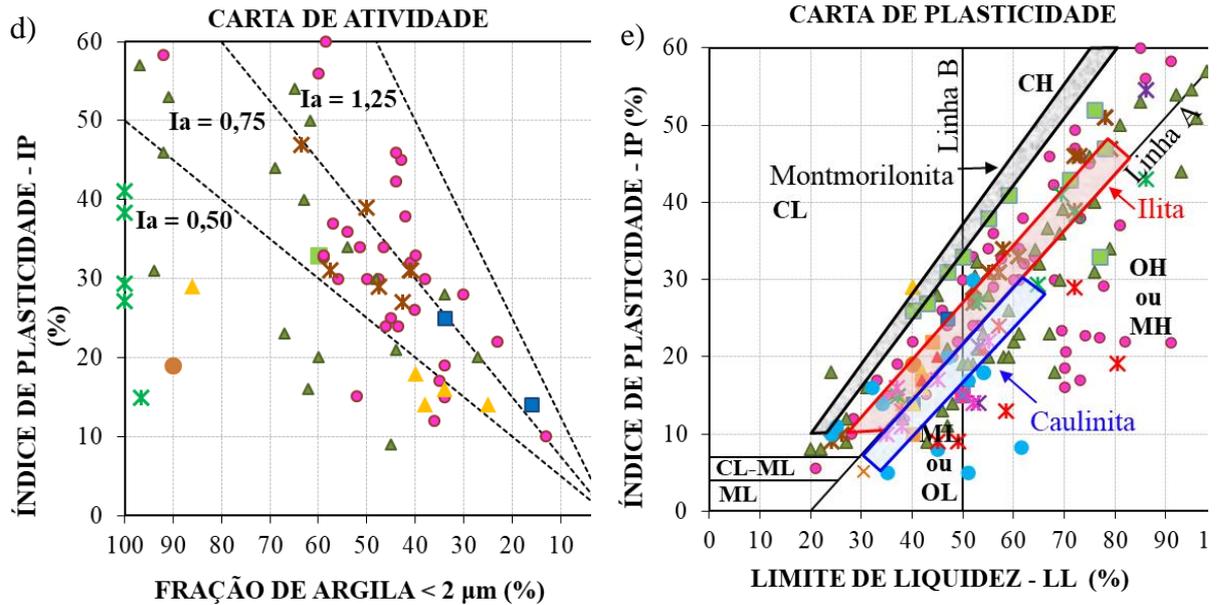
Mediante a disponibilidade dos dados e tendo em vista as variáveis de entrada para enriquecer o aprendizado da RNA, as amostras de solo situadas em mesmas coordenadas geográficas, porém coletadas em diferentes profundidades, foram consideradas como pontos distintos na coleta de dados deste trabalho.

Na Figura 10 pode ser visto as Cartas de Atividade e Plasticidade das 393 amostras do Brasil provenientes do BANDASE aplicadas ao desenvolvimento das RNA, além das características referentes as 133 amostras destinadas as etapas de validação. Na Figura 10 (a) e (c) a atividade varia nas três categorias – Ativa ($I_a = 1,25$), Normal ($I_a = 0,75$) e Inativa ($I_a = 0,50$) (Skempton, 1953), com uma concentração maior de solos expansivos nas faixas Normal e Ativa, e de solos colapsíveis na faixa Inativa.

Na Figura 10 (b) e (d) os valores de Limites de Liquidez (LL) e Índices de Plasticidade (IP) das amostras estão em grande maioria situados ligeiramente acima da Linha A, garantindo argilas inorgânicas de alta plasticidade e pouco solos orgânicos ou siltosos. Predominantemente, a fração argila está na faixa da Ilita ou entre a faixa da Montmorilonita e Ilita; em geral tais amostras apresentam grande resistência quando secas, perdendo sua capacidade de suporte por saturação, sob variações de umidade, apresentam ainda grandes variações volumétricas, retraindo-se em processos de secagem e expandindo-se sob processos de umedecimento. Quanto a compressibilidade, à direita da Linha B da Carta de Plasticidade ($LL < 50\%$) com alta compressibilidade e alta plasticidade, há alguns solos de PE, BA e SE; as demais amostras de estados (PI, CE, PB e RN) apresentam-se com baixa compressibilidade e baixa plasticidade.

Figura 10 - Cartas de Atividade (Skempton, 1953) (a) e Plasticidade dos Solos Colapsíveis (b), Critério de Identificação de Solos Colapsíveis Segundo Reginatto e Ferrero (1973) (c), Cartas de Atividade (Skempton, 1953) (d) e Plasticidade dos Solos Expansivos (e) e Critério de Identificação de Solos Expansivos Segundo Jimenes e Salas (1980) (f)





LEGENDA



Fonte: A Autora (2022).

Um outro padrão de variáveis buscou dar continuidade a pesquisa de Amorim (2004), atribuindo às redes neurais as variáveis qualitativas com características de origem e formação do solo. Dessa forma, pôde-se selecionar três grupos de variáveis de entrada, um apenas com índices físicos, outro com a adição de variáveis qualitativas (geologia, pedologia e climatologia)

aos índices físicos e outro apenas com as variáveis qualitativas; os resultados dessas seleções podem ser vistos no Capítulo IV.

As variáveis qualitativas estão relacionadas a origem e formação do solo e foram baseadas na metodologia de Amorim (2004), a qual apresentou três condicionantes do meio físico como os mais representativos para o zoneamento geotécnico dos solos colapsíveis e expansivos. Outras pesquisas já mencionadas na revisão bibliográfica também classificam os solos, quanto ao colapso e expansão, a partir de algumas dessas características de origem e formação do solo. A utilização, para fins científicos, de dados relacionados ao meio físico pressupõe que seja conhecida as coordenadas geográficas dos locais onde foram extraídas as amostras, ou seja, suas coordenadas geográficas.

A classificação Climática de Thorntwaite e Mather (1955) ocorre em função do índice hídrico e índice de aridez em 9 categorias, numa escala de umidade que vai do árido ao superúmido. Tal mapa temático fundamental foi obtido a partir da EMBRAPA (2021), com uma escala na ordem de 1:1.000.000 e possibilitou identificar a presença de solos colapsíveis e expansivos em todos os tipos climáticos, do árido ao superúmido. A classificação Pedológica (IBGE, 2019) ocorre em classes do 1º nível categórico (ordem), numa escala de 1:250.000. E a classificação Geológica (IBGE, 2019) decorre de uma unidade litoestratigráfica de mapeamento de 1:250.000.

Após a definição das variáveis de entrada e do banco de dados disponível para cada RNA, foi definida a proporção otimizada de amostras para cada etapa do desenvolvimento da rede; essa proporção segue (com pequenas diferenças) a média já consolidada nas bibliografias de RNA, inclusive nos estudos que referenciam essa pesquisa, que é 60% para treinamento, 20% para seleção e 20% para teste.

A nomenclatura das RNA desenvolvidas vem das características de variáveis de entrada e procedência dos dados de treinamento, como pode ser visto no Quadro 4. A rede PE04 foi desenvolvida com base em 87 amostras de solos de Pernambuco e com 4 variáveis de entrada (4 índices físicos): porcentagem da fração de argila e de areia, índice de plasticidade e índice de atividade. Do total de amostras, houve uma melhor proporção com: 61% (53 amostras) para treinamento, 19,5% (17 amostras) para seleção e 19,5% (17 amostras) para teste. Vale ressaltar que é fundamental que, a distribuição dos dados esteja equivalente, para que a rede possa identificar ambas as situações: colapso e expansão; e para que as amostras colapsíveis e expansivas sejam escolhidas aleatoriamente pelo *Neural Designer*.

A rede PE07 foi desenvolvida com o mesmo banco de dados da rede anterior (rede PE04), entretanto, foi acrescentado às variáveis de entrada geotécnicas, três variáveis condicionantes do meio físico: climatologia, geologia e pedologia, totalizando 7 variáveis de entrada. Como o banco de dados é o mesmo, a divisão das amostras em grupos de desenvolvimento seguiu as mesmas quantidades da rede PE04.

Já a terceira rede, BR03, foi desenvolvida com base em 393 amostras de solo de todo o território brasileiro, usando apenas 3 variáveis de entrada: climatologia, geologia e pedologia e seguindo o mesmo parâmetro de subdivisão das redes anteriores, para as etapas de desenvolvimento, distribuiu-se em: 237 (60,3%) amostras para o treinamento, 78 (19,8%) para seleção e 78 (19,8%) para teste.

Quadro 4 - Características das Principais RNA Desenvolvidas

Rede	Variáveis de Entrada	Amostras para desenvolvimento da rede	Nº de neurônios	Função de Ativação	Validação
PE04	% de areia, % de argila, IP e Ia.	Treinamento – 53 Seleção – 17 Teste – 17 Total – 87.	2 - 1 neurônios em 2 camadas ocultas.	Função Logarítmica para as duas camadas ocultas/ <i>perceptron</i> e Função Probabilística para a camada de saída.	Amostras do Nordeste – 45.
PE07	% de areia, % de argila, IP, Ia, Clima, Geologia e Pedologia.	Treinamento – 53 Seleção – 17 Teste – 17 Total – 87.	3 - 1 neurônios em 2 camadas ocultas.		Amostras do Nordeste – 284. Amostras do Brasil – 463.
BR03	Clima, Geologia e Pedologia.	Treinamento – 237 Seleção – 78 Teste – 78 Total – 393.	3 – 1 neurônios em 2 camadas ocultas.		Amostras do Brasil – 133.

Fonte: A Autora (2022).

Vale salientar, que as redes não usaram valores quantitativos ou valores qualitativos da classificação da suscetibilidade do SIGSEC-PE, essa usou métodos indiretos, a partir de opinião experta para caracterizar e classificar os solos. No caso das RNA, houve interesse de deixá-la isenta de quaisquer possibilidades de subjetividade de interpretações.

No caso das variáveis com registros ausentes, aqueles que, por algum motivo, faltam no conjunto de dados, mas existem no contexto em que a variável foi dimensionada na rede, geralmente, são indicadas por um rótulo no conjunto de dados. Alguns rótulos padrões usados no programa para representar valores ausentes são: ‘NA’ (não disponível), ‘NaN’ (não é um número), ‘Desconhecido’ ou ‘?’. Nesses casos não se deve usar valores numéricos nulos como zeros, para que a falta de informação não seja confundida com um valor real.

A Tabela 1 mostra os valores usados para caracterizar as variáveis de entrada, que incluem o mínimo, o máximo, a média e o desvio padrão para cada uma das variáveis de entrada, das três redes; segundo o banco de dados (colapso e expansão) usado por cada uma delas, nas três etapas de desenvolvimento. No caso das variáveis de entrada qualitativas, esses valores são representados pela menor e maior quantidade de ocorrência de amostras nas classes indicadas, sendo assim inexistentes para os valores de média e desvio padrão.

Tabela 1 – Estatística das variáveis que compõem as amostras para o desenvolvimento das RNA. Rede PE04 (a), Rede PE07 (b) e Rede BR03 (c).

Rede PE04					Rede PE07					
	Mínimo	Máximo	Média	Desvio		Mínimo	Máximo	Média	Desvio	
a)	Argila	2.10	94.00	30.78	20.78	Argila	2.10	94.00	31.31	20.73
	Areia	3.00	94.00	48.16	25.80	Areia	3.00	94.00	47.66	25.73
	IP	0.00	60.00	18.47	14.86	IP	0.00	60.00	17.25	15.04
	la	0.11	2.75	0.66	0.43	la	0.11	2.75	0.66	0.42
	EXP	0.00	1.00	0.53	0.50	GEO1	Reconcavo	Borborema	-	-
						PED1	Planossolo	Neossolo	-	-
						CLIM1	Subumido	Árido	-	-
						EXP	0.00	1.00	0.54	0.50
c)					Rede BR03					
		Mínimo	Máximo	Média	Desvio					
	PEDOLOGIA	Plintossolo	Latossolo	-	-					
	CLIMATOLOGIA	sub u seco	umido	-	-					
	GEOLOGIA	Amazonia	SaoFrancisco	-	-					
	SOLO(EXP/COL)	0.00	1.00	0.47	0.50					

Fonte: A Autora (2022).

Para as variáveis quantitativas, pode-se observar que essas contam com uma grande variabilidade em cada uma de suas características, melhorando a qualidade do banco de dados,

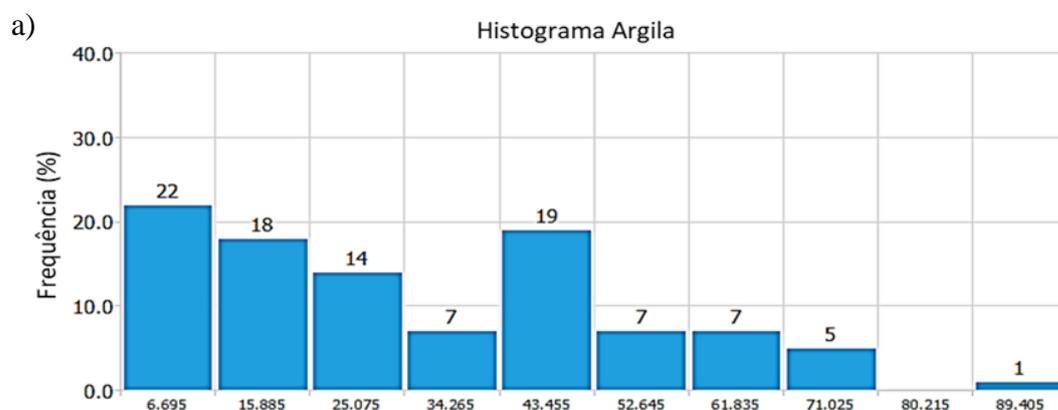
tendo em vista a importância de reconhecimento dos padrões de colapso e expansão. O ideal de variabilidade é apresentar características à rede, na etapa de treinamento e seleção, de modo a obter uma melhor generalização de padrões.

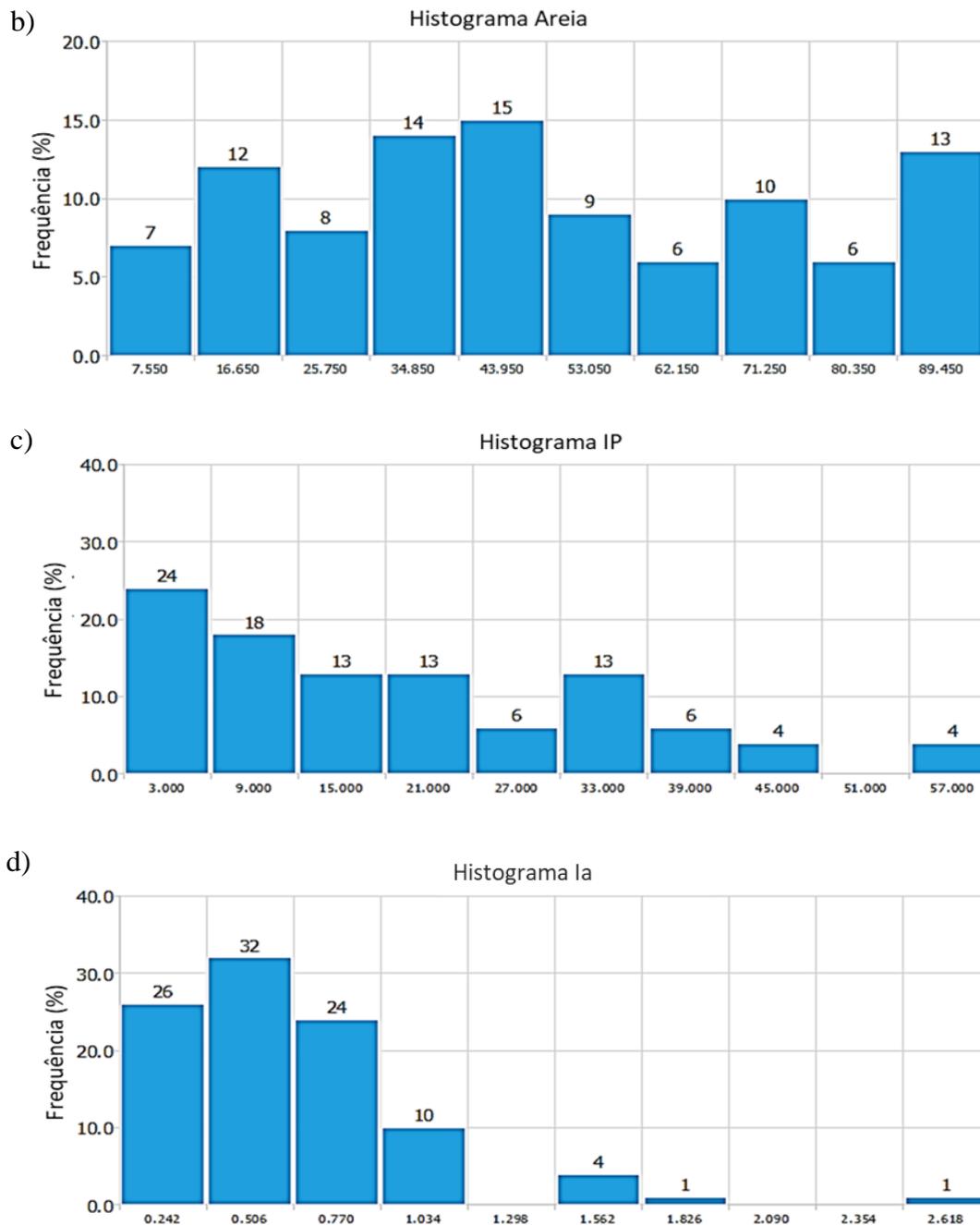
3.2.3 Características e Correlações dos Dados de Entrada das Redes Neurais Artificiais

a) Rede PE04

A primeira rede neural aqui detalhada, conta com 4 variáveis de entrada de caráter geotécnico. Na Figura 11 são apresentados os histogramas dos dados de entrada. Segundo as 87 amostras da rede PE04 e de acordo com a distribuição dos dados de cada variável tem-se: % de Argila = bimodal, % de Areia = multimodal e IP e Ia = distorcido a direita. A variação na apresentação dos dados está associada a reduzida quantidade de amostras disponibilizadas para essa etapa de elaboração de gráficos. Os histogramas de Argila e Areia retratam uma mistura de dados com médias diferentes, obtidas em condições distintas; enquanto para IP e Ia a distribuição é característica de parâmetros que possuem apenas um limite de especificação, além da representação de picos isolados que podem significar anormalidades temporárias no processo e erros de medição ou registro.

Figura 11 – Histograma das variáveis escalares, segundo as 87 amostras da rede PE04: % de Argila (a), % de Areia (b), Índice de Plasticidade - IP (c) e Índice de Atividade - Ia (d)





Fonte: A Autora (2022).

Para identificar redundâncias entre as variáveis de entrada, é usada uma matriz de correlação, de modo a identificar correlações perfeitas (1,00), possibilitando a remoção de uma dessas sem qualquer perda de informação. Os resultados mediante a correlação entre as variáveis definidas pelo *Neural Designer* foram apresentados na matriz da Figura 12, em função dos valores absolutos das correlações lineares entre todas as entradas, confirmando o que geotecnica é esperado para características opostas de colapso e expansão. A correlação mínima entre as variáveis % de Argila e % de Areia é de -0,84. A correlação máxima é de 0,83 entre as variáveis:

% de Argila e IP. A variável Ia embora não tenha atingido correlação acima de 0,50, foi mantida como entrada da rede neural; a grande ausência dessa informação nas 87 amostras desse banco de dados proporcionou essa tendência, entretanto, os estudos bibliográficos nos direcionam a mantê-la, e buscar melhorar a representatividade e variação dessa entrada ao longo da pesquisa.

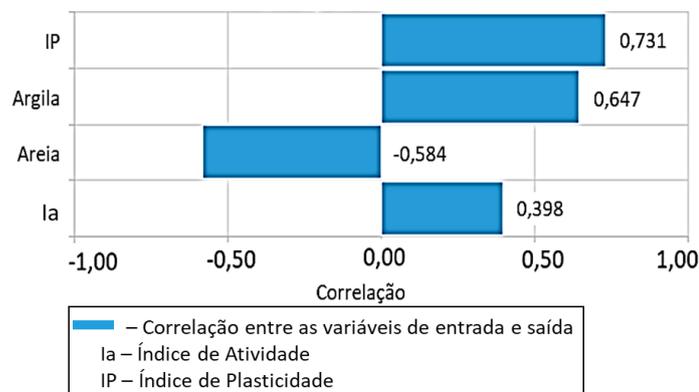
Figura 12 – Matriz de correlação entre as variáveis principais pelo *Neural Designer*

	Argila	Areia	IP	Ia
Argila	1	-0.84	0.83	-0.23
Areia		1	-0.76	0.11
IP			1	0.25
IA				1

Fonte: A Autora (2022).

Além disso, é importante identificar as dependências entre variáveis entrada e a variável de saída. A Figura 13, mostra esses valores de correlação, definidos pela rede neural, para cada variável de entrada da rede PE04 e sua respectiva saída. Correlações próximas de 1 significam que a variável de saída está integralmente correlacionada com a dita variável de entrada. Correlações próximas de 0 implicam em ausência de correlação entre a entrada e a variável de saída. Em geral, os resultados de saída das redes neurais dependem de muitas entradas simultaneamente.

Figura 13 - Principais correlações entre as variáveis de entrada e saída



Fonte: A Autora (2022).

Ratificando a aplicabilidade de tais variáveis de entrada, tem-se a classificação de solos com algumas características que baseiam a identificação de solos expansivos segundo Seed *et al.*

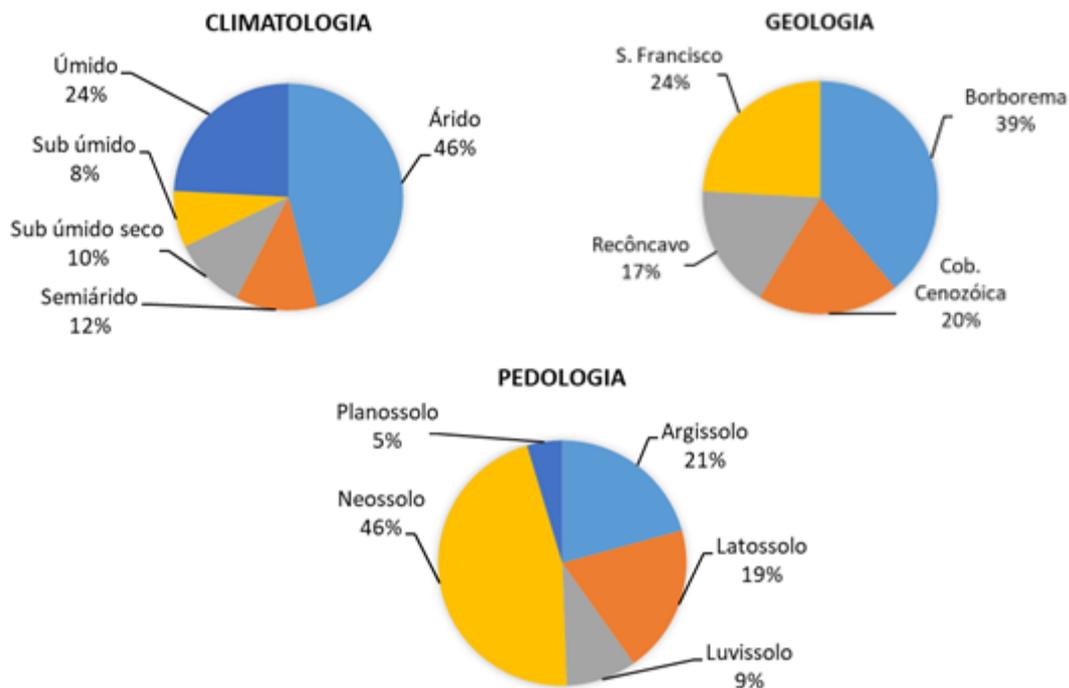
(1962) como: porcentagem da fração argila e índice de atividade, e ainda segundo Chen (1975), com Índice de Plasticidade diretamente proporcionais, enquanto o percentual de Areia tem sua correlação inversa. A pouca variabilidade dos valores de IA também influenciam em uma correlação menor que as demais variáveis de entrada.

b) Rede PE07

Com a mesma metodologia descrita para definir todos os parâmetros da rede PE04, também será desenvolvida a rede PE07, detalhando as características das novas variáveis de entrada atribuídas à rede, bem como a correlação entre elas e as variáveis de saída.

Como o banco de dados das amostras não sofreu qualquer alteração, contando ainda com 87 amostras de solos colapsíveis e expansivos de Pernambuco, será descrita a informação adicional de caracterização e correlação das amostras que é referente as novas variáveis de entrada. A caracterização das variáveis categóricas de entrada, referentes as 87 amostras usadas no desenvolvimento da rede, foi representada por gráficos na Figura 14.

Figura 14 – Caracterização das amostras da rede PE07, segundo as variáveis categóricas: Climatologia, Geologia e Pedologia.



Fonte: A Autora (2022).

A variável climatológica que apresenta predominância das amostras de Pernambuco é o clima Árido. A colapsibilidade e expansividade se manifestam em função da mudança de umidade, formando zonas de suscetibilidade a esses solos que coincidem com essas áreas climáticas deficitárias de umidade, principalmente em climas áridos e semiáridos.

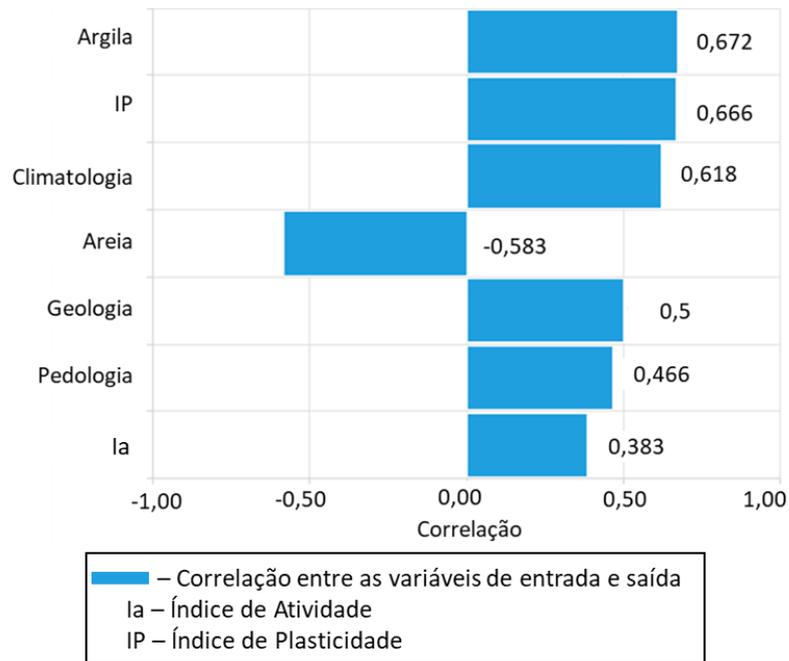
Os critérios geológicos que determinam se uma unidade geológica é ou não geradora de um solo colapsível ou expansivo foram baseados em Ferreira (1999) e Vasconcelos (2001) na sua composição mineralógica primária, nos prováveis produtos do intemperismo e presença ou não de formação dos sedimentos. Dessa forma, a predominância da Província da Borborema que apresenta estratigrafia e geocronologia complexas, definindo uma série de compartimentos tectônicos, caracterizado por diferentes aspectos geológicos e geofísicos, apresenta uma suscetibilidade média ao colapso e expansão.

Para a Pedologia, há a presença predominante do Neossolo que são desenvolvidos de sedimentos recentes, geralmente origem fluvial, e do Argissolo que, com um acentuado teor de argila em profundidade, apresenta condutividade hidráulica baixa, causando restrição a percolação de água.

As representações numéricas das correlações (Figura 15) destacam a correlação diretamente proporcional entre todas as variáveis de entrada e saída, com exceção da variável % de Areia que apresentou comportamento inversamente proporcional, como já havia sido observado na rede PE04. Destaque para as variáveis % de argila e IP, que confirmaram as maiores correlações entre as variáveis de entrada e saída, com a variável categórica Clima.

Quanto aos fatores climáticos, a sazonalidade da região varia a umidade da área, modificando tanto o nível dos lençóis freáticos como as condições hidro geológicas, possibilitando uma correlação diretamente proporcional de 0,618. Os fatores geológicos trazem a presença de micas e minerais félsicos no tipo de formação dos sedimentos e uma correlação de 0,5 para a rede. E, tendo a Pedologia como uma variável categórica, que permite uma correspondência com a ocorrência do lençol d'água, o grau de saturação, a atividade do solo, a presença de minerais colapsíveis ou expansivos, as características de drenagem, erodibilidade e de plasticidade, a ocorrência de solos porosos, permeáveis e impermeáveis, além de trazer informações acerca dos solos de jazidas, apresentando uma correlação diretamente proporcional de 0,466.

Figura 15 - Principais correlações entre as variáveis de entrada e saída.



Fonte: A Autora (2022).

c) Rede BR03

Observando o desempenho das redes anteriores e com foco no objetivo inicial de reduzir a quantidade de informações necessárias para a classificação da amostra pela rede neural, foi desenvolvida a rede BR03, de modo a ser necessário apenas a georreferencia das amostras, para que sejam obtidas as informações de origem e formação dos solos (Climatologia, Geologia e Pedologia).

Como as variáveis de entrada da rede são integralmente categóricas, nenhuma média e desvio padrão são apresentados para elas, entretanto é possível observar as características de clima, geologia e pedologia, que se repetem com o menor e maior frequência nas amostras destinadas ao desenvolvimento. O detalhamento da distribuição dessas variáveis é representado pelo gráfico da Figura 16.

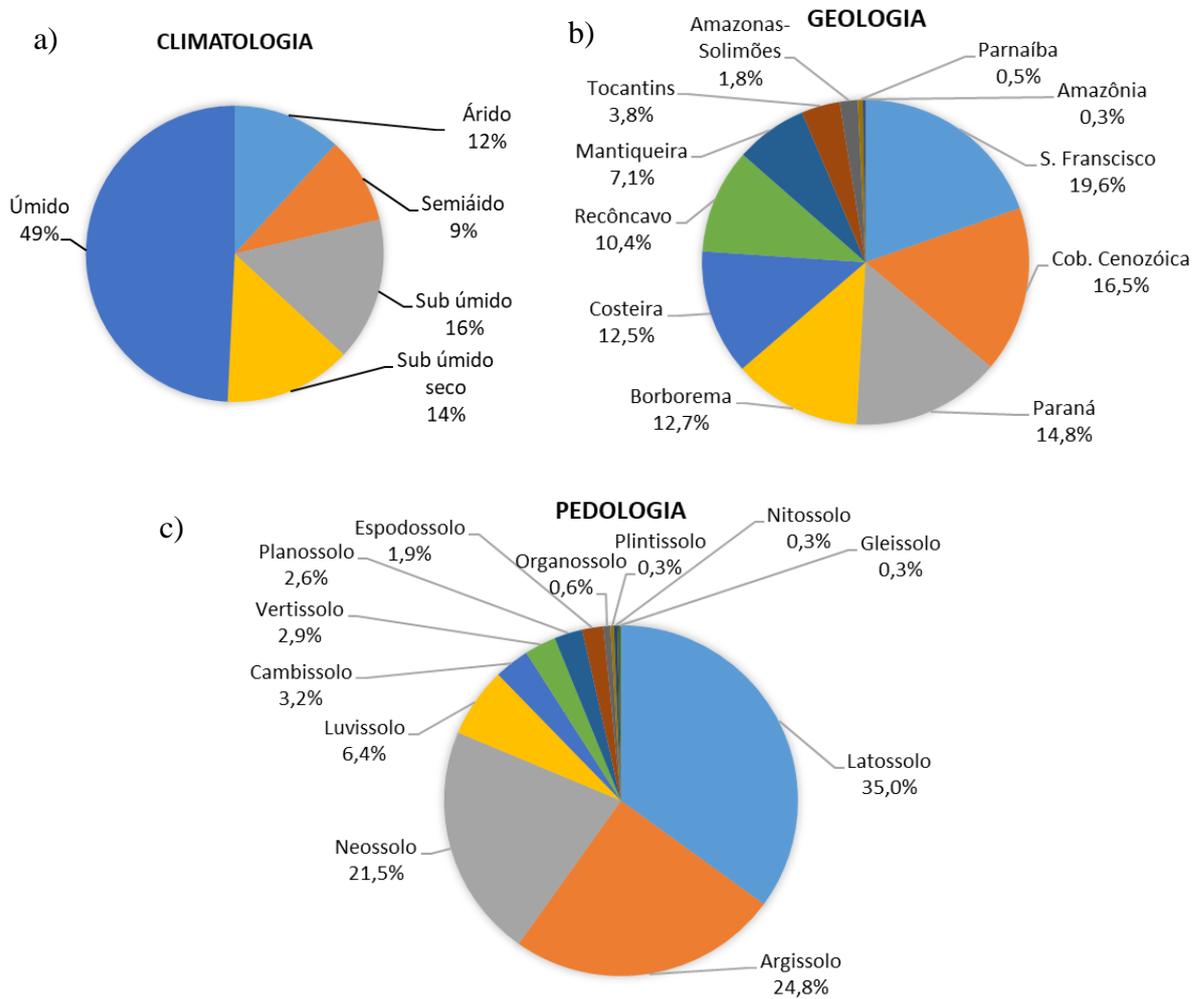
A maioria das amostras do Brasil apresentou características de áreas de clima úmido e sub úmido (Figura 16-a), consideradas com suscetibilidade baixa ao colapso e a expansão de acordo com Amorim (2004), por caracterizarem áreas de superávit hídrico, não apresentando estações contrastadas (seca e chuvosa), observadas em climas áridos e semiáridos; propiciando ao solo teores de umidade mais elevados nas camadas superficiais do terreno.

Na Geologia (Figura 16-b), as unidades que predominaram a composição das amostras foram Belém do São Francisco e a Cobertura Cenozoica. Na unidade Belém do São Francisco, os solos são ricos em minerais félsicos (quartzo e feldspato), considerados potencialmente colapsíveis, embora em um grau mais baixo, tendo em vista a maior velocidade de intemperização, devido a estrutura dos gnaisses e micaxistos. Os solos sedimentares da Cobertura Cenozoica, de texturas arenosas, se distribuem de modo irregular, ocorrendo de modo restrito e sem associação direta com as Bacias; apesar de serem consideradas potencialmente colapsíveis, a parte mais oriental dessas coberturas apresenta propriedades diferenciadas quando próximas ao litoral, sendo consideradas potencialmente expansivas onde há concentração de argilas (AMORIM, 2004).

Na Pedologia há uma maior concentração de Latossolos e Argissolos (Figura 16-c) que compõem 60% do território nacional. São solos profundos, altamente intemperizados, ácidos, de baixa fertilidade natural e, em certos casos, com alta saturação por alumínio. Os Latossolos são solos não hidromórficos em avançado estágio de intemperização, geralmente são fortemente drenados e muito profundos, assim não apresentam características típicas de solos expansivos, entretanto apresentam uma média suscetibilidade ao colapso. Já os Argissolos são minerais não hidromórficos, apresentando moderada cerosidade, associada a estrutura fraca a moderada e argilas de baixa atividade.

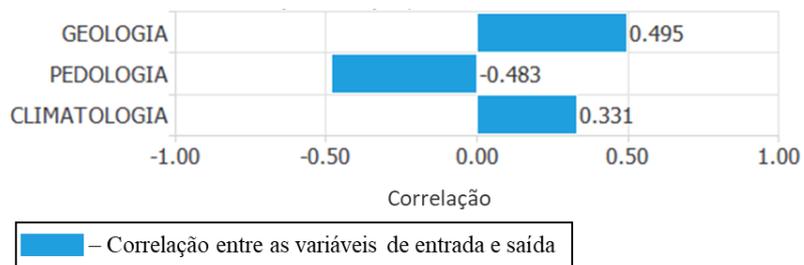
O grau de correlação entre as variáveis de entrada e saída destacam a relação diretamente proporcional entre as variáveis geologia (0,495) e climatologia (0,331), e uma correlação inversamente proporcional quanto a variável pedologia, para essa classe foi verificado que 60% das amostras pertencem a classe dos Latossolos e Argissolos (Figura 17) justificando a inversão da correlação.

Figura 16 – Correlação entre as variáveis categóricas da Rede BR03. Climatologia (a), Geologia (b) e Pedologia (c).



Fonte: A Autora (2022).

Figura 17 - Principais correlações entre as variáveis de entrada e saída



Fonte: A Autora (2022).

3.3 Estratégias de Treinamento das RNA

Nas estratégias de treinamento das redes neurais são definidos os critérios de parada de minimização dos erros, para as etapas de treinamento e seleção (variando para cada algoritmo escolhido), além de estratégias aplicadas ao conjunto de dados, como: definição do ‘Índice de Perda’ e do ‘Algoritmo de Otimização’, que ajustarão a rede neural às informações do banco de dados.

O Índice de Perda define uma medida da qualidade da representação necessária para o aprendizado das classes, sendo caracterizado pelo somatório entre o Termo de Erro e o Termo de Regularização. O Termo de Erro estima como a rede se ajusta ao conjunto de dados, sejam dados de treinamento ou dados de seleção; o Termo de Regularização é um parâmetro ponderador, de modo a reduzir a extensão de ocorrência do *overfitting*; ocorrendo quando há uma rede com quantidade de dados de treinamento limitada, por meio do ajuste de pesos sinápticos e bias, para o cálculo do gradiente iterativo (critério de parada), aumentando a qualidade de previsão da rede e reduzindo os erros em cada etapa, treinamento e seleção; se a resposta da rede for muito suave, o peso da regularização deverá ser diminuído, se a resposta da rede oscilar demais, o peso aumentará. Para as redes criadas no desenvolvimento dessa pesquisa, o tipo de erro definido para calcular o Índice de Perda foi: Erro ao Quadrado Ponderado (WSE) que é usado em aplicações de classificação binária com saídas desequilibradas, ou seja, quando o número de amostras positivas e negativas é diferente, que é o caso do banco de dados utilizado com uma quantidade diferente de amostras colapsíveis e expansivas.

O Algoritmo de Otimização também depende da quantidade e das características das amostras disponíveis no banco de dados, determinando como ocorre o ajuste dos parâmetros na rede neural, identificando o melhor critério de parada entre: um número máximo de iterações atingido, o Índice de Perda minimizado para um valor objetivo, o tempo máximo de operação computacional atingido, etc. O Método Quase-Newton (QNM), como algoritmo de otimização, usa a função Hessiana da perda, uma matriz de segundas derivadas, para calcular a direção da aprendizagem, que aponta para o mínimo da função de perda com maior precisão.

As mesmas estratégias de treinamento foram aplicadas aos conjuntos de dados de todas as redes, criadas de modo a manter os padrões de comparação entre elas.

A seleção da melhor arquitetura de rede é determinada usando o dimensionamento da camada de entrada, a partir do Desvio Médio Padrão, resultando em um modelo cuja complexidade ajuste adequadamente os dados e seja responsável por encontrar o número ideal de neurônios na rede, com maior capacidade de generalização.

Na etapa de testes, os resultados de previsão da rede são comparados com as condições observadas por cada amostra, a partir dos métodos de classificação binária: Matriz de Confusão, Métricas de Avaliação Binária, Gráfico ROC e Gráfico de Ganho Acumulado, confirmando assim a qualidade da RNA.

Na análise do desempenho das redes a partir da matriz de confusão, as colunas representam os resultados observados, medidos em campo no conjunto de dados, e as linhas representam os resultados previstos pela rede neural. Alguns dados derivados da matriz de confusão podem ser extremamente úteis na avaliação do modelo, ao aplicá-lo a um conjunto inteiro de exemplos, cada um desses quatro resultados (parâmetros derivados – TP, FP, TN, FN) e seus derivados serão empregados nos critérios de desempenho. Como cada estudo tem suas peculiaridades, que devem ser levadas em consideração na escolha de como o teste de classificação será avaliado, não se deve pensar em um critério como melhor ou pior que o outro, de maneira geral, o que deve ser feito é analisar o problema e escolher os critérios que melhor se adequem. Para esta pesquisa serão observados os valores de precisão da classificação, taxa de erro, sensibilidade e especificidade, tendo em vista que ambas as situações de erro, acarretam problemas de reconhecimento das características do solo. Os parâmetros derivados da matriz de confusão dão origem ao Gráfico ROC, que apresenta algumas vantagens usado nas análises de classificação como: poder fazer uma análise independentemente de certas condições, tais como o limiar de classificação e a distribuição a priori das classes; calibrar a rede e ajustar os modelos de classificação, quando necessário, analisando-se os possíveis ajustes entre as classificações colapsíveis e expansivas. Dada uma condição operacional, é possível derivar o limiar apropriado, quanto mais distante a curva estiver da diagonal principal (acima dessa), melhor será o desempenho do sistema de aprendizado para aquele domínio.

A capacidade de previsão também está diretamente relacionada a quantidade e qualidade do banco de dados, quanto maior o banco de dados e quanto menor a ausência de informações, melhor o desempenho da RNA; além disso, quanto maior a variabilidade das características das amostras, alcançando os limites de extrapolação das variáveis de entrada, melhor a qualidade de previsão das RNA.

A nomenclatura dessas etapas (treinamento, seleção e teste) pode variar para algumas literaturas (MEROUANE E MAMOUNE, 2018; BARROS, 2019; IKEAGWUANI E NWONU, 2019; CHOUDHURY E COSTA, 2019), entretanto, a nomenclatura adotada nessa pesquisa segue o padrão usado pelo *Neural Designer*, o que não interfere nos resultados da divisão e proporção do banco de dados.

3.4 Aplicação dos Padrões das RNA

A partir do conceito de generalização de padrões no aprendizado de máquina, a aplicação de uma rede para a previsão de padrões em amostras totalmente desconhecidas - que não foram consideradas durante o processo de aprendizagem ou teste; são propagadas de maneira *feed-forward*, através das camadas *perceptron* e da camada probabilística, respectivamente.

Para generalizar os padrões da rede PE04, treinada apenas com amostras de solos de Pernambuco, foram utilizadas 45 amostras de solos dos demais estados do Nordeste. Na generalização dos padrões da Rede PE07 – também treinado apenas de amostras de solos de Pernambuco, dois grupos diferentes foram verificados: um com 284 novas amostras de solos do Nordeste e outro com 463 amostras de solos de todo o Brasil. Para generalizar os padrões da Rede BR03, desenvolvida com amostras de todo o Brasil, 133 diferentes amostras desse universo foram utilizadas para validar os padrões da rede, conforme descrito, anteriormente, no Quadro 4.

A progressão nas quantidades de amostras dos grupos de generalização dos padrões também ocorreu à medida que o banco de dados foi sendo expandido, esse banco de dados continua em constante crescimento, alimentado pelas pesquisas do GNsat.

Cada previsão feita pelas redes para cada amostra testada foi comparada as informações observadas em ensaios laboratoriais, de modo a certificar o erro ou acerto da previsão da rede, possibilitando uma avaliação do desempenho.

3.5 Elaboração de Programa de Classificação de Solos

A partir da equação resultante da rede neural, desenvolvida com o melhor desempenho, e com base nos testes de classificação binária, foi possível elaborar um programa de classificação de

amostras de solos, criado em linguagem C# com sintaxe e programação orientada a componentes, que de acordo com as coordenadas de latitude e longitude, localiza os valores de entrada nos mapas fundamentais e aplica à equação da rede.

O programa deve fazer previsões de amostras quanto as características de colapso e expansão, classificando-as quanto as probabilidades de suscetibilidade ao colapso em intervalos de: alta (0,0 a 0,2), média (0,2 a 0,3) e baixa (0,3 a 0,5), e quanto as probabilidades de suscetibilidade a expansão em: alta (0,5 a 0,7), média (0,7 a 0,8) e baixa (0,8 a 1,0), de acordo com o índice assumido pela amostra após classificação, como pode ser observado na Tabela 2.

Tabela 2 - Índice de Classificação dos Solos pelo Programa.

Classificação do Solo	Índice de Classificação
Suscetibilidade Alta - Colapso	0,000 – 0,200
Suscetibilidade Média - Colapso	0,201 – 0,300
Suscetibilidade Baixa - Colapso	0,301 – 0,500
Suscetibilidade Baixa - Expansão	0,501 – 0,700
Suscetibilidade Média - Expansão	0,701 – 0,800
Suscetibilidade Alta - Expansão	0,801 – 1,000

Fonte: A Autora (2022).

As amostras de solos que estiverem localizadas em áreas com característica Geológica identificada como ‘Corpo d’água’ ou característica Pedológica identificada como ‘Corpo D’Água’ ou ‘Outros’, não prosseguirão com a classificação, tendo em vista o caráter desses atributos e a falta de propriedades do solo.

Deve-se observar que as amostras de solos que não forem colapsíveis ou expansivas deverão tender a um índice de suscetibilidade com uma menor certeza de classificação, prevendo uma suscetibilidade baixa às características, assumindo valores entre o intervalo de 0,30 a 0,70. As demais classificações, como possuem origem em uma variável qualitativa, para manter-se isentas de opinião experta, tem seus valores de classificação definidos a partir das informações do banco de dados atribuídas à RNA na etapa de treinamento (fundamentos da rede neural supervisionada).

Os índices de classificação, apresentados na Tabela 2, foram definidos a partir das informações provenientes do histograma das populações em um processo iterativo de comparação entre os

resultados da classificação com a real classificação das amostras. Assim, foi possível identificar um intervalo numérico de maior incerteza, e de maior certeza, de classificação pelo programa.

3.6 Carta de Classificação de Solos Colapsíveis e Expansivos do Brasil

Apesar de o programa gerar respostas para qualquer ponto localizado dentro do território brasileiro, desde que se localize em área com informações das variáveis condicionantes, essa seria uma informação pontual. Diante disso, houve o interesse de estender a informação para toda a área do território brasileiro, cobrindo espacialmente todo o mapa.

Segundo a equação proveniente da melhor rede desenvolvida, foi possível elaborar as cartas de suscetibilidade ao colapso e a expansão, a partir da união dos mapas fundamentais de Clima (EMBRAPA, 2021) – escala de 1:1.000.000, Geologia (IBGE, 2019) – escala de 1:250.000 e Pedologia (IBGE, 2019) - escala de 1:250.000, que resultou em 1.264 unidades de mapeamento. Cada área teve suas características submetidas a classificação da rede de melhor desempenho e foram identificadas as suscetibilidades (alta, média ou baixa) ao colapso ou expansão para todo o território do Brasil (Figura 18).

Figura 18 - Processo de Elaboração das Cartas de Suscetibilidade



Fonte: A Autora (2022).

4 RESULTADOS

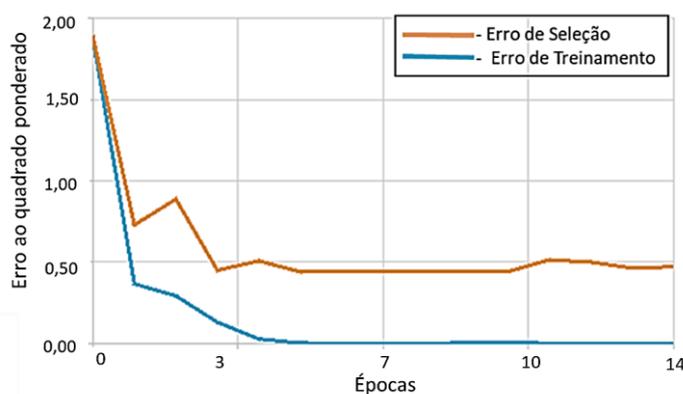
São apresentados os resultados alcançados na definição das variáveis usadas para elaboração das RNA, bem como o desenvolvimento dessas redes mediante à aplicação de estratégias ao banco de dados, o resultado dos testes de desempenho de cada rede, e a generalização dos padrões de previsão a novos dados desconhecidos pelas redes; e por fim as orientações a respeito do programa de classificação dos solos e apresentação da carta de suscetibilidade de solo colapsível e de solo expansivo do Brasil.

4.1 Rede Neural PE04

4.1.1 Composição da RNA PE04

Como resultado a composição da rede PE04, para um algoritmo de retro propagação do erro, o critério de parada procura encontrar um valor mínimo global. Os pesos sinápticos da rede, obtidos a partir das determinações dos erros de treinamento e seleção foram sendo ajustados segundo a minimização da função do erro, pelo método do gradiente descendente, observado na Figura 19, a cada iteração. A linha azul representa o erro de treinamento, e tende a diminuir rapidamente; a linha laranja representa o erro de seleção. O valor inicial do erro de treinamento é 1,9 e o valor final após 14 iterações/épocas é 0. O valor inicial do erro de seleção é 1,9 e o valor final após 14 iterações/épocas é 0,47.

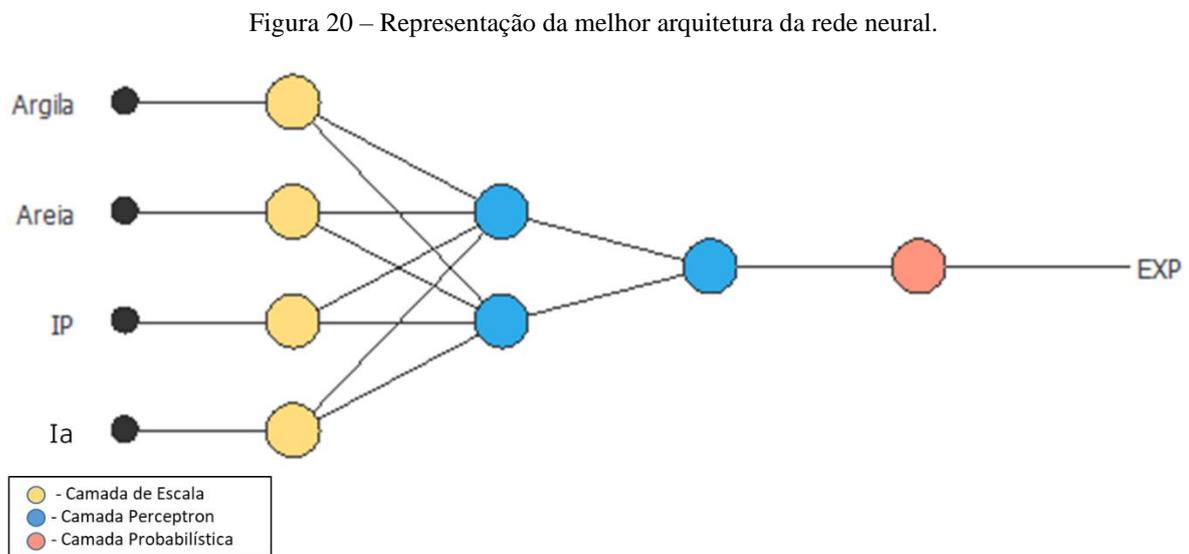
Figura 19 – Gráfico do histórico dos erros do método *Quase-Newton*.



Fonte: A Autora (2022).

A escolha da arquitetura definida para a rede PE04, ocorreu a partir de testes na etapa de seleção de modelos usando algoritmos genéticos para apontar o melhor número de camadas e de

neurônios por camadas, desse modo, contando com 4 variáveis de entrada (círculos pretos), uma camada escalar (círculos amarelos), na qual ocorre um pré-tratamento dos dados de entrada, ocorrendo a normalização de dados escalares ou binarização de dados categóricos; duas camadas *perceptron* ou rede neural propriamente dita (círculos azuis), que têm seus neurônios ativados pela função logística, devido a característica binária das variáveis de saída; e uma camada probabilística (círculos vermelhos), com resultados variando entre 0 e 1 (Figura 20).



Fonte: A Autora (2022).

4.1.2 Teste – Análise do Desempenho

Das 17 amostras destinadas para esta etapa, e 4 (23,6%) foram classificadas incorretamente, enquanto 13 (76,5%) tiveram uma correta classificação (Tabela 3-a). A partir desses parâmetros pode-se observar que a rede PE04 alcançou uma acurácia de 76,5% e, conseqüentemente, uma taxa de erro de 23,5%; da taxa de acertos (acurácia), a característica melhor identificada foi o colapso, tendo em vista que a taxa de especificidade - o percentual de amostras colapsáveis, classificadas corretamente pela rede PE04, foi de 77,8%, enquanto a taxa de sensibilidade - o percentual de amostras expansivas, classificadas corretamente, foi de 75% (Tabela 3-b).

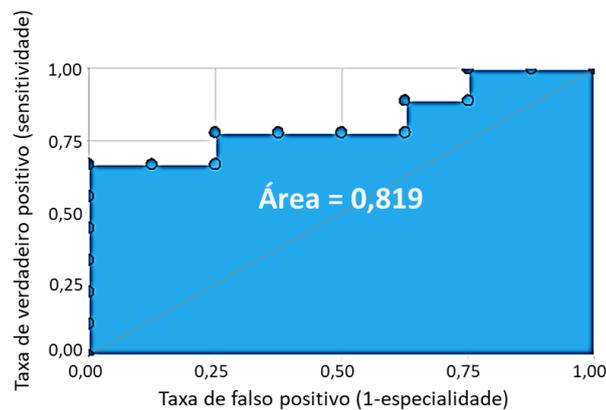
Tabela 3 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b).

a)		b)	
6 (Verdadeiras Expansivas)	2 (Falsas Expansivas)	Acurácia	76,5%
2 (Falsas Colapsíveis)	7 (Verdadeiras Colapsíveis)	Taxa de Erro	23,5%
		Sensitividade	75%
		Especificidade	77,8%

Fonte: A Autora (2022).

Pela análise do gráfico ROC, foi observado uma área sob a curva para a rede PE04 de 0,819 (Figura 21) e o limiar ideal de decisão de classificação das amostras foi de 0,476. A curva se apresenta com uma característica com “degraus” ou assimetrias, devido a utilização de dados experimentais não normais.

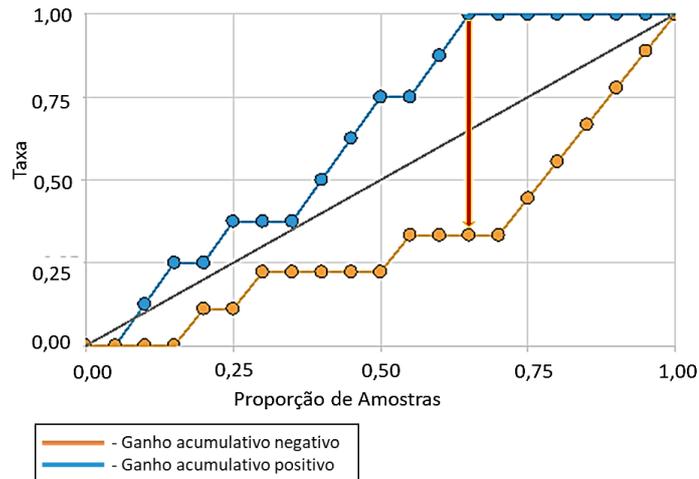
Figura 21 – Gráfico ROC.



Fonte: A Autora (2022).

O teste do Ganho Acumulado – potencial da rede diferenciar mais rapidamente as características dos solos, apresentou um momento de melhor desempenho de classificação das amostras expansivas em detrimento das colapsíveis para a rede PE04, ao serem classificadas 65% das amostras (Figura 22). A partir dessa análise, verifica-se também que no momento de análise de 50% das amostras, 75% das amostras de solo expansivo já haviam sido identificadas, em detrimento das amostras de solo colapsível (22,2%), enquanto que na aleatoriedade haveria uma probabilidade de classificação de 50% para cada característica.

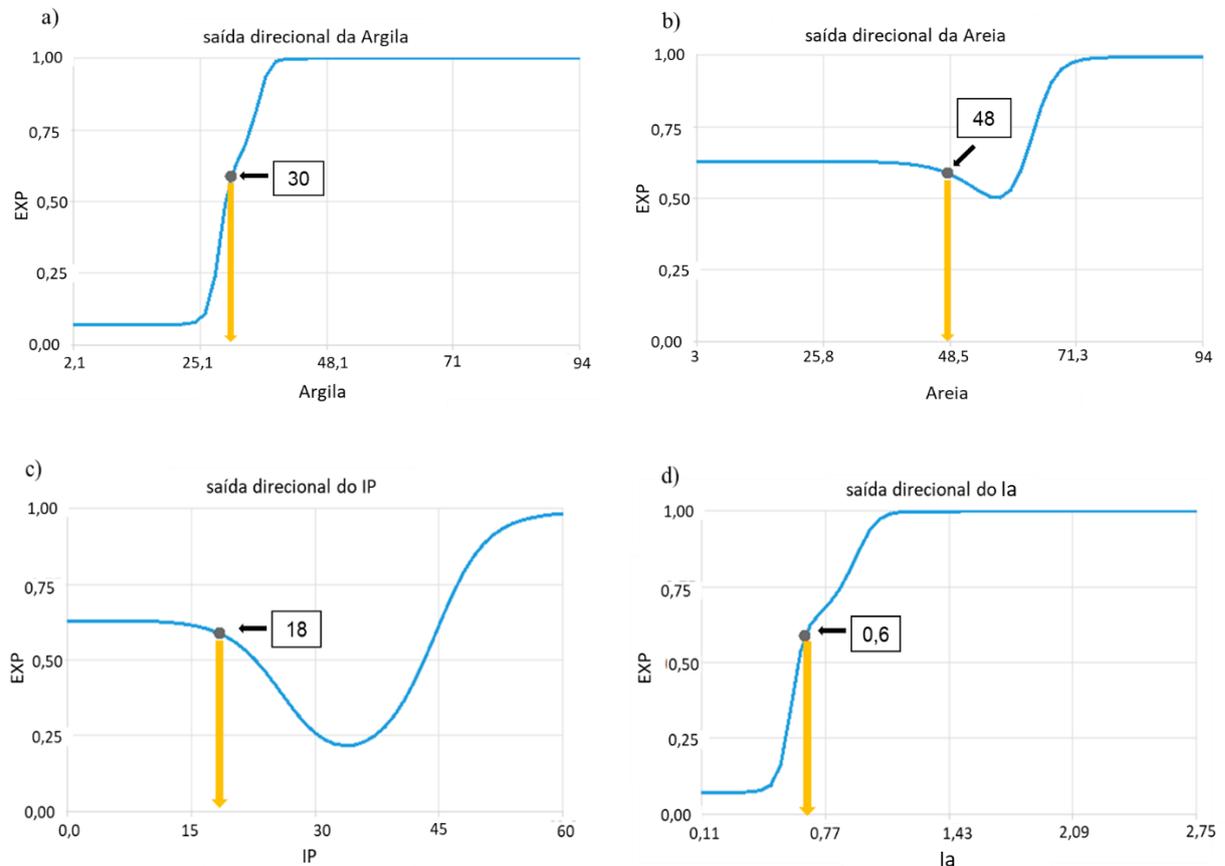
Figura 22 - Gráfico de Ganho Acumulado.



Fonte: A Autora (2022).

Depois que o modelo foi testado, ele está apto para fazer previsões. Na Figura 23 pode ser observado o ponto de referência dos valores críticos para cada variável de entrada (quando todos os limites combinados influenciam para a mudança de característica do solo), mediante as amostras do banco de dados usado para o desenvolvimento da rede PE04. Os pontos críticos, observados a partir do ponto de corte de classificação das amostras, indicam a mudança de comportamento do solo, demonstrando tendência a comportamento expansivo para valores de ‘% de Argila’ maiores que 30,78%, para o ‘% Areia’ menores que de 48,16%, para o ‘Índice de Plasticidade’ maiores que 18,47%, e finalmente para o ‘Índice de Atividade’ maiores que 0,66; a variável I_a tem valor crítico aproximado ao observado na literatura, em comparação com os limites definidos por Skempton (1953), as amostras ativas e normais apresentam limites quanto a expansão para valores acima de 0,75.

Figura 23 - Valores Críticos da Variável Argila (a), Areia (b), Índice de Plasticidade -IP (c) e Índice de Atividade - Ia (d) para as Amostras de Solos Expansivos



Fonte: A Autora (2022).

Atenta-se para a variável ‘% de Areia’ que teve uma correlação inversamente proporcional a característica expansiva (Figura 23-b), desse modo, amostras com valores de % de areia maiores que 48,16% têm tendência a característica colapsível, e menores que 48,16% poderá apresentar característica expansiva.

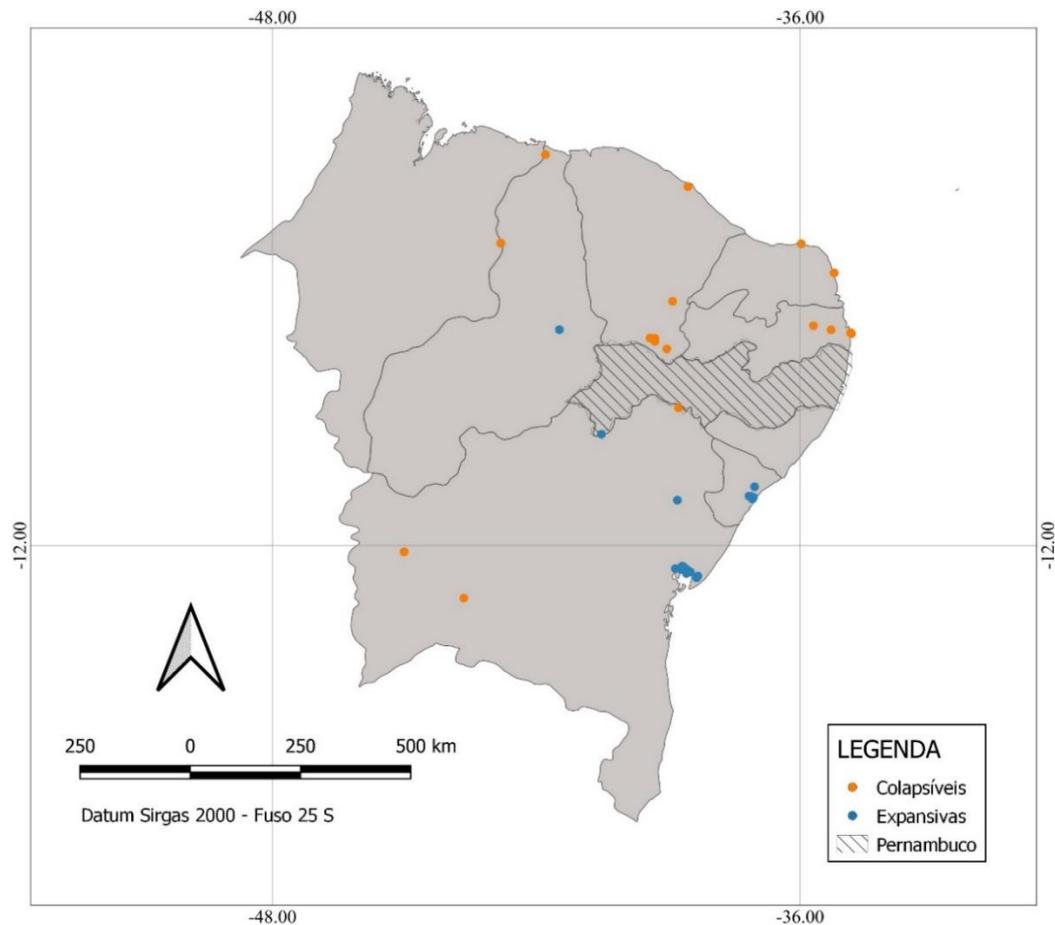
Após o entendimento dos padrões de comportamento dessas variáveis, segundo o banco de dados, pode-se analisar as características das amostras classificadas incorretamente. Nas 2 amostras de solo colapsível, classificadas como solo expansivo, há ausência de dados em uma delas (de 4 dados, 3 são ausentes) e na outra, embora apresente um Índice de Plasticidade baixo, possui um Índice de Atividade alto (1,80), inferindo uma maior tendência a característica expansiva. Para os solos expansivos, caracterizados como colapsíveis, uma mesma problemática é comum às classificações equivocadas: um percentual de areia que poderia indicar uma classificação de colapso, está situado em um ponto de inflexão na curva (Figura 23-b), ponto em que a tendência de comportamento expansivo, para solos com percentual de

areia entre 48 e 59%, nesse banco de dados, se ajustam, reduzindo de volume, possivelmente associado a liberação da pressão de sucção ao qual o solo estava submetido.

4.1.3 Aplicação dos Padrões da Rede PE04 para Solos dos Estados do Nordeste

A generalização dos padrões da rede PE04 foi aplicada a 45 amostras de solos dos estados do Nordeste (Figura 24), é importante destacar que os estados de Alagoas e Maranhão não tiveram amostras disponíveis a serem aplicadas nessa generalização, devido à falta de dados disponíveis para classificação pela RNA, tendo em vista que, para a correta classificação, é necessário que a maioria dos dados de entrada sejam definidos, como foi visto na etapa de desenvolvimento da rede, quanto maior a ausência de informação, mais difícil a correta classificação.

Figura 24 – Localização das 45 amostras de solos colapsíveis e expansivos do Nordeste (exceto Pernambuco, Alagoas e Maranhão) implementadas a rede PE04



Fonte: A Autora (2022).

Com os resultados obtidos pela simulação, foi calculada a Matriz de Confusão e analisadas as Métricas de Avaliação Binária (Tabelas 4-a e 4-b) classificando 4 amostras incorretamente, como falsas expansivas e 41 amostras corretamente, alcançando uma taxa de acurácia de 91,11%. Na investigação detalhada do banco de dados, observou-se que as amostras classificadas incorretamente não continham duas das quatro variáveis solicitadas pela entrada da rede, ratificando a ideia de que a ausência de informações de entrada compromete a capacidade de previsão.

É observado também uma importante característica do modelo a exportabilidade. A rede PE04, treinada em Pernambuco, pôde ser aplicada para amostras de outros estados, classificando corretamente com uma boa margem de precisão e baixa taxa de erro (8,89%), identificando com excelente precisão os solos expansivos, confirmada no alto percentual de Sensibilidade (100%), conforme Tabela 4.

Tabela 4 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b)

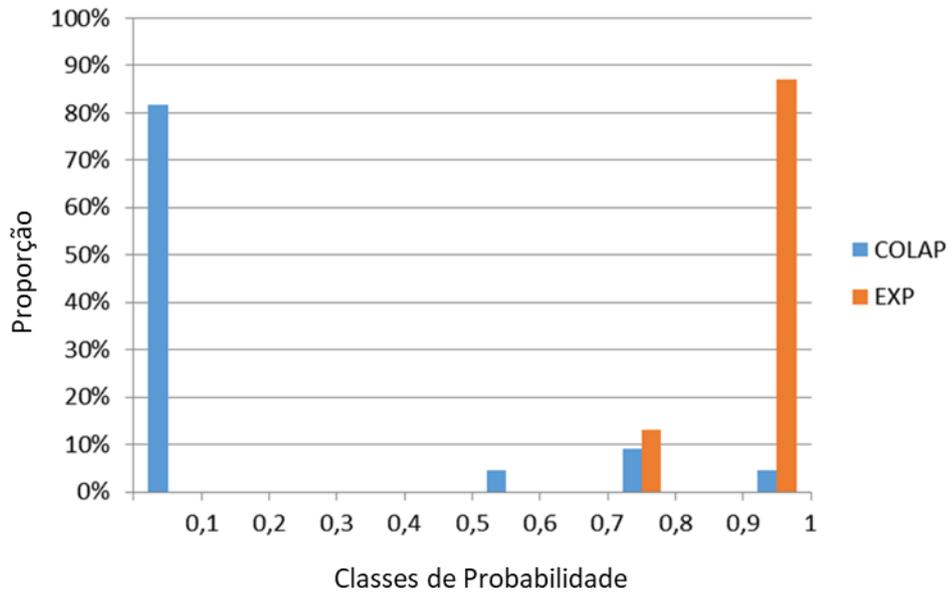
		a)		b)	
		Verdadeiras Expansivas	Falsas Expansivas		
a)	Verdadeiras Expansivas	22	4	Acurácia	91,11%
	Falsas Colapsíveis	0	19	Taxa de Erro	8,89%
				Precisão	85,18%
				Sensibilidade	100%
				Especificidade	81,82%

Fonte: A Autora (2022).

Na Figura 25, é apresentado o Histograma de População, sendo possível identificar as características de classificação feitas pela previsão da RNA em uma escala, que vai de 0,00 a 0,50 para características de colapso e de 0,51 a 1,00 para características de expansão. Além de identificar a proporção de amostras classificada erroneamente, quanto se tem sobreposição.

Com a aplicação dos parâmetros, a rede apresentou um comportamento melhor que o observado na etapa de testes. Com um limiar de separação das características em 0,5, observou-se um importante afastamento entre as classes de amostras. As amostras expansivas foram classificadas com suscetibilidade entre os intervalos de 0,7-0,8 e 0,9-1,0. Enquanto mais de 80% das amostras colapsíveis concentrou-se no intervalo 0,0-0,1.

Figura 25 – Histograma de População



Fonte: A Autora (2022).

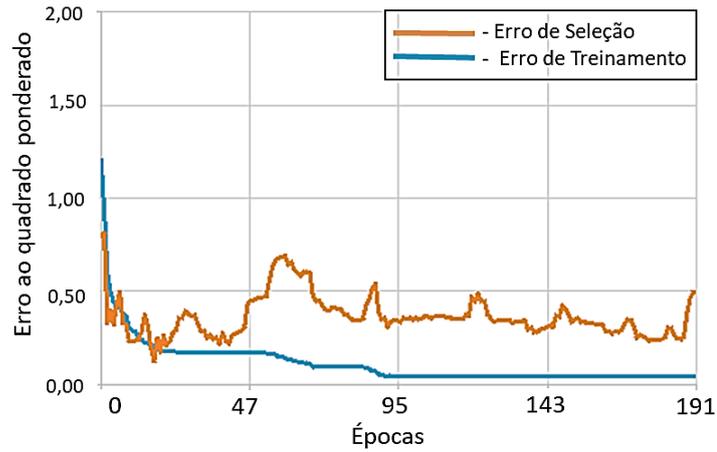
Assim, a RNA desenvolvida com amostras de Pernambuco e com dados geotécnicos como variáveis de entrada, continuou sendo representativa às características do Nordeste, tendo em vista a acurácia de mais de 91%; isso se deve a fatores como: variabilidade possível alcançada pelo banco de dados da etapa de desenvolvimento e possivelmente pela repetição nas características das amostras utilizadas na etapa de validação.

4.2 Rede Neural PE07

4.2.1 Composição da RNA PE07

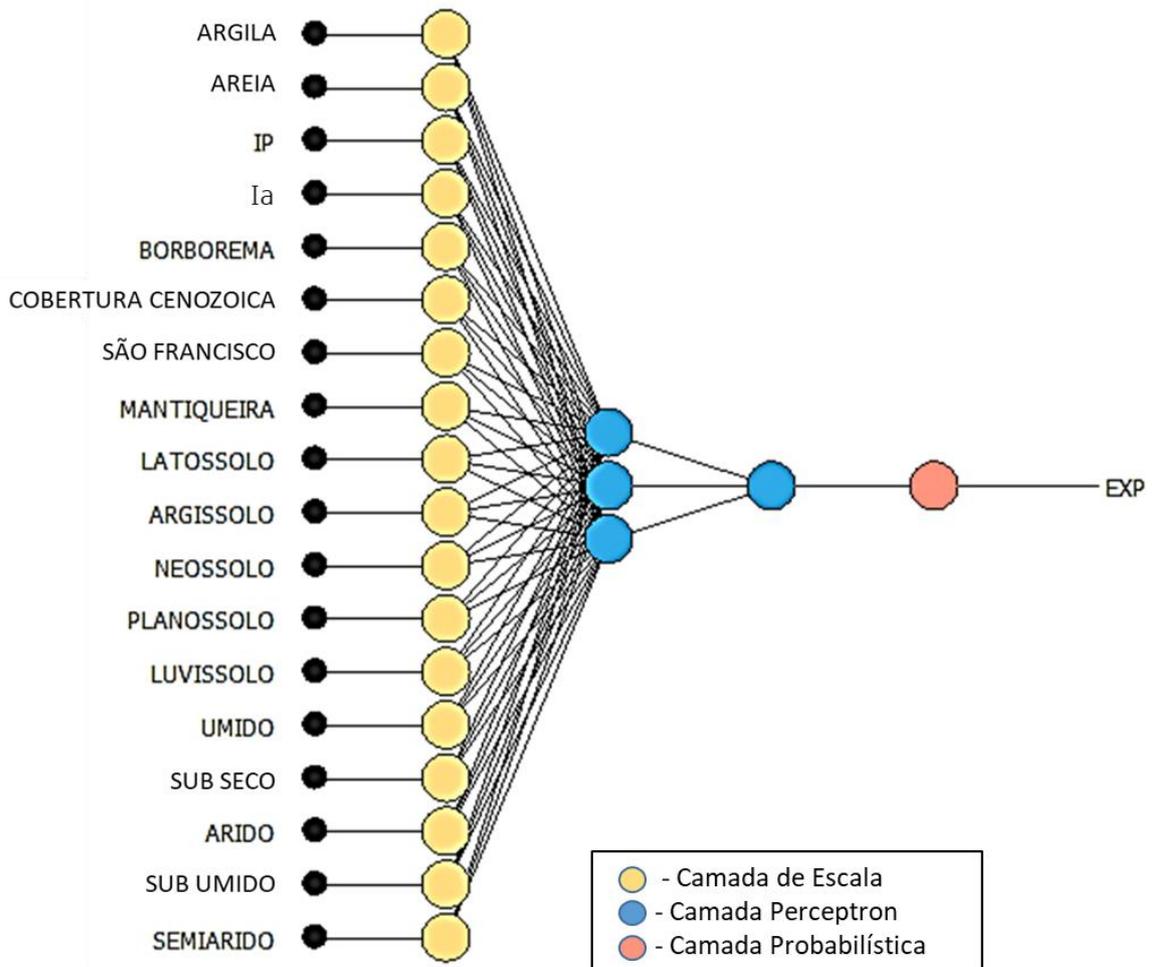
Os erros de treinamento e seleção em cada iteração são detalhados na Figura 26. O valor inicial do erro de treinamento é 1,2 e o valor final após 191 épocas é 0,04. O valor inicial do erro de seleção é 0,8 e o valor final após 191 épocas é 0,5. A adição das variáveis categóricas ocasionou em um maior número de iterações para que a rede PE07 alcançasse os mínimos valores de erro possíveis. Essa característica já era previsível, pois quanto mais simples a rede neural (menor quantidade de variáveis de entrada) menor a quantidade de iterações a ser executada pela rede. Já o parâmetro de arquitetura, definido para esta, não difere dos parâmetros da rede PE04, contando com duas camadas *perceptron* e neurônios das camadas ocultas ativados pela função logística (Figura 27).

Figura 26 - Histórico dos erros do método Quase-Newton



Fonte: A Autora (2022).

Figura 27 – Representação da melhor arquitetura da rede neural



Fonte: A Autora (2022).

4.2.2 Teste – Análise do Desempenho

As 17 amostras destinadas a etapa de testes, de acordo com a fração de 19,5% do banco de dados, apresentaram uma melhoria na classificação da rede PE07 após a inserção das variáveis categóricas, 15 amostras foram classificadas corretamente (88,2%) e 2 amostras classificadas incorretamente (11,8%), Tabela 5-a.

Observando com maior detalhe os falsos colapsíveis, casos em que o solo expansivo não é identificado pela RNA, tem-se uma boa avaliação, tendo em vista que não ocorreu esse tipo de erro. Para as 2 amostras classificadas incorretamente, falsos expansivos, os parâmetros, da primeira delas não contam com a granulometria da Areia e Argila e os parâmetros da segunda amostra tem como ausências o percentual de Argila, Índice de Plasticidade e Índice de Atividade. Dessa forma, entende-se que embora a rede possua bons ajustes de parâmetros para identificação do comportamento de colapso ou expansão, a ausência de algumas variáveis compromete significativamente o resultado da classificação.

Os dados derivados da matriz de confusão aplicados ao modelo classificador são utilizados para analisar os critérios de desempenho na Tabela 5-b. A partir desses parâmetros de precisão da classificação, taxa de erro, sensibilidade e especificidade, pode-se ver que o modelo está classificando com excelência os casos de solos expansivos (sensibilidade = 100%), assim, nenhum solo com tal característica foi previsto como solo colapsível. A rede classificou com boa precisão os solos colapsíveis (especificidade = 81,8 %), identificando apenas duas amostras de solos expansivos (com ausência de dados geotécnicos) como colapsíveis.

A inclusão das variáveis categóricas influenciou diretamente na classificação das amostras expansivas, que tiveram todas suas amostras reconhecidas pela classificação da rede, enquanto as amostras colapsíveis mantiveram o número de erros já observados no desenvolvimento da rede PE04 (2 amostras classificadas erroneamente).

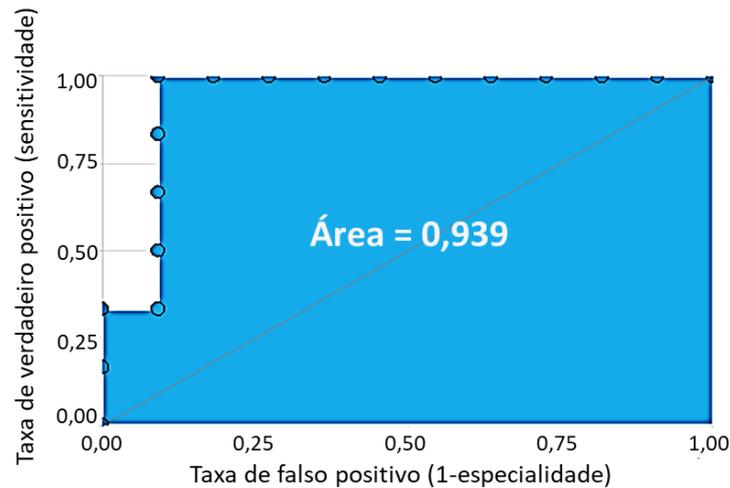
Tabela 5 – Matriz de confusão (a) e Métricas da Avaliação Binária (b)

a)		b)	
9 (Verdadeiras Expansivas)	2 (Falsas Expansivas)	Acurácia	88,2%
0 (Falsas Colapsíveis)	6 (Verdadeiras Colapsíveis)	Taxa de Erro	11,8%
		Sensibilidade	100%
		Especificidade	81,8%

Fonte: A Autora (2022).

A melhoria de classificação também refletiu na área do gráfico ROC (Figura 28), com um aumento da área sob a curva passando para 0,939, indicando como melhor ponto de corte o limiar de decisão em 0,48.

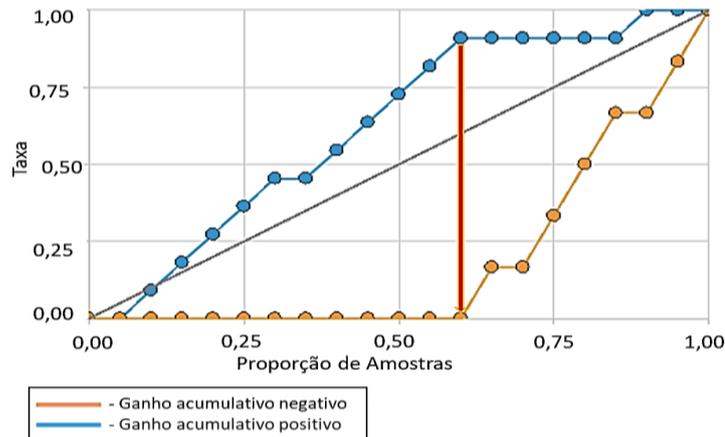
Figura 28 – Gráfico ROC



Fonte: A Autora (2022).

O gráfico de ganho acumulado, Figura 29, também refletiu a melhoria de classificação, antecipando o momento de melhor desempenho com a mais rápida classificação das amostras expansivas em detrimento das colapsíveis, de 65% (caso da rede PE04) para 60% na classificação da rede PE07. Nesse momento ótimo 91% das amostras expansivas foram classificadas. A melhoria no resultado desse teste, em comparação com a rede PE04, era prevista, devido ao aumento na taxa de sensibilidade (identificação da característica expansiva corretamente). Em comparação a uma situação de classificação aleatória, na qual ao ser analisadas 50% das amostras disponíveis para essa etapa, haveria 50% de chance de identificar tanto uma classe como outra; a classificação a partir do treinamento da rede PE07 aumenta de 50 para 75% a taxa de classificação das amostras expansivas, em detrimento das amostras colapsíveis.

Figura 29 – Gráfico de Ganho Acumulado.

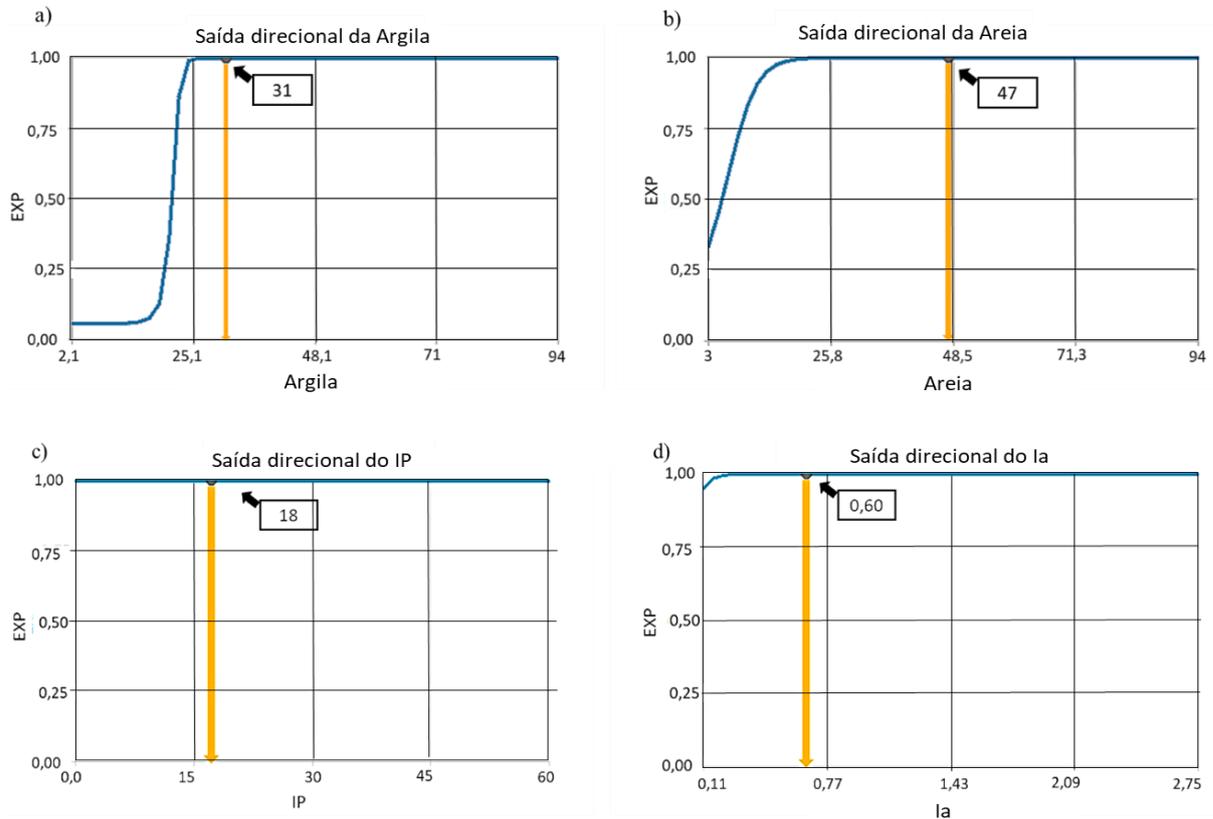


Fonte: A Autora (2022).

A Figura 30, mostra o ponto de referência para a mudança de característica do solo, identificando os valores críticos que melhor representam as variáveis de entrada, mediante a mudança de comportamento do solo quanto a probabilidade de colapso ou expansão para as amostras do banco de dados utilizada. Como o banco de dados para o desenvolvimento das redes PE04 e PE07 não mudou, tem-se valores muito semelhantes para as variáveis geotécnicas. Já as variáveis qualitativas não integram esse formato de representação e classificação, pois cada uma delas tem suas características que influenciam de forma particular, com suas peculiaridades, nas condições de propiciar as características de colapso ou expansão aos solos.

No caso da variável ‘% de Argila’ o ponto crítico para caracterizar a expansividade ocorre em amostras com valores maiores que 31,31%; para o Índice de Plasticidade, percentuais maiores que 17,25%, e para o Índice de Atividade valores maiores que 0,66. Já para o ‘% de Areia’, que manteve um comportamento inversamente proporcional, a maior tendência de classificação da amostra como expansiva é com percentuais menores que 47,66% e consequentemente como colapsíveis com percentuais maiores que isso.

Figura 30 – Valores Críticos da Variável Argila (a), Areia (b), IP (c) e Ia (d) para as Amostras de Solos Expansivos



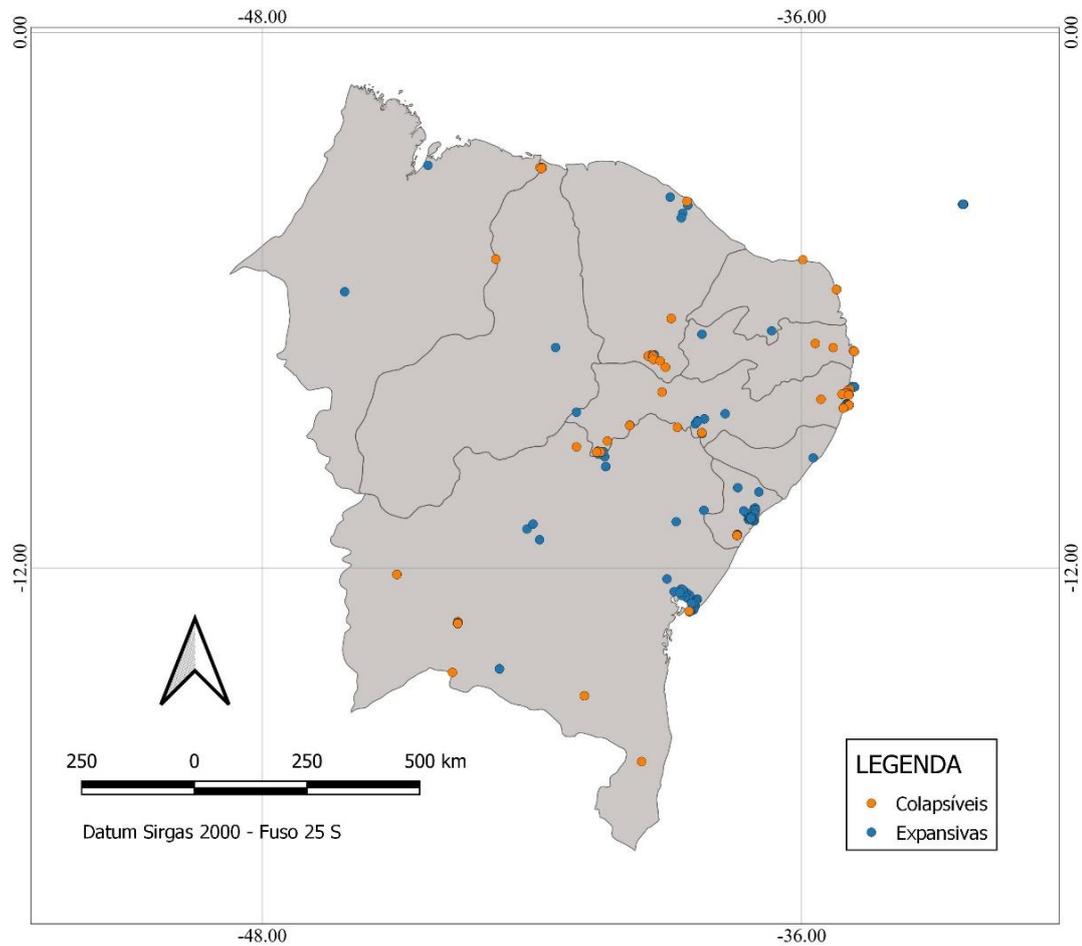
Fonte: A Autora (2022).

Após o desenvolvimento dessa rede, o banco de dados sofreu um considerável avanço em suas atualizações, com isso, a aplicação da rede PE07 foi feita a 284 amostras do Nordeste do Brasil e em um segundo momento a 463 amostras de todo o território do Brasil, a fim de que pudesse ser comparada sua exportabilidade, para essas duas dimensões.

4.2.3 Aplicação dos Padrões da Rede PE07 para Solos dos Estados do Nordeste

Os padrões da rede PE07 foram aplicados à 284 amostras dos demais estados do Nordeste, incluindo novas amostras do estado de Pernambuco e o arquipélago de Fernando de Noronha (Figura 31). Desse total, 159 amostras eram de solo expansivo e 126 amostras de solo colapsível; na matriz de confusão (Tabela 6-a), pôde-se observar que 52 amostras foram classificadas como falsas expansivas, enquanto 46 amostras de solo foram identificadas como falsas colapsíveis, resultando em métricas da matriz de confusão (Tabela 6-b) com taxa de erro de 34,51% e uma precisão de 68,29%.

Figura 31 – Localização das 284 amostras do Nordeste brasileiro validadas às cegas com a rede PE07



Fonte: A Autora (2022).

Tabela 6 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b)

a)

112 (Verdadeiras Expansivas)	52 (Falsas Expansivas)
46 (Falsas Colapsíveis)	74 (Verdadeiras Colapsíveis)

b)

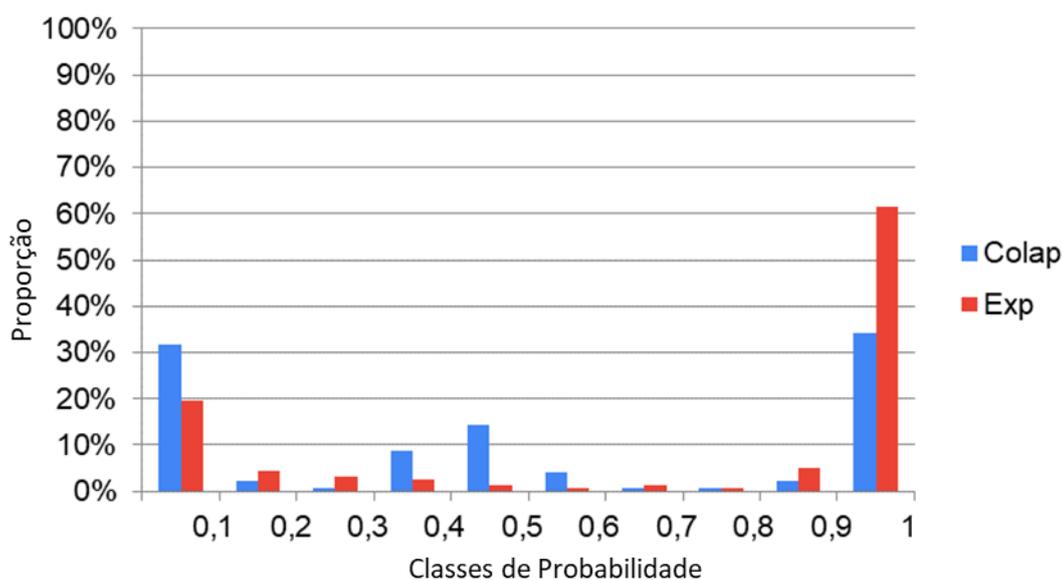
Acurácia	65,49%
Taxa de Erro	34,51%
Precisão	68,29%
Sensitividade	70,89%
Especificidade	58,73%

Fonte: A Autora (2022).

No Histograma de Populações (Figura 32) é possível observar as classificações feitas por essa rede, que além de ter um desempenho consideravelmente baixo, comparado a rede PE04, por ter a classificação generalizada a padrões de pedologia e geologia desconhecidos na etapa de treinamento da rede, classificou amostras na zona central do gráfico (próximas ao limiar de decisão, 0,5). Essa área de classificação representa amostras classificadas com um menor índice de certeza, tanto para colapso, como para expansão. Isso pode ocorrer tanto pelo

desconhecimento de características presentes nessas novas amostras, como por ocorrência de padrões com alta/média suscetibilidade ao colapso e expansão, comuns às duas classificações, como por exemplo, as classificações pedológicas: Chernossolo e Neossolo, os climas: árido e semiárido, e as classes geológicas: Recôncavo – Tucano – Jatobá e Costeira e Margem Continental.

Figura 32 – Histograma de População



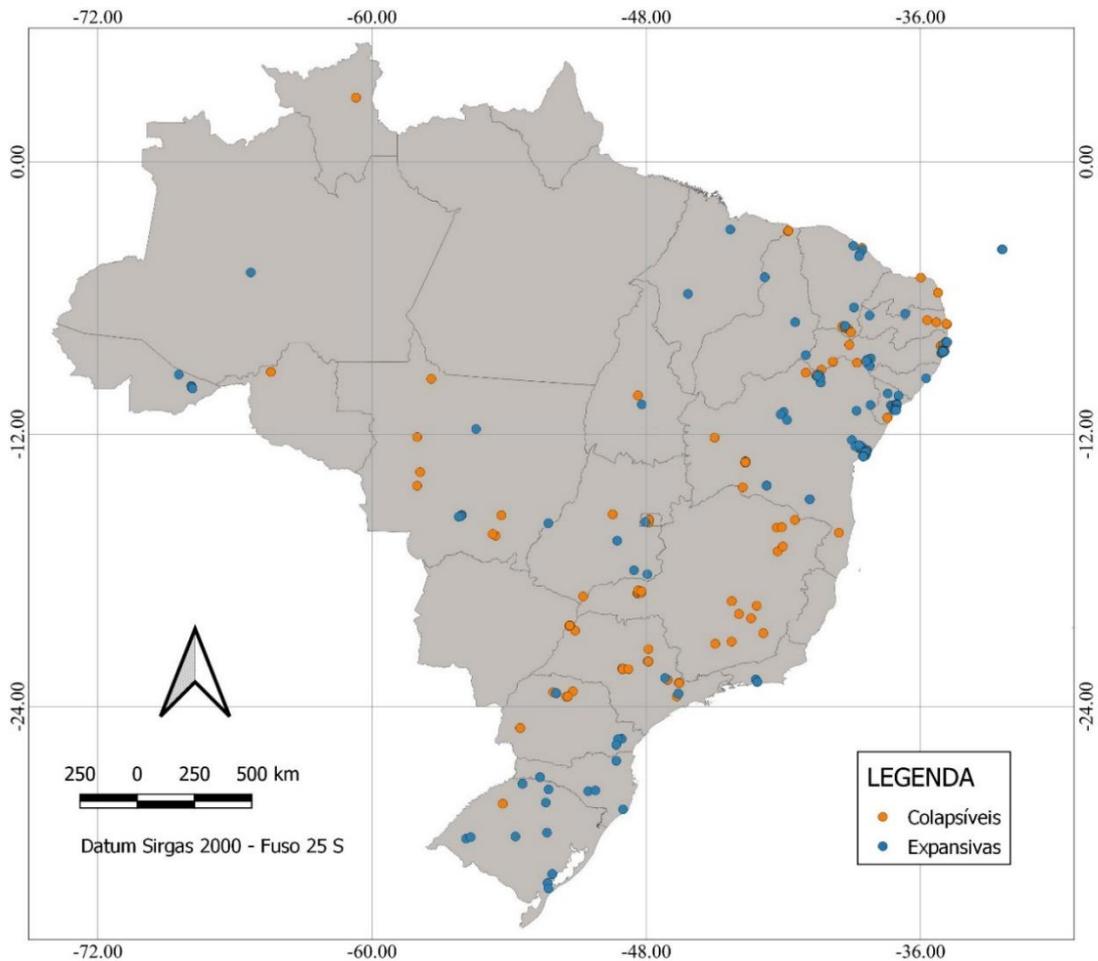
Fonte: A Autora (2022).

A generalização dos padrões às amostras do Nordeste classificou características de geologia e pedologia desconhecidas pela rede na fase de desenvolvimento, como por exemplo geologia: Amazonas-Solimões, Costeira e Margem Continental e Tocantins; e classe pedológica: Gleissolo.

4.2.4 Aplicação dos Padrões da Rede PE07 para Solos do Brasil

Para essa validação, foram consideradas 463 amostras pertencentes as 5 regiões do país (Figura 33), apenas os estados do Amapá, Mato Grosso do Sul e Pará não foram localizados registros de pesquisas desenvolvidas nessa área, e por isso não integraram o banco de dados. É observada uma grande concentração de amostras na região Nordeste, seguida da região Sudeste; onde se concentram os principais grupos de pesquisa de solos não saturados, reunindo os maiores bancos de dados nessa área.

Figura 33 – Localização das 463 amostras do Brasil validadas às cegas com a rede PE07



Fonte: A Autora (2022).

Os resultados dessa generalização de padrões destacam uma insuficiência de conhecimento da rede PE07 para os padrões das amostras validadas, tendo em vista que a rede PE07 foi treinada apenas com amostras do estado de Pernambuco. Padrões qualitativos e limites das variáveis quantitativas, desconhecidos anteriormente, refletiram no resultado dessa generalização de padrões com 238 amostras colapsíveis e 225 amostras expansivas, que resultaram em uma matriz de confusão (Tabela 7-a) com 128 amostra classificadas como falsas expansivas e 68 falsas colapsíveis, assim, a rede passou a classificar, segundo as métricas da avaliação binária (Tabela 7-b) com uma precisão de 55,09% e uma taxa de erro de 42,33%, um valor muito próximo da classificação por aleatoriedade.

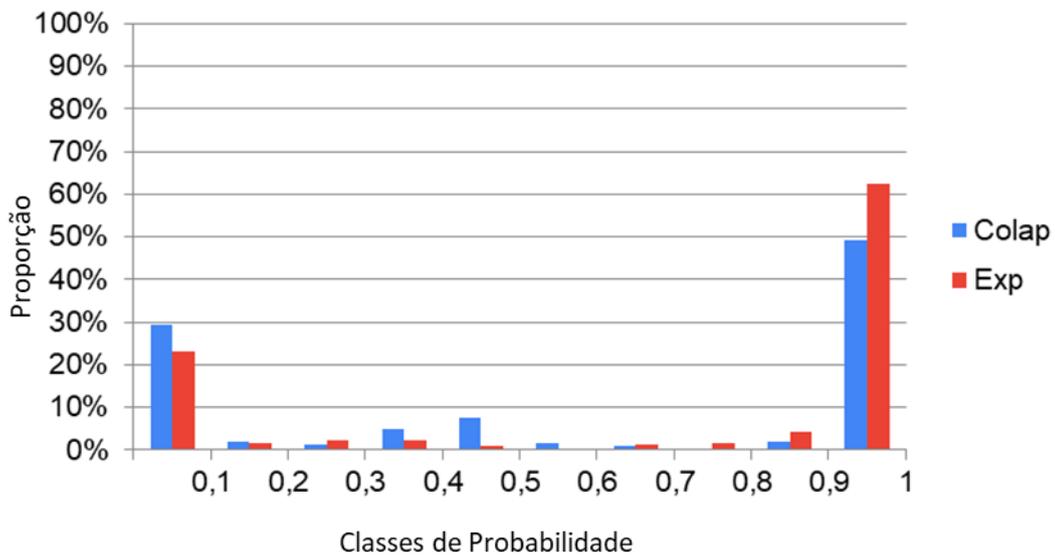
Tabela 7 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b)

a)		b)	
157 (Verdadeiras Expansivas)	128 (Falsas Expansivas)	Acurácia	56,67%
68 (Falsas Colapsíveis)	110 (Verdadeiras Colapsíveis)	Taxa de Erro	42,33%
		Precisão	55,09%
		Sensitividade	69,78%
		Especificidade	46,22%

Fonte: A Autora (2022).

No Histograma de População (Figura 34) é possível observar que 15% das amostras expansivas e 20% das amostras colapsíveis concentram-se dispersas nas áreas de menor certeza de classificação, entre 0,20 e 0,8, enquanto o restante das amostras se dividiu em classificações corretas e errôneas. Para o caso das amostras colapsíveis apenas 30% das amostras foram classificadas com alta probabilidade de suscetibilidade; para as amostras expansivas esse percentual foi acima de 60%.

Figura 34 – Histograma de População



Fonte: A Autora (2022).

Como era previsto para esta generalização de padrões, mais características de variáveis de entrada são desconhecidas pela rede treinada com amostras de Pernambuco, como por exemplo clima: super úmido; geologia: Amazonas-Solimões, Amazonia, Costeira e Margem Continental, Mantiqueira, Paraná, Parnaíba e Tocantins; e pedologia: Cambissolo, Chernossolo, Dunas, Espodossolo, Gleissolo, Nitossolo, Organossolo, Plintossolo e Vertissolo.

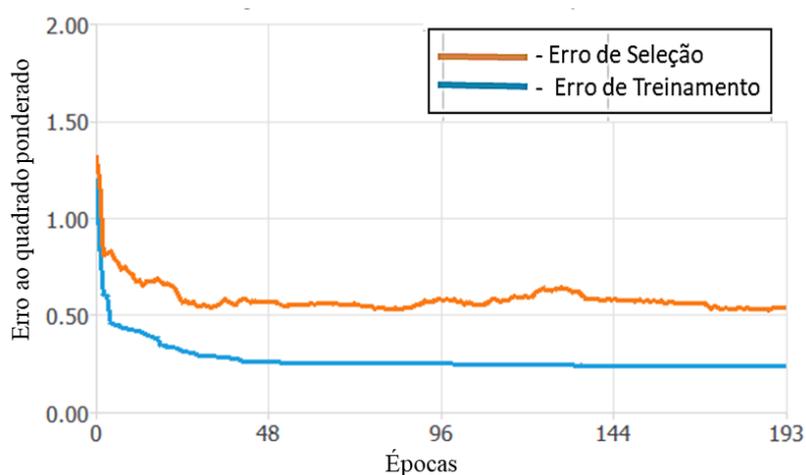
Assim, entende-se que a generalização dos padrões de uma rede desenvolvida com dados de Pernambuco, tendo como de entrada variáveis geotécnicas e de origem e formação do solo, não atende a critérios de classificação quando aplicada a solos de quaisquer regiões do Brasil, tendo uma acurácia de classificação próxima a uma classificação aleatória, e isso ocorre pelo fato de algumas características de clima, pedologia e geologia das demais regiões do Brasil não terem representação no estado de Pernambuco, estando assim ausentes no processo de treinamento da rede.

4.3 Rede Neural BR03

4.3.1 Composição da RNA BR03

O histórico dos erros de treinamento e seleção, pelo método Quase-Newton em cada uma das 193 épocas/iterações, é detalhado na Figura, 35. Iniciando com valores de 1,2 para erro de treinamento e 1,31 para erro de seleção e finalizando as iterações com erros de 0,2 e 0,5, respectivamente.

Figura 35 – Histórico dos erros do método Quase-Newton

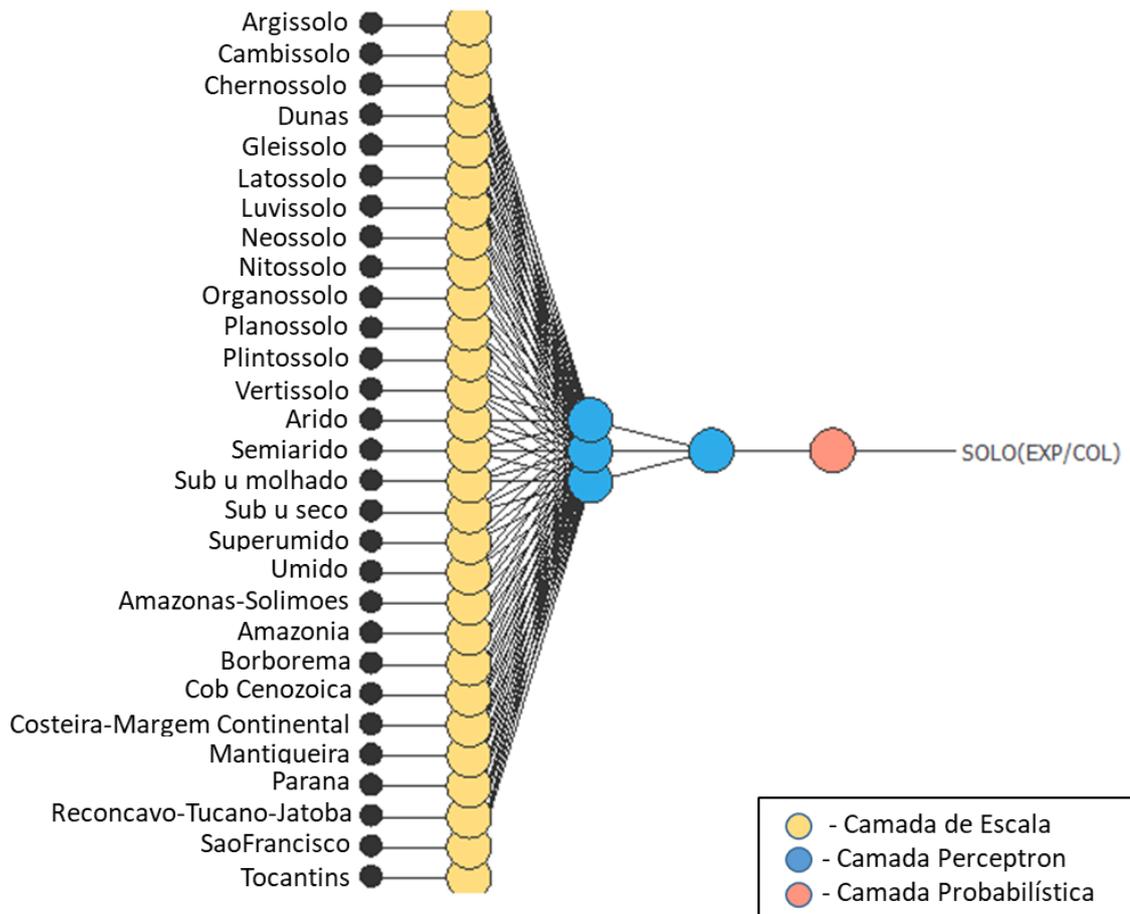


Fonte: A Autora (2022).

A representação gráfica da arquitetura da rede, Figura 36, conta com duas camadas *perceptron* e cada uma delas com neurônios ativados pela função logística. Para essa rede, com três variáveis de entrada de caráter categórico, cada unidade de mapeamento dessas variáveis é representada como uma entrada individual no tipo de rede e no tipo de equação gerada,

inferindo ou não valores a amostra que está sendo classificada, por isso há uma entrada específica para cada variável de clima, geologia e pedologia, assim a arquitetura é composta de 29 entradas diferentes.

Figura 36 - Representação gráfica da melhor arquitetura da rede neural.



Fonte: A Autora (2022).

4.3.2 Teste – Análise do Desempenho

Das 78 amostras destinadas a etapa de testes (o equivalente a 19,8% das amostras destinadas ao desenvolvimento dessa rede), foi observado na matriz de confusão (Tabela 8-a) que 70 amostras foram classificadas corretamente como verdadeiras expansivas (50%) e verdadeiras colapsíveis (39,7%) e 8 amostras foram classificadas incorretamente, 4 como falsas expansivas (5,13%) e outras 4 como falsas colapsíveis (5,13%).

Os parâmetros derivados da matriz de confusão (Tabela 8-b) destacam a baixa taxa de erro de classificação (10,3%) complementar a alta acurácia da rede (89,7%) e com um alto reconhecimento das características de solos expansivos e colapsíveis, com taxa de sensibilidade de 90,7% e de especificidade em 88,6%.

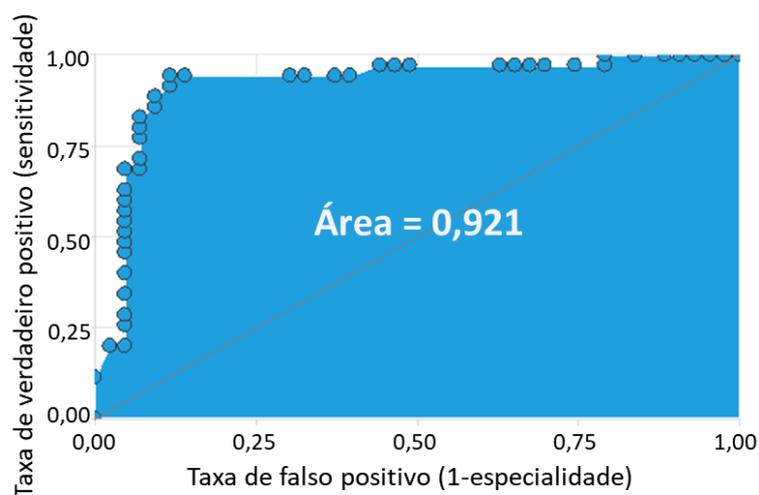
Tabela 8 – Matriz de confusão (a) e Métricas da Avaliação Binária (b).

a)		b)	
39 (Verdadeiras Expansivas)	4 (Falsas Expansivas)	Acurácia	89,7%
4 (Falsas Colapsíveis)	31 (Verdadeiras Colapsíveis)	Taxa de Erro	10,3%
		Sensibilidade	90,7%
		Especificidade	88,6%

Fonte: A Autora (2022).

A partir desse resultado de sensibilidade é observada uma curva no gráfico ROC (Figura 37) que se aproxima muito do ponto (0, 1), elevando assim a área desse gráfico a 0,921, a melhor área atingida entre todas as redes comparadas nessa pesquisa, e tendo como limite ótimo de decisão o ponto 0,832, de modo a se concluir que nessa etapa de testes as amostras têm suas características de classificação (colapsíveis e expansivas) bastante sobrepostas.

Figura 37 – Gráfico ROC.

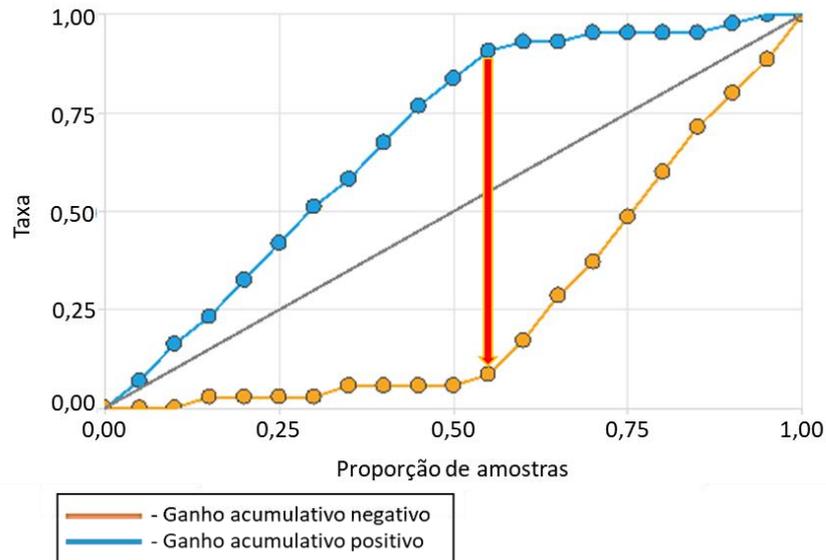


Fonte: A Autora (2022).

Por fim, o gráfico de ganho acumulado (Figura 38), como já era esperado pela observação do comportamento da curva ROC (Figura 37), teve aproximadamente 90% das amostras com

características expansivas classificadas inicialmente; só depois foram classificadas as amostras colapsíveis, esse momento de melhor desempenho na classificação das amostras expansivas é no momento de classificação de 55% das amostras.

Figura 38 – Gráfico de Ganho Acumulado.

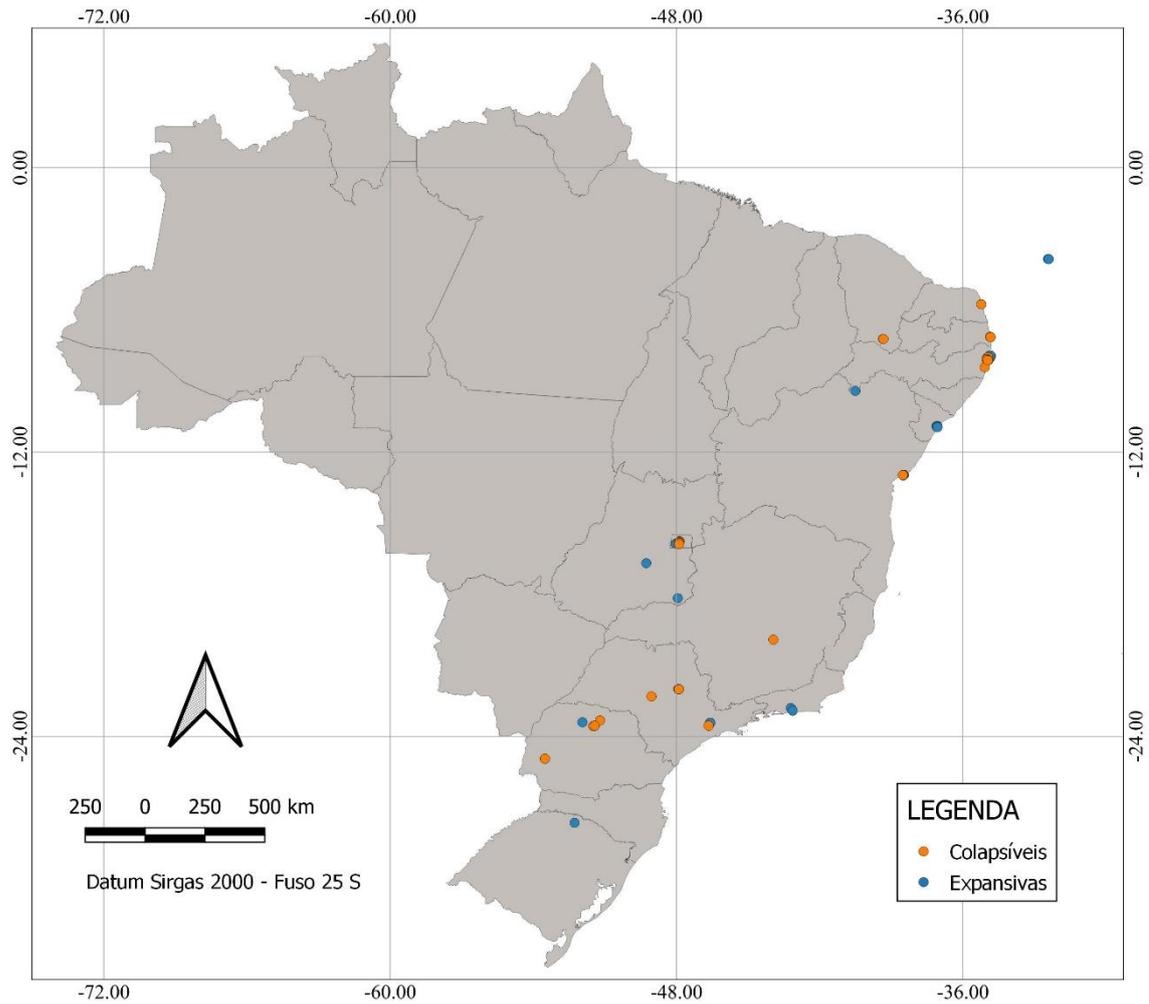


Fonte: A Autora (2022).

4.3.3 Aplicação dos Padrões da Rede BR03 para Solos do Brasil

Com o banco de dados já definido, no desenvolvimento da rede BR03, foram separadas as amostras com algum dado de característica ausente e essas foram destinadas a etapa de validação, totalizando 133 amostras (Figura 39). Preferiu-se usar amostras com todas as características presentes na etapa de desenvolvimento da rede para que a qualidade do aprendizado das características fosse a maior possível.

Figura 39 – Localização das 133 amostras do Brasil implementadas a rede BR03



Fonte: A Autora (2022).

A aplicação da rede BR03 resultou em uma matriz de confusão (Tabela 9–a) com 13 amostras classificadas como falsas expansivas e 11 amostras classificadas como falsas colapsíveis, resultando em uma taxa de erro de 18,05%. Do total de amostras validadas (133 amostras), 130 situavam-se em localização com ausência de classificação pedológica, áreas intensamente urbanizadas, e 3 e áreas sem classificação climatológica (localizadas no arquipélago de Fernando de Noronha), ainda assim, a rede BR03 apresentou como métricas resultantes da avaliação binária (Tabela 9-b) uma precisão de 82,89%, para amostras expansivas uma sensibilidade de 85,14% e para amostras colapsíveis uma especificidade de 77,97%.

Tabela 9 – Matriz de Confusão (a) e Métricas da Avaliação Binária (b)

a)

63 (Verdadeiras Expansivas)	13 (Falsas Expansivas)
11 (Falsas Colapsíveis)	46 (Verdadeiras Colapsíveis)

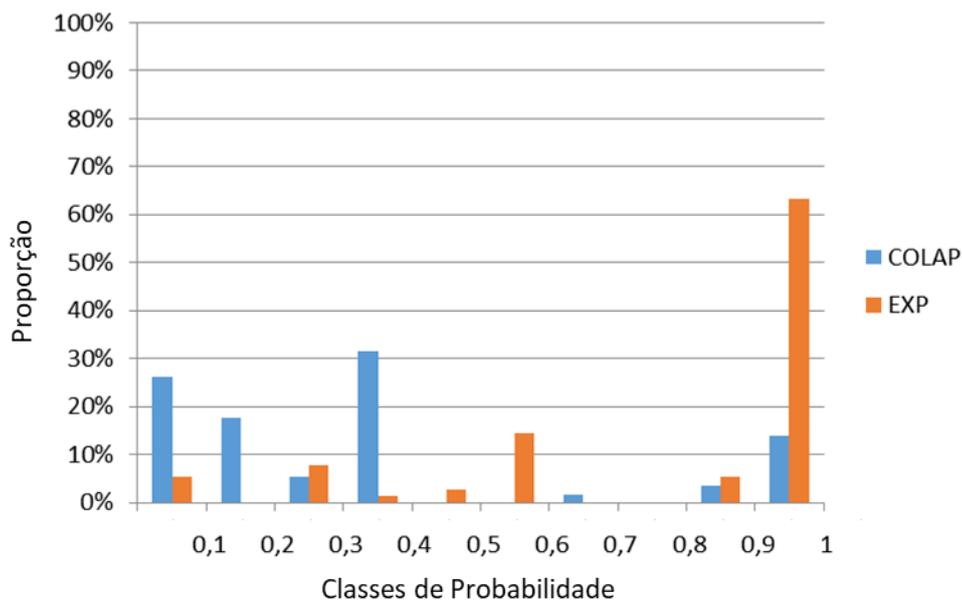
b)

Acurácia	81,95%
Taxa de Erro	18,05%
Precisão	82,89%
Sensitividade	85,14%
Especificidade	77,97%

Fonte: A Autora (2022).

O histograma de população (Figura 40) apresentou uma característica de sobreposição dos resultados de previsão de comportamento das amostras e destaca uma maior incerteza de resultados principalmente para solos colapsíveis, que tiveram mais de 30% das amostras, embora classificadas corretamente como colapsíveis, com baixa suscetibilidade ao colapso, percentual muito próximo ao limite de dúvida atribuído a classificação da rede neural (entre 0,30 e 0,50); 28% das amostras colapsíveis, com uma probabilidade maior que 90% de apresentarem comportamento colapsível, suscetibilidade alta ao colapso. Enquanto as amostras expansivas, mais de 60% obtiveram probabilidade maior que 90% de indicação de comportamento expansivo, suscetibilidade alta a expansão.

Figura 40 – Histograma de População



Fonte: A Autora (2022).

A síntese dos desempenhos das RNA detalhadas nessa pesquisa pode ser comparada a partir das informações descritas na Tabela 10 e Figura 41. Os índices analisados não necessariamente devem ter crescimento para representar um desempenho positivo, o caso da taxa de erro e o percentual de Ganho Acumulado são positivos para o desempenho da rede quando decrescem.

Na etapa de testes (Figura 41-a), os critérios de avaliação binária apresentam um crescimento da taxa de acurácia e na taxa de especificidade (capacidade de reconhecer a característica de colapsividade) e um importante decréscimo na taxa de erro, pontos positivos para a evolução das redes. Entretanto, há um decréscimo na taxa de sensibilidade (capacidade de reconhecer a característica de expansividade) entre as redes PE07 e BR03; ponderou-se, nesse caso, as características qualitativas de cada uma das redes, uma testada com amostras de Pernambuco, outra com amostras do Brasil, a variabilidade de características, das variáveis de entrada geológicas e pedológicas, a serem reconhecidas pela rede BR03 foi consideravelmente maior com algumas possivelmente não representadas na etapa de treinamento, ficando o reconhecimento para a etapa de testes e validação às cegas.

A etapa de validação às cegas (Figura 41-b) deve ter seus padrões de desempenho analisados caso a caso, pois a característica preponderante nessa situação é o domínio ao qual está sendo aplicado os padrões de generalização, se esse for o mesmo aplicado nas etapas de treinamento, seleção e teste da RNA o desempenho deve ser melhor que uma validação aplicada a um domínio que extrapola o padrão de características ao qual a rede foi desenvolvida. Dessa forma, observa-se que as redes PE04 e PE07 tiveram seus padrões de classificação aplicados a domínios extrapolados em maior e menor dimensão. A extrapolação da rede PE04 às amostras do Nordeste apresentou avanços no desempenho diminuindo erros (de 23,53% para 8,89%) e aumentando taxas de acurácia (de 76,47% para 91,11%), sensibilidade (de 75% para 100%) e especificidade (de 77,78% para 81,82%), inferindo às amostras validadas às cegas um padrão de características geotécnicas pouco variável em relação aos reconhecidos no desenvolvimento da RNA.

Tabela 10 - Síntese de Desempenho das RNA

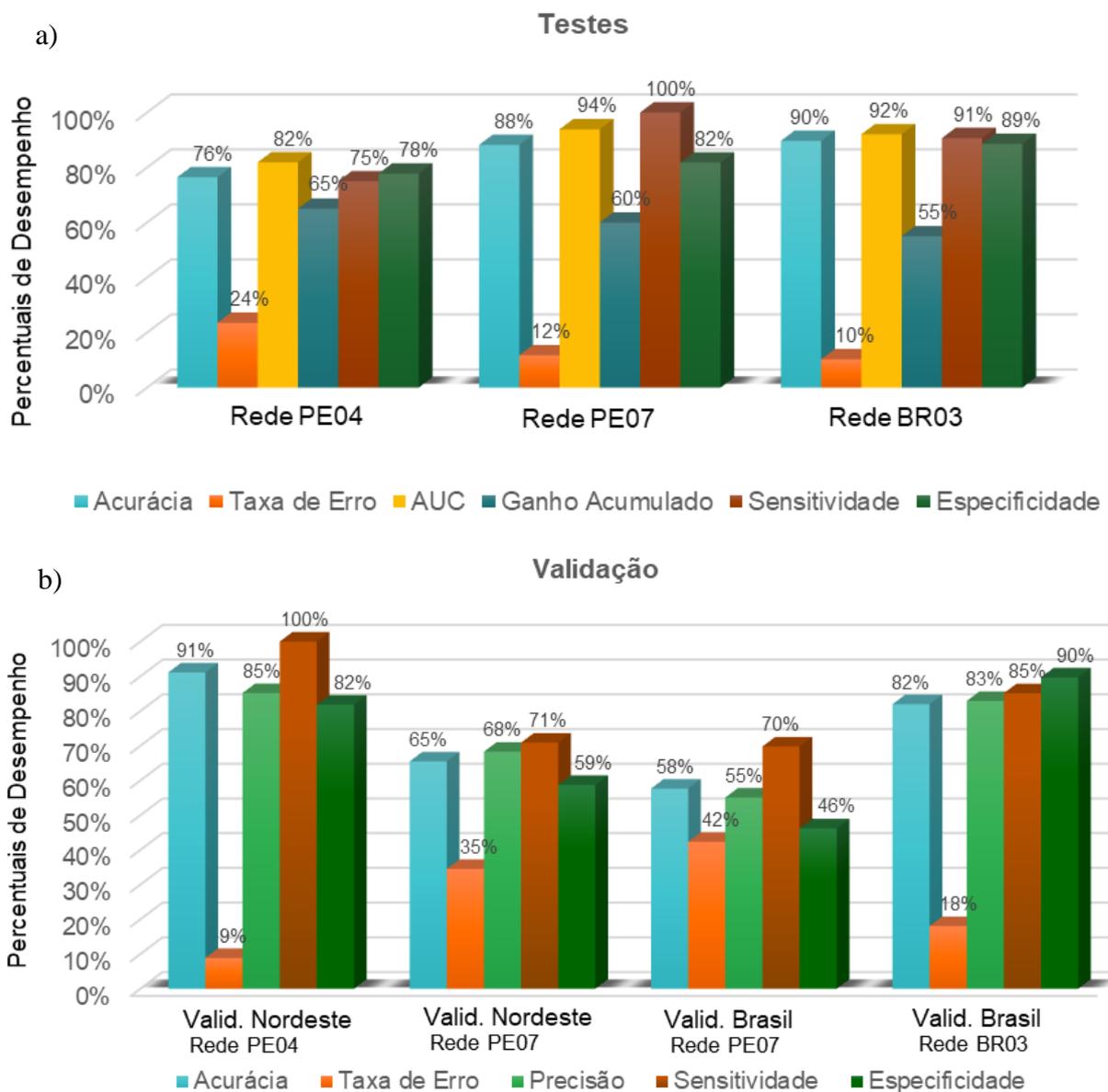
Redes/Métricas de Desempenho		Rede PE04	Rede PE07		Rede BR03
Testes	Acurácia	76,47%	88,23%		89,7%
	Taxa de Erro	23,53%	11,76%		10,3%
	Sensitividade	75%	100%		90,7%
	Especificidade	77,78%	81,82%		88,6%
	AUC	0,819	0,939		0,921
	Ganho Acumulado	Em 65%	Em 60%		Em 55%
Validação	Acurácia	91,11%	65,49%	57,67%	81,95%
	Taxa de Erro	8,89%	34,51%	42,33%	18,05%
	Precisão	85,18%	68,29%	55,09%	82,89%
	Sensitividade	100%	70,89%	69,78%	85,14%
	Especificidade	81,82%	58,73%	46,22%	89,7%

Fonte: A Autora (2022).

A influência de aplicação dos padrões dentro de um mesmo domínio ou extrapolando-o é mais claramente observada com a validação da Rede PE07 (Figura 41-b), aplicada a duas situações distintas: a amostras do Nordeste e a amostras do Brasil, nesse caso, todos os parâmetros indicadores de crescimento no desempenho da rede decrescem da etapa de testes para a validação de amostras do Nordeste e ainda mais para a validação de amostras do Brasil, chegando a padrões de classificação muito próximos da aleatoriedade.

A rede BR03 foi a única rede que não extrapolou o domínio de validação das amostras, já sendo desenvolvida com dados do mesmo domínio ao qual seria validada às cegas posteriormente, obtendo assim padrões mais equivalentes de desempenho entre as etapas de testes e de validação às cegas (Figura 41-b).

Figura 41 - Indicadores de Desempenho das Redes nas Etapas de Testes (a) e Validação (b)



Fonte: A Autora (2022).

O conhecimento adquirido ao construir um modelo de previsão precisa ser organizado e apresentado de maneira usual. Dependendo dos requisitos, a fase de implantação pode ser tão simples quanto gerar um relatório ou tão complexa quanto implementar um processo de aprendizado contínuo. O modelo preditivo assume a forma de uma função matemática e a expressão lógica representada pela rede neural pode ser usada para incorporá-la em outro programa com diferentes linguagens e sistemas operacionais.

4.4 Elaboração de Programa de Classificação dos Solos – Aplicação da equação da RNA BR03

A partir da tabela resumo dos testes de desempenho das RNA (Tabela 10) e das análises feitas segundo do desempenho das redes desenvolvidas, foi possível identificar que a rede BR03 alcançou melhores critérios de classificação, por isso, teve a equação matemática aplicada para o desenvolvimento de um programa de classificação de solos, que propõe uma alternativa de fácil uso e interpretação para a amostras a partir apenas de suas coordenadas geográficas, no Datum SIRGAS 2000, UTM zona 25 S, o SECBRAApp – Aplicativo de Solos Expansivos e Colapsíveis do Brasil.

A interface intuitiva localiza pelas coordenadas, os valores de cada um dos mapas temáticos fundamentais de climatologia, geologia e pedologia do Brasil, aplicando a equação da RNA, possibilitando ao usuário analisar qual a probabilidade, de um ponto em questão, assumir as características de colapso ou expansão, de acordo com os índices de incerteza da rede, e de acordo com as classes de suscetibilidade alcançadas, dispensando a necessidade de que o usuário localize em cada mapa as características de origem e formação do solo e calcule aplicando a equação e parâmetros da rede.

Alguns exemplos de classificações de solos pelo SECBRAApp são apresentados a seguir. Na classificação de amostras expansivas, Figura 42, ao inserir as coordenadas lat./long. (-49,025; -23,710) e definir o tipo de resultado esperado pelo usuário (expansão), é iniciada a execução da classificação que identificou, no referido local, clima: Sub úmido, como unidade geológica: Cobertura Cenozoica e como classe pedológica: Latossolo; resultando em uma indicação de solo expansivo, com um índice de probabilidade igual a 0,629603, que de acordo com as classes de suscetibilidade representa uma amostra com suscetibilidade baixa a expansão.

Figura 42 - Classificação de Amostra Expansiva para previsão de expansão pelo SECBRAApp

UNIVERSIDADE FEDERAL DE PERNAMBUCO GÑSAT

ENTRADA:

Coordenada X: -49,025 Coordenada Y: -23,710 Tipo de Resultado: Expansão Colapso Executar

RESULTADO:

Climatologia: Subúmido Geologia: Cobertura Cenozoica Pedologia: Latossolo Probabilidade: 0,629603 Indicativo De: Expansivo

Fonte: A Autora (2022).

Caso a mesma coordenada fosse usada para prever o tipo de resultado oposto (colapso), a classificação seguiria o mesmo processo e identificaria as mesmas características para clima, geologia e pedologia, entretanto o indicativo de tipo de solo seria 'Não colapsível' com grau de probabilidade de 0,629603, Figura 43.

Figura 43 - Classificação de Amostra Expansiva para previsão de colapso pelo SECBRAApp

UNIVERSIDADE FEDERAL DE PERNAMBUCO GÑSAT

ENTRADA:

Coordenada X: -49,025 Coordenada Y: -23,710 Tipo de Resultado: Expansão Colapso Executar

RESULTADO:

Climatologia: Subúmido Geologia: Cobertura Cenozoica Pedologia: Latossolo Probabilidade: 0,629603 Indicativo De: Não Colapsível

Fonte: A Autora (2022).

Em outra situação, ao inserir as coordenadas lat./long. (-35,237788; -8,75286) e definir o tipo de resultado previsto pelo usuário (colapso), é iniciada a execução da classificação que identificou um clima: Úmido, como unidade geológica: Borborema e como classe pedológica: Latossolo; resultando em uma indicação de solo colapsível, com um índice de probabilidade igual a 0,98565, que de acordo com as classes de suscetibilidade representa uma amostra com suscetibilidade alta ao colapso, Figura 44.

Figura 44 - Classificação de Amostra Colapsível para previsão de colapso pelo SECBRAApp

UNIVERSIDADE FEDERAL DE PERNAMBUCO GÑSAT

ENTRADA:

Coordenada X: -35,237788 Coordenada Y: -8,75286 Tipo de Resultado: Expansão Colapso Executar

RESULTADO:

Climatologia: Umido Geologia: Borborema Pedologia: Latossolo Probabilidade: 0,98565 Indicativo De: Colapso

Fonte: A Autora (2022).

Para situações em que a coordenada aplicada refere-se a uma área com indicação de unidade geológica identificada como 'Corpo d'água' e/ou classe pedológica identificada como 'Corpo d'água' ou 'Outros', essa classificação passa a não ser executada, por não apresentar precisão nessas informações de entrada, geologia e/ou pedologia, não informando o grau de suscetibilidade nem a característica de solo indicada, entretanto ainda é possível identificar as características das outras variáveis de origem e formação do solo, como clima, Figura 45.

Figura 45 - Classificação de Amostra 'Não-classificável' para previsão de colapso pelo SECBRAApp.

UNIVERSIDADE FEDERAL DE PERNAMBUCO GÑSAT

ENTRADA:

Coordenada X: -51,34251667 Coordenada Y: -20,428341667 Tipo de Resultado: Expansão Colapso Executar

RESULTADO:

Climatologia: Subumido Geologia: Parana Pedologia: Outros Probabilidade: ----- Indicativo De: Não Classificado

* Caso alguma variável de origem e formação do solo resulte em característica não conclusiva a classificação não será efetuada.

Fonte: A Autora (2022).

4.5 Carta de Classificação de Solos Colapsíveis e Expansivos do Brasil

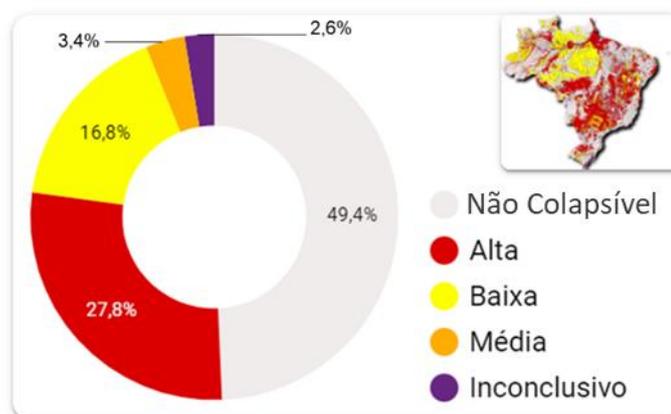
A aplicação dessas classificações para o território brasileiro, a partir da equação da rede BR03 possibilitaram a elaboração de cartas de suscetibilidade também disponíveis em ambiente *web*, o que permite a classificação de amostras, de forma espacial e não mais pontual, obtendo como

resultado - de forma gráfica, a representação em mapa da região classificada (Figuras 48 e 49), denominado SECBRMap – Mapa de Solos Expansivos e Colapsíveis do Brasil.

As cartas do Brasil para classificação de solos em ambiente *web* apresentam-se, também, a partir da inserção de coordenadas (latitude e longitude) em Graus Decimais georreferenciadas no Datum SIRGAS 2000, em escala 1:1.000.000, e teve as áreas de suscetibilidade a colapso e expansão definidas a partir do cálculo direto, baseado na equação e nos valores dos condicionantes da rede BR03.

Na carta de suscetibilidade ao colapso, Figura 48, a indicação de ocorrência desses solos se dá com uma suscetibilidade baixa em 16,8% (1.432.022 km²) da área, média em 3,4% (293.374 km²) e alta em 27,8% (2.367.032 km²), totalizando 48% do território do Brasil. As regiões classificadas na Carta como ‘não colapsível’ correspondem a 49,4% (4.203.760 km²) e 2,6% (221.849 km²) correspondem a classificação inconclusiva (localizada em corpos de água, áreas urbanas ou com registros incompletos), Figura 46.

Figura 46 - Percentuais de ocorrência de solo colapsível no Brasil.



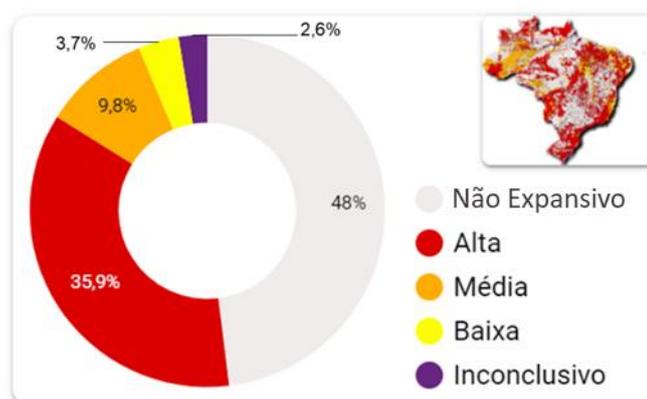
Fonte: A Autora (2022).

Os maiores destaques para o grau alto de suscetibilidade ao colapso estão na região Centro-Oeste, nos estados de Mato Grosso, sul de Mato Grosso do Sul e Goiás, entretanto esse comportamento de solos é observado em todas as regiões do país. Na região Norte destaca-se o estado do Pará em sua região nordeste. No Nordeste do país observa-se uma maior concentração na Bahia e na região litorânea do Rio Grande do Norte, Paraíba, Pernambuco e Alagoas. No Sudeste destaque para todo o interior de São Paulo e para as regiões Norte, Noroeste, Central, Sul e do Triângulo Mineiro em Minas Gerais, com níveis de suscetibilidade alta e média ao

colapso. E na região Sul há uma maior evidência no norte dos estados Rio Grande do Sul e do Paraná.

Ao analisar a carta de suscetibilidade de ocorrência de solos expansivos no Brasil, Figura 49, verifica-se que um total de 3,7% (315.878 km²) da área apresenta suscetibilidade baixa, 9,8% (830.959 km²) média e 35,9% (3.056.921 km²) alta. As demais áreas foram enquadradas como comportamento 'não expansivo' e de classificação inconclusiva (localizada em corpos de água, áreas urbanas ou com registros incompletos) que correspondem aos percentuais de 48% (4.092.428 km²) e 2,6% (221.849 km²), respectivamente, Figura 47.

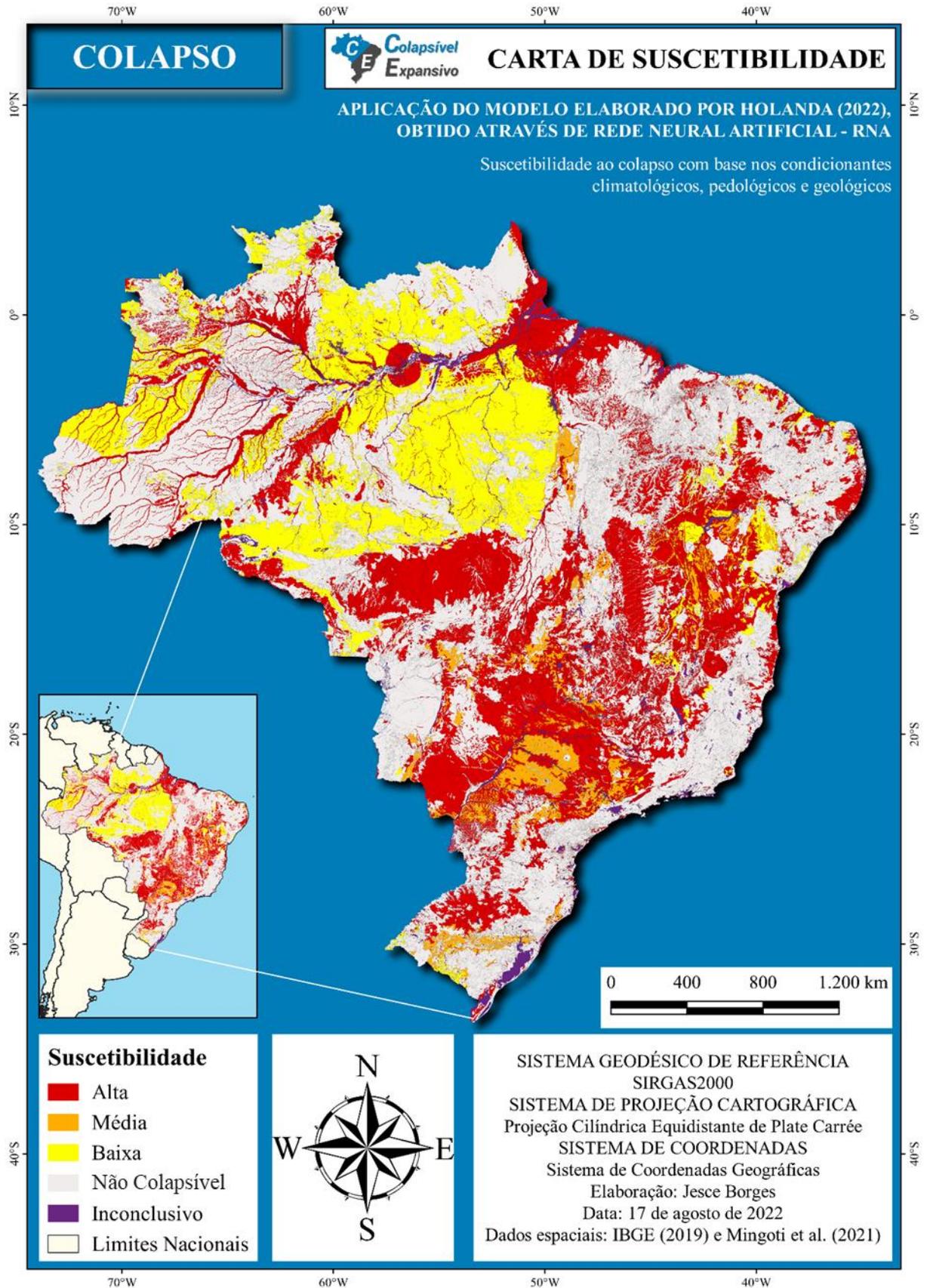
Figura 47 - Percentuais de ocorrência de solo expansivo no Brasil



Fonte: A Autora (2022).

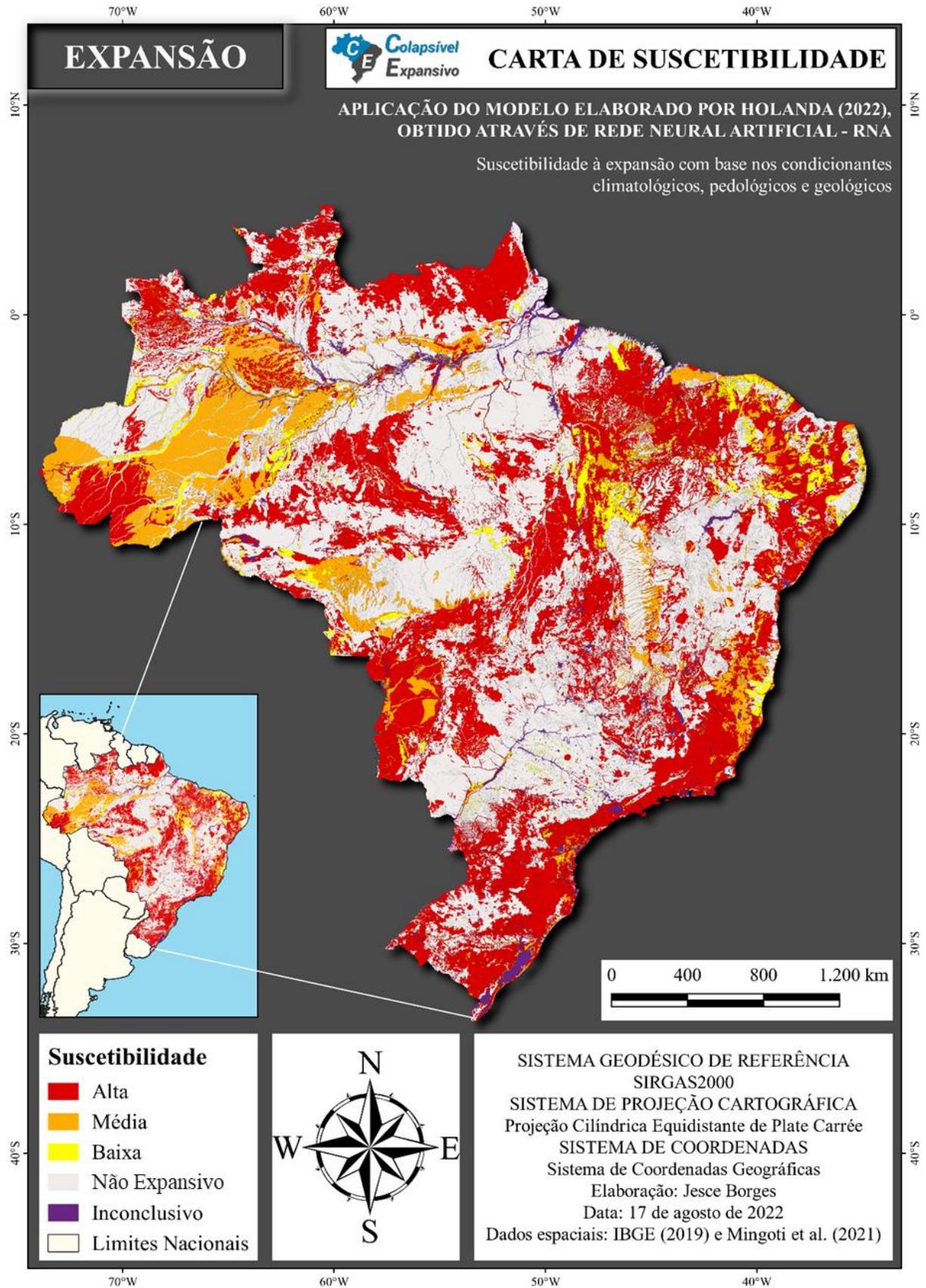
Quanto à suscetibilidade alta, é possível perceber que, na região Norte, os estados do Amapá, Roraima e Acre apresentam destaque em áreas com maior criticidade. Esse nível se faz presente também na Região Nordeste, abrangendo a maioria dos estados. No Centro-Oeste, destaca-se o Mato Grosso do Sul, que apresenta extensão de área com níveis de suscetibilidade alta e média. Em relação ao Sudeste, observa-se a suscetibilidade alta, abrangendo a parte leste da região. Na região Sul, prevalece o nível de maior criticidade à expansão.

Figura 48 – Carta derivada de suscetibilidade de ocorrência de solos colapsíveis no Brasil



Fonte: Ferreira *et al.* (2022).

Figura 49 - Carta derivada de suscetibilidade de ocorrência de solos expansivos no Brasil



Fonte: Ferreira *et al.* (2022).

5 CONCLUSÕES E SUGESTÕES

São apresentadas as principais conclusões e sugestões resultantes do trabalho desenvolvido nesta pesquisa e que podem ser aplicadas em estudos futuros de solos colapsíveis e expansivos do Brasil.

5.1 Conclusões Finais

Relembrando a nomenclatura das RNA, a rede PE04 foi desenvolvida com amostras de solos de Pernambuco a partir de 4 variáveis de entrada; a rede PE07 foi desenvolvida com amostras de solos de Pernambuco a partir de 7 variáveis de entrada; e a rede BR03, desenvolvida com amostras de solo de todo o território brasileiro, usando apenas 3 variáveis de entrada.

As Redes Neurais Artificiais possibilitam a resolução de problemas complexos e ainda não totalmente compreendidos, resultando em estimativas de probabilidade de classificação de solos colapsíveis e expansivos com base no conjunto das variáveis intrínsecas e extrínsecas, que são variáveis de simples obtenção. Dessa forma, as principais conclusões são:

- O desenvolvimento das redes neurais acompanhou o crescimento do BANDASE, partindo de 132 amostras (87 de Pernambuco + 45 de outros estados do Nordeste) para 526 amostras (157 de Pernambuco + 369 dos demais estados do Brasil);
- Os valores das variáveis de entrada que mudam a classificação da saída para a rede PE04, (4 variáveis quantitativas de entrada) indicam que o solo expansivo deve apresentar um conjunto de características físicas com: percentual da fração argila maior que 30%, percentual da fração areia menor do que 48%, índice de plasticidade maior que 18% e atividade maior que 0,60; e para o solo ser colapsível os limites indicados deverão ser inversos;
- Os valores das variáveis de entrada que mudam a classificação da saída para a rede PE07 (4 variáveis quantitativas e 3 variáveis condicionantes do meio físico como entradas) indicam que para o solo ser expansivo deve apresentar um conjunto de características físicas com: percentual da fração argila maior que 31%, percentual da fração areia menor do que 47%, índice de plasticidade maior que 18% e atividade maior que 0,60 e para ser colapsível os limites indicados para o solo são inversos. As correlações das variáveis condicionantes do meio físico

(categóricas) contribuem com escores de 0,50 para a climatologia, 0,62 para pedologia e 0,48 para geologia; para o solo ser expansivo.

- A correlação entre as variáveis de entrada e saída para a rede BR03 (3 variáveis condicionantes do meio físico como entradas) indica que para o solo ser expansivo deve apresentar variáveis condicionantes do meio físico (categóricas) para a climatologia contribuindo com escore com 0,33, pedologia com escore de -0,48 e geologia com 0,5.
- A rede PE04 apresenta taxas de probabilidade de ocorrência de solos colapsíveis e expansivos com acurácia acima de 76%, embora tenha sido verificado uma grande ausência de informações de entrada nesse banco de dados usado para desenvolver a rede. Todas as amostras classificadas erroneamente na etapa de testes apresentavam ausências de informações, destacando a importância de todas as informações de entrada. Ao ser validada às cegas com amostras de outros estados do Nordeste a taxa de acurácia foi maior que 91%.
- A rede PE07 demonstra taxa de acurácia maior que 88%, apresentando excelência no reconhecimento das características de solos expansivos (taxa de sensibilidade 100%). A aplicação das amostras foi a outros estados do NE e posteriormente a amostras de todo o Brasil e apontou um considerável decréscimo em seus critérios de classificação, sendo 65,49% classificando amostras do Nordeste e 57,67% na classificação de amostras de todo o Brasil, situação em que a rede PE07 apresenta-se ineficaz por ter uma classificação próxima a aleatoriedade;
- A rede BR03 apresenta um percentual de acerto maior que 89%, e durante a validação da rede às amostras do mesmo domínio, manteve a taxa de acurácia em 82%, sendo identificada como a rede de melhor desempenho;
- O programa – SECBRApp desenvolvido a partir da equação da rede BR03 classifica o solo em expansivo ou não expansivo, colapsível ou não colapsível; e é disponibilizado em ambiente web, permitindo a identificação dos solos a partir, apenas, da localização geográfica;
- As cartas de suscetibilidade com classificações alta, média, baixa e não colapsível/não expansivo, desenvolvidas para os solos do Brasil, podem ser usadas para o planejamento de uso e ocupação dos solos, bem como para o anteprojeto de obras em geral. Entretanto, não devem substituir os ensaios de campo e de laboratório em projetos, que são indispensáveis para quaisquer obras de engenharia;

- Por fim, conclui-se que o modelo de redes neurais proposto representa uma ferramenta útil e promissora, sendo possível a melhoria do modelo com cada novo caso testado, tendo em vista que a utilização dessa ferramenta no Brasil ainda seja pouco disseminada no contexto da geotecnia. Para melhor eficiência destas identificações de solos é necessário a adequada precisão das coordenadas geográficas;
- É possível identificar e classificar a suscetibilidade de ocorrência de solos colapsíveis ou expansivos utilizando condicionantes climatológicos, geológicos, pedológicos e geotécnicos por meio de Redes Neurais Artificiais, podendo-se inferir que a hipótese da pesquisa é verdadeira.
-

5.2 Sugestões Para Futuras Pesquisas

- a) Continuar atualizando o BANDASE com mais dados de solos colapsíveis e expansivos de outras regiões brasileiras, para aumentar a representatividade do conjunto experimental utilizado neste trabalho;
- b) Aplicar a metodologia desta pesquisa na elaboração de cartas de suscetibilidade ao colapso e expansão regionais utilizando mapas em escala maior;
- c) Comparar a função estimada pelo modelo de RNA desenvolvido com as funções previstas por outros métodos de identificações, como: Seed *et al.* (1960), Daksanamurthy e Raman (1973); Rodrigues Ortiz (1975), Jennings e Knight (1975);
- d) Verificar a sensibilidade do modelo, aplicando o programa cada vez mais à identificação de novos solos de diferentes regiões;
- e) Desenvolver um modelo de Redes Neurais Artificiais com classificação múltipla (ou reconhecimento de padrões), de modo a reconhecer de maneira discriminada as características de solos colapsíveis, expansivos e não colapsíveis e não expansivos.

REFERÊNCIAS

- ABIODUN, A., NALBANTOGLU, Z. Técnicas de estacas de cal para a melhoria de solos argilosos. **Revista Brasileira de Geotecnia**, 52 (6) (2015), pp. 760 - 768.
- ABNT - ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 13133**: Execução de Levantamento Topográfico. Rio de Janeiro, 1994.
- ALAN, S. K.; MONDAL, A.; SHIULY, A. Prediction of CBR Value of Fine-Grained Soils of Bengal Basin by Genetic Expression Programming, Artificial Neural Network and Krigging Method. **J Geol Soc India** **95**, 190–196 (2020). <https://doi.org/10.1007/s12594-020-1409-0>
- AL-BUSULTAN, S.; ASWED, G. K.; ALMUHANNA, R. R.; RASHEED, S. E. (2020). Application of artificial neural networks in predicting subbase CBR values using soil indices data. In IOP Conference Series: **Materials Science and Engineering** (Vol. 671, No. 1, p. 012106).
- ALMEIDA, G. B. de O. Avaliação da Resistência à Compressão Simples de Misturas de Solo e Escória de Cobre Aditivadas com Cimento Portland Aplicadas a Pavimentação. In: XVIII Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica, 2016.
- ALONSO, E. E.; GENS, A.; JOSA, A. (1990). A constitutive model for partially saturated soil. **Geotechnique**, 40(3): 405-430.
- AMORIM, S. F. **Contribuição à cartografia geotécnica: sistema de informações geográficas dos solos expansivos e colapsíveis do Estado de Pernambuco (SIGSEC – PE)**. 2004. Dissertação (Mestrado em Engenharia Civil), Programa de Pós-Graduação em Engenharia Civil, Universidade Federal de Pernambuco, Recife, 2004.
- AMORIM, S. F. **Estudio comparativo de métodos para la evaluación de la susceptibilidad del terreno a la formación de deslizamientos superficiales: Aplicación al Pirineo Oriental**. 2012. Tesis Doctoral. Universitat Politècnica de Catalunya. Barcelona, 2012.
- AQUINO, A. E. B.; FERREIRA, S. R. M. Análise Geotécnica e Caracterização Geomecânica de um Solo Expansivo de Teresina-PI. In: ANAIS DO XX COBRAMSEG, 2022, Campinas. **Anais [...]**. Campinas, Galoá, 2022.
- ASHAYERI, I.; YASREBI, S. Free-swell and swelling pressure of unsaturated compacted clays; experiments and neural networks modeling. **Geotechnical and Geological Engineering**. 27 (1), 137-153. 2009.
- AYALA, J.F.; GIJON, M.F.; MOZO, C.O.; RODRIGUEZ J.L.S. Mapa Previsor de Riesgos por Expansividad de Argilas Em España a Escala 1:1.000.000. Série: **Geologia Ambiental**. Instituto Geologico y Minero de España. Centro de Estudios y Experimentacion de Obras Públicas. España, Madrid, 1986.
- BARBOSA, F. A. S. **Análise do comportamento hidromecânico e da dinâmica de fissuração de um solo expansivo**. 2019. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal de Pernambuco, Recife, 2019.

BARBOSA, V. (2013). **Estudo do comportamento geotécnico de um solo argiloso de Cabrobó, potencialmente expansivo, estabilizado com cal**. Dissertação de MSc, Universidade Federal de Pernambuco, Recife, PE, 115p.

BARROS, L. B. **Aplicação de Redes Neurais Artificiais no Contexto de Estimativa de Custos de Construção de Rodovias**. Programa de Pós-graduação em Estruturas e Construção Civil. Brasília, janeiro de 2019.

BARROS, L. B.; MARCY, M.; CARVALHO, M. T. M. Construction Cost Estimation of Brazilian Highways Using Artificial Neural Networks. **International Journal of Structural and Civil Engineering Research**. Vol. 7, No. 3, August 2018.

BASHEER, I. A.; REDDI, L. N.; NAJJAR, Y. M. (1996) Caracterização do local por neuronetes: uma aplicação ao problema de assentamento em aterros sanitários. **Água Subterrânea 34**: 610–617.

BASMA, A. & KALLAS, N. (2004). Modeling soil collapse by artificial neural networks. **Geotechnical and Geological Engineering**. 22. 427-438. 10.1023/B:GEGE.0000025044.72718.db.

BASMA, A. A.; BARAKAT, S. A.; OMAR, M. Modeling time dependent swell of clays using sequential artificial neural networks. **Environmental and Engineering Geoscience**, 9 (3) (2003), p. 279-288.

BASTOS, E. G. (1994). **Variação volumétrica de uma argila expansiva do litoral de Pernambuco**. 155f. Dissertação (Mestrado em Engenharia Civil). Universidade Federal da Paraíba, Campina Grande, 1994.

BAUCHSPIESS, A. (2002). **Sistemas Inteligentes Aplicados à Solução de Problemas de Engenharia**. Semana da Engenharia Civil, Mini-Curso, Faculdade de Tecnologia, Departamento de Engenharia Civil, Universidade de Brasília, Brasília, DF, 50p.

BAZIAR, M. H.; GHORBANI, A. (2005) Evaluation of lateral spreading using artificial neural networks. **Soil Dyn Earthq Eng** 25(1):1–9.

BENATTI, J. C. B. **Colapsividade com sucção controlada de um solo coluvionar e laterítico de Campinas/SP**. Campinas, 2010. 189 p. Dissertação (Mestrado em Engenharia Civil) - Faculdade de Engenharia Civil, Arquitetura e Urbanismo, Universidade Estadual de Campinas.

BESQ, A.; MALFOY, C.; PANTET, A.; MONNET, P.; RIGHI, D. Physicochemical characterization and flow properties of some bentonite muds. **Applied Clay Science**, 23 (5–6) (2003), pp. 275-286.

BISHOP, C. M. Neural networks for pattern recognition. **Oxford University Press**, New York, 1995.

BORGES, M. V. (2019). **Cartografia geotécnica da cidade de Rio Branco – Acre: uso de estatística descritiva, geoestatística e estatística multivariada**. Publicação G.DM-150/19. Tese de Doutorado em Geotecnia. Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 290p.

BRAGA, A.P.; LURDERMIR, T.B.; CARVALHO, A.C.L.F. (2000). **Redes Neurais Artificiais: Teoria e Aplicações**. Editora LTC, Rio de Janeiro, RJ, 262p.

BRIAUD, J-L.; ZHANG, X.; MOON, S. Shrink test-water content method for shrink and swell predictions. **Journal of Geotechnical and Geoenvironmental Engineering**, 2003; 129(7):590–600.

BURGOS, P. C.; CAMPOS, L. E. P. (2002). Estudo do Potencial Expansivo de um Solo Saprolítico de Granulito. **In.:** XII COBRAMSEG. Vol. 2. pp. 639-648.

CABRAL, M. A. L. **Classificação Automatizada de Falhas Tribológicas de Sistemas Alternativos com o Uso de Redes Neurais Artificiais Não Supervisionadas**. Tese (doutorado) - Universidade Federal de Rio Grande do Norte, Centro de Tecnologia, Natal, RN, 2017.

CAL, Y. (1995) Soil classification by neural-network. **Adv Eng Softw** 22(2):95–97.

CAMAPUM DE CARVALHO, J. *et al.* (Orgs.) (2015). Solos não saturados no contexto geotécnico. **ABMS**, São Paulo, cap. 25: 687-731.

CARVALHO, P. G; ANDRADE, T.B. Solos Problemáticas: expansivos e colapsíveis. Universidade Potiguar – RN. 2015. **In.:** Prezi. Natal – RN.

CAVALCANTE, E. H.; CAVALCANTI Jr.; SANTOS, W. J.; SOUZA NETO, J. B. (2006). Propriedades geotécnicas de um solo expansivo de Sergipe. **In.:** II GEOJOVEM, 2006, Nova Friburgo - RJ. GEOJOVEM 2006.

CAVALCANTI, R. S.; SILVA, F. A. N. Utilização de Ensaio Não Destrutivo de Ultra-Som e Redes Neurais Artificiais na Estimativa da Resistência à Compressão do Concreto. **In.:** 54 Congresso Brasileiro do Concreto, 2012, Maceió. 54 Congresso Brasileiro do Concreto. São Paulo: Ibracon, 2012. v. 1.

CELIK, S.; TAN, O. (2005). Determination of pre consolidation pressure with artificial neural network. **Civil Eng Environ Syst** 22(4):217–231.

CHAKRABORTY, A.; GOSWAMI, D. (2017). Prediction of slope stability using multiple linear regression (MLR) and artificial neural network (ANN). **Arab J Geosci** 10:11.

CHEN, F. H. (1975). Foundation on Expansive Soils. **Amsterdam: Elsevier**, 275 - 463 p.

CHOUDHURY, T; COSTA, S. Prediction of Parallel Clay Cracks Using Neural Networks – A Feasibility Study. **School of Science**, Engineering and IT, Federation University Australia, Ballarat, VIC, Australia. 2019.

CHRISTOPHER, I. C.; CHIMOB, N. D. Emerging trends in expansive soil stabilisation: A review. **Journal of Rock Mechanics and Geotechnical Engineering**, p. 37, 2019.

CLEMENCE, S.P.; FINBARR, A.O. (1981). “Design consideration for collapsible soils”. **Journal of Geotechnical Engineering**, ASCE 107 (3), pp. 305– 317.

CODUTO, D. P. Geotechnical Engineering Principles and Practices. **Prentice-Hall: Upper Saddle River**, NJ, 1999.

CPRM - Campanha de Pesquisa de Recursos Minerais (2021). **Geologia e Recursos Minerais do Estado de Pernambuco**. DNPM, GOVERNO DO ESTADO DE PERNAMBUCO, AD/Diper., Recife.

DANTAS NETO, A. S.; INDRARATNA, B.; OLIVEIRA, D. A. F. Prediction of the Shear Behaviour of Clean Joints in Soft Rocks Using Perceptron. In: VII SIMPÓSIO BRASILEIRO DE MECÂNICA DAS ROCHAS, 2016, Belo Horizonte. **Anais[...]**. Belo Horizonte, Minas Gerais: SBMR, 2016. p. 1817-1831.

DANTAS NETO, S. A.; INDRARATNA, B.; OLIVEIRA, D. A. F.; ASSIS, P. Modelling the Shear Behaviour of Clean Rock Discontinuities Using Artificial Neural Networks. **Rock Mech. Rock Eng.**, v. 50, p. 1817-1831, 2017.

DEHGHAN, S.; SATTARI, G.; CHELGANI, C. S.; ALIABADI, M. A. Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression and artificial neural networks. **Min. Sci. Technol.**, v. 20, p. 41-46, 2010.

DELAZZERI, M. L. C. **Estimativa da Capacidade de Carga de Fundações Profundas Utilizando Rede Neural Artificial**. 2018. Dissertação (Mestrado em Engenharia Civil e Ambiental) - Universidade Estadual de Feira de Santana.

DELAZZERI, M.L.C.; GADÉA, A.S.M.; NAGAHAMA, K.J.; MATEUS, M.S.C.S. Utilização de redes neurais artificiais na engenharia geotécnica. **In.:** 1º Seminário Geotécnico Bahia/Sergipe, 2017, Feira de Santana. 1º Seminário Geotécnico Bahia/Sergipe, 2017.

DORIS, J. J.; RIZZO, M. M.; DEWOOLKAR, M. M. Forecasting vertical ground surface movement from shrinking/swelling soils with artificial neural networks. **Int. J. Numer. Anal. Meth. Geomech.** 32:1229–1245. 2008. DOI: 10.1002/nag.666

DUDLEY, H. H. (1970) Review of Collapsing Soils. **ASCE Journal of Soil Mechanics and Foundation Division**, 96, 925-947.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária (2001). **Zoneamento Agroecológico do Estado de Pernambuco**. CD-ROM. Serviço Nacional de Levantamento e Conservação de Solos, Documentos; nº. 35, Recife.

ENGEGRAUT. GROUTING - SOLOS COLAPSÍVEIS E EXPANSIVOS TRATAMENTO. Dezembro, 2015.

ERZIN, Y. (2007). Artificial neural networks approach for swell pressure versus soil suction behavior. **Can Geotech J** 44(10):1215–1223.

ESTABRAGH, A. R.; PARSAEI, B.; JAVADI, A. A. Laboratory investigation of the effect of cyclic wetting and drying on the behavior of an expansive soil. **Soils and Foundations**, v. 55, n. 2, p. 304–314, 2015.

FATEHNIA, M. AMIRINIA, G. A review of genetic programming and artificial neural network applications in pile foundations. **International Journal of Geo Engineering**, 9 (2018), p. 2. DOI: <https://doi.org/10.1186/s40703-017-0067-6>.

FERNÁNDEZ-STEEGER, T. M.; ROHN, J.; CZURDA, K. Identification of landslide areas with neural nets for hazard analysis. **In.:** Landslides. Routledge, 2018. p. 163-168.

FERREIRA, R. N.; BASTOS, E. G.; JUCÁ, J. F. T (1998b). Estudo de Laboratório e Campo de um Solo Expansivo. **In.:** XI COBRAMSEG. Vol. 1. pp. 21-26.

FERREIRA, R. N.; JUCÁ, J. F. T; SILVA, J. M. J. (1998a). Resistência de uma Argila Expansiva Utilizando o Equipamento de Cisalhamento Direto com Sucção Controlada. **In.:** XI COBRAMSEG. Vol. 1. pp. 1-6.

FERREIRA, S. R. M. Classes de Solos com Potencialidade de Apresentarem Solos Colapsíveis e Expansivos. **In.:** XXIII Congresso Brasileiro de Ciência dos Solos, 1991, Porto Alegre - RS. XXIII Congresso Brasileiro de Ciência dos Solos, 1991. v. 1. p. 135-135.

FERREIRA, S. R. M. Mapa de Risco a Colapsibilidade e Expansividade de Solo do Estado de Pernambuco – Geotecnia Ambiental. **Relatório Técnico do CNPq**. 1999.

FERREIRA, S. R. M.; BORGES, J. J. S.; HOLANDA, M. J. O.; AMORIM, S. F. Analysis of volume change and crack propagation in swelling soil due to water content change. **In.:** **Revista Brasileira de Ciência do Solo** (no prelo). 2022.

FERREIRA, S. R. M.; FERREIRA, M. G. V. X. Mudanças de volume devido à variação de teor de umidade em um Vertisol no Semi-Árido de Pernambuco. **Revista Brasileira de Ciência do Solo** (Impresso), v. 33, p. 779-791, 2009.

FERREIRA, S. R. M.; TEIXEIRA, D. C. L. Collapsible Soil - A Practical Case in Construction Pernambuco - Brazil. **In.:** 12th International Conference on Soil Mechanics and Foundation Engineering, 1989, Rio de Janeiro. v. 1. p. 603-609.

FERREIRA, S. R. M. Banco de Dados de Solos Especiais - Colapsíveis no Estado de Pernambuco. **In.:** 6º CBGE/IX COBRANSEF. 1990.

FERREIRA, S. R. M. **Colapso e Expansão em Solos Naturais não Saturados Devido à Inundação**. Tese de Doutorado. COPPE/UFRJ, Rio de Janeiro, RJ, 1995.

FERREIRA, S. R. M.; AMORIM, S. F.; VAREJAO-SILVA, M. A. "Tecnologia da geoinformação aplicada aos estudos dos solos expansivos e colapsíveis do estado de Pernambuco", **In.:** Simpósio Brasileiro de Ciência Geodésicas e Tecnologias de Geoinformação, 2, v.1. p. 91-99, Recife, 2008.

FRANCO, J. A. M.; PIRES FILHO, C. J.; MACÊDO JÚNIOR, P. M. *et al.*, "A Contribuição de Estruturas Reliquiarias para Rupturas Localizadas dos Taludes em Solo Residual nas Obras do AHE Simplício Queda-Única – RJ". **In.:** XV COBRAMSEG: pp. 1852 – 1857, Gramado, Ago. 2010.

FREDLUND, D. G.; XING, A. Equations for the soil-water characteristic curve. **Canadian Geotechnical Journal**, v. 31 n° 3, p. 521-532, mar. 1994.

FREDLUND, D. G.; RAHARDJO, H. (1993) Soil Mechanics for Unsaturated Soils. **John Wiley and Sons**, Inc., New York. DOI: <https://doi.org/10.1002/9780470172759>

FREITAS, M. L. R. A. (2017) **Comportamento Geomecânico de um Solo Colapsível de Petrolina- PE avaliado por meio de Ensaio de Campo e Laboratório**. Tese de Doutorado, Recife: UFPE, 143p.

FUTAI, M. M. Comportamento Colapsível De Solos Tropicais Brasileiros. Congresso Nacional de Geotecnia. **In:** Congresso Nacional de Geotecnia, 2000, Porto. Congresso Nacional de Geotecnia. Porto, 2000.

FUTAI, M. M.; ALMEIDA, M. S. S.; LACERDA, W.A (2001). Propriedades geotécnicas de algumas argilas do Rio de Janeiro. Encontro sobre solos moles. COPPE-UFRJ.

GARDNER, B. J.; GRANSBERG, D. D.; RUEDA, J. A. Stochastic conceptual cost estimating of highway projects to communicate uncertainty using bootstrap sampling. **Journal of Risk and Uncertainty in Engineering Systems**, Part A: Civil Engineering. 2016. DOI: 10.1061/AJRUA6.0000895

GOH, A. T. (2002). Probabilistic neural network for evaluating seismic liquefaction potential. **Can Geotech J** 39(1):219–232.

GOMES, R. L.; MARQUES, E. A. G.; FRANCO, G. B. The waste disposal suitability of Almada River Watershed. **Eng Sanit Ambient** | v.22 n.4 | jul/ago 2017 | 731-747. DOI: 10.1590/S1413-41522017157490.

GORDAN, B.; ARMAGHANI, D. J.; HAJIHASSANI, M.; MONJEZI, M. Prediction of seismic slope stability through combination of particle swarm optimization and neural network. **Eng. Comput.**, 32 (1) (2016), pp. 85-97.

GRANATYR, J. IA Expert. Conceitos sobre IA, IA simbólica X IA conexionista. 2017. Disponível em: <<https://iaexpert.com.br/index.php/2017/03/23/ia-simbolica-x-ia-conexionista/>>. Acesso em: 02.03.2020.

GRIBB, M. M.; GRIBB, G. W. (1994). Use of neural networks for hydraulic conductivity determination in unsaturated soil. **In.:** Proceedings of the 2nd international conference on ground water ecology, Bethesda, pp 155–163.

HALLAL, R. R. **Características de Depósitos de Argilas Moles no Estado do Rio Grande do Sul**. 2003. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Porto Alegre. 2003, 150p.

HANNA, A. M.; URAL, D.; SAYGILI, G. (2007). Neural network model for liquefaction potential in soil deposits using Turkey and Taiwan earthquake data. **Soil Dyn Earthq Eng** 27(6):521–540.

HAYKIN, S. (2001). Redes Neurais: Princípios e Prática. **Bookman**, Porto Alegre, RS. 900p.

HAYKIN, S. Neural network: a comprehensive foundation. 2. ed. Michigan, USA: Macmillan, 2004.

HAYKIN, S. *Neural Networks and Learning Machines*. 3rd ed. Ontario, Canada: Pearson, 2008.

IBGE - Instituto Brasileiro de Geografia Estatística. Banco de Dados de Informações Ambientais. <<https://www.ibge.gov.br/geociencias/informacoes-ambientais/pedologia/10871-pedologia.html>>. Acesso em fevereiro de 2022.

_____. - Instituto Brasileiro de Geografia Estatística. Brazil: geology [map]. 1:250,000. 2019b [cited 2022 Oct 12]. Available from: <https://portaldemapas.ibge.gov.br/portal.php#mapa222449>

_____. - Instituto Brasileiro de Geografia Estatística. Brazil: pedology [map]. 1:250,000. 2019a. [cited 2022 Oct 12]. Available from: <https://portaldemapas.ibge.gov.br/portal.php#mapa222451>

IKEAGWUANI, C. C.; NWONU, D. C. Emerging trends in expansive soil stabilization: A review. **Journal of Rock Mechanics and Geotechnical Engineering** 11 (2019) 423-440.

IKIZLER, S. B.; AYTEKIN, M.; VEKLI, M.; KOCABAŞ, F. Prediction of swelling pressures of expansive soils using artificial neural networks. **Advances in Engineering Software**. Volume 41, Issue 4, April 2010, p. 647-655.

JAKSA, M. B. The Influence of Spatial Variability on the Geotechnical Design Properties of a Stiff, Over consolidated Clay. **Advisor: W. S. Kaggwa, P. I. Brooker**. November, 1995.

JIMENEZ SALAS, J. A. 1980. Cimentaciones en terrenos expansivos o colasables. **Geotecnia y Cimientos III**, 1, 533-650.

JONES, D. E.; HOLTZ, W.G. 1974. Expansive soils – the hidden disaster. Civil Eng., **ASCE**, New York, NY pp 87-89, Aug. 1973.

JUANG, C. H., NI, S. H. LU, P. C. (1999), Training Artificial Neural Networks with the Aid of Fuzzy Sets. **Computer-Aided Civil and Infrastructure Engineering**, 14: 407-415. DOI: <https://doi.org/10.1111/0885-9507.00159>

JUSTINO DA SILVA, J. M. Solos Expansivos. Livro “Geotecnia no Nordeste”, **ABMS-NRNE**, Cap. 8, pp. 397-418. 2005.

JUSTINO DA SILVA, J. M. **Variação Volumétrica de uma Argila Expansiva não Saturada Submetida a Diferentes Condições Climáticas**. 249 f. Tese (Doutorado em Engenharia Civil). Escola Politécnica da Universidade de São Paulo, São Paulo. 2001.

JUSZCZYK, M.; LEŚNIAK, A.; ZIMA, K. ANN based approach for estimation of construction costs of sports fields. **Complexity**, v. 2018, 2018.

KHANDELWAL, U.; HE, H.; QI, P.; JURAFSKY, D. Sharp Nearby, Fuzzy Far Away: How Neural Language Models Use Context. **Computer Science Department Stanford University**. 2018.

KIM, G. H.; AN, S.H.; KANG, K.I. "Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. **Building and Environment**, vol. 39, no. 10, pp. 1235–1242, October 2004.

KIM, Y., KIM, B. (2006). Use of artificial neural networks in the prediction of liquefaction resistance of sands. **J Geotech Geoenviron Eng** 132(11):1502–1504

LATIFI, N. *et al.* Tropical residual soil stabilization: A powder form material for increasing soil strength. **Construction and Building Materials**, v. 147, p. 827–836, 2017.

LEE, S. J.; LEE, S. R.; KIM, Y. S. (2003). An approach to estimate unsaturated shear strength using artificial neural network and hyperbolic formulation. **Comput Geotech** 30(6):489–503.

LEITE, A. R. S. **Aplicações de redes neurais artificiais para previsão do comportamento cisalhante em descontinuidades de maciços rochosos**. Dissertação (mestrado) – Universidade Federal do Ceará, Programa de Pós-Graduação em Engenharia Civil: Geotecnia, Fortaleza, 2019.

LEPSCH, I. F. **19 Lições de Pedologia**. São Paulo. Oficina de Textos. 2011, 456p

LIAN, C.; ZENG, Z.; YAO, W.; TANG, H. Multiple neural networks switched prediction for landslide displacement. **Engineering Geology**. Volume 186, 24 February 2015, p.91-99.

LODI, P. C.; LOLLO, J. A. Propriedades dos solos colapsíveis. **In.:** LOLLO, J. A. (Org.). Solos colapsíveis: identificação, comportamento, impactos, riscos e soluções tecnológicas. São Paulo. Cultura Acadêmica: Universidade Estadual Paulista, Pró Reitoria de Graduação, 2008.

LOLLO, J. A. (2008). **Solos Colapsíveis: Identificação, comportamento, impactos, riscos e soluções tecnológicas**. Editora Cultura Acadêmica. UNESP/Ilha Solteira.

LUNDSTROM, K.; LARSSON, R.; DAHLIN, T. (2009). Mapping of quick clay formations using geotechnical and geophysical methods. **Landslides** 6:1–15

LYTTON, R. L. Prediction of movement in expansive clays. **Proceedings of the Settlement '94 Conference on Geotechnical Special Publication** N. 40, Reston, Virginia, 1994; 1827–1845.

MAGALHÃES, P. R. F.; SOARES, A. B. Uso de argilas expansivas em camadas de impermeabilização de coberturas de aterros sanitários no Nordeste do Brasil. **Engenharia Sanitária e Ambiental [online]**. 2021, v. 26, n. 6 [Acessado 19 Agosto 2022], pp. 1069-1077. Disponível em: <<https://doi.org/10.1590/S1413-415220200206>>.

MARCY, M. *et al.* Locating damages in beams with artificial neural network. **International Journal of Lifecycle Performance Engineering**, v. 1, n. 4, p. 398–413, 2014.

MARINHO, F. A. M. Medição de Sucção em Solos. **In.:** SIMPÓSIO BRASILEIRO DE SOLOS NÃO SATURADOS, 3., Rio de Janeiro, 1997, **Anais[...]**. pp. 399-416. Rio de Janeiro: Ed. Freitas Bastos, 1997.

MAYORAZ, F.; CORNU, T.; VUILLET, L. (1996). Using neural networks to predict slope movements. **In.:** Proceedings of VII international symposium on landslides, Trondheim, June 1966, 1. Balkema, Rotterdam, p.295–300

MCDOWELL, C. Interrelationships of loads, volume change, and layer thickness of soils to the behavior of engineering structures. **In.:** Proceedings of the 35th Annual Meetings, Publication N. 426, Washington, DC, 1956; 754–772.

MCKEEN, G. R. A model for predicting expansive soil behavior. **Proceedings of the 7th International Conference on Expansive Soils**. Texas Tech University: Lubbock, Texas, 1992; 1–6.

MEDEIROS, B. **Análise por meio de redes neurais artificiais dos dados de monitoramento dos piezômetros da barragem de concreto de Itaipu**. 2013. Dissertação (Mestrado em Geotecnia) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2013.

MELO, T. M.; PEDROLLO, O. C. (2015). **Artificial Neural Networks for Estimating Soil Water Retention Curve Using Fitted and Measured Data**, 2015.

MENDONÇA NETO, H. C. **Análise da variação de volume devido à inundação de um solo colapsível utilizado em um canal de irrigação**. Dissertação de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco. 2011.

MEROUANE, F. Z.; MAMOUNE, S. M. Prediction of Swelling Parameters of Two Clayey Soils from Algeria Using Artificial Neural Networks. **Mathematical Modelling in Civil Engineering** 14(3):11-26. Vol. 14-No. 3: 11-26- 2018. DOI: 10.2478/mmce-2018-0008

MEULENKAMP, F.; GRIMA, M. A. Application of neural networks for the prediction of the unconfined compressive strength (UCS) from Equotip hardness. **Int. J. Rock Mech. Min. Sci.**, v. 36, p. 29-39, 1999.

MILLER, D. J.; DURKEE E. B.; CHAO, K. C.; NELSON, J. D. Simplified heave prediction for expansive soils. **Proceedings of the 1st International Conference on Unsaturated Soils**, Rotterdam, Netherlands, 1995; 891–897.

MOAYEDI, H.; REZAEI, A. An artificial neural network approach for under-reamed piles subjected to uplift forces in dry sand. **Neural Comput & Applic** 31, 327–336 (2019). DOI: <https://doi.org/10.1007/s00521-017-2990-z>

MOLAABASI, H.; SABERIAN, M.; KHAJEH, A.; LI, J.; CHENARI, R. J. Settlement predictions of shallow foundations for non-cohesive soils based on CPT records-polynomial model, **Computers and Geotechnics**, Volume 128, 2020. DOI: <https://doi.org/10.1016/j.compgeo.2020.103811>.

MONICO, J. F. G. **Posicionamento pelo GNSS: Descrição, Fundamentos e Aplicações**. 2 ed. São Paulo, Editora UNESP, 2008.

MONROY, R.; ZDRAVKOVIC, L.; RIDLEY, A. M. Mechanical behaviour of unsaturated expansive clay under Ko conditions. **Engineering Geology**, v. 197, p. 112–131, 2015.

MOOSAVI, M., YAZDANPANA, M. J., DOOSTMOHAMMADI, R. Modeling the cyclic swelling pressure of mudrock using artificial neural networks. **Eng. Geol.**, 2006,87, 178–194.

MORALES, E. M.; MORALES, M. K. 2004. “**Expansive Soils – Identification, Detection and Remediation Strategies**” 3–5. www.pgatech.com.ph.

MOSHREFII, S.; SHAHRIAR, K.; RAMEZANZADEH, A.; GOSHTASBI, K. Prediction of ultimate strength of shale using artificial neural network. **Journal of Mining & Environment**, v. 9, n. 1, p. 91-105, 2018.

MURTHY, V. N. S. (2003). Geotechnical Engineering: Principles and Practices of Soil Mechanics and Foundation Engineering. **Marcel Dekker**, Inc. New York.

NADIM, F.; PEDERSEN, S. A. S.; SCHMIDT-THOMÉ, P.; SIGMUNDSSON, F.; ENGDahl, M. (2008). **Natural Hazards in Nordic Countries**. Episodes 31(1):176–184.

NAKASHIMA, C. H.; SAITO, L. A.; HELLA, M. **Análise da estabilização de solo expansivo em plataforma ferroviária**. 2014. 115 f. Trabalho de Conclusão de Curso (Graduação) – Universidade Tecnológica Federal do Paraná, Curitiba, 2014.

NELLES, O. (2001). Nonlinear system identification: from classical approaches to neural networks and fuzzy models. **Springer**, New York. 785pp. DOI: <http://dx.doi.org/10.1007/978-3-662-04323-3>

NEVES, J. M. C. **Estabilização de solos com cal – Estudo e aplicações rodoviárias**. Dissertação (Mestrado) da Universidade Nova Lisboa, Lisboa.1993.

NYS DHSES - New York Division of Homeland Security and Emergency Services. 2014. **New York State Standard Multi-Hazard Mitigation Plan Draft**, Section 3.13: Land Subsidence and Expansive Soils Hazard. Disponível em:<<http://www.dhSES.ny.gov/oem/mitigation/documents/2014-shmp/2014-SHMP-full.pdf>>.

NOGUEIRA SILVA, F. A.; MAIA FERREIRA GOMES, M. E.; ALMEIDA DE OLIVEIRA, R. Dimensionamento de pilares de concreto armado de seções retangular e circular maciça submetidos à flexão oblíqua composta utilizando redes neurais artificiais. **Ciência & Engenharia**, v. 25, p. 27-33, 2016.

NOLA, I. T. S. **Avaliação de dados geológico-geotécnicos prévios para elaboração de carta de eventos perigosos de movimentos de massa gravitacionais por meio de redes neurais artificiais e probabilidade**. Dissertação (Mestrado) - Escola de Engenharia de São Carlos da Universidade de São Paulo, 2015.

OLIVEIRA FILHO, A. G.; TOTOLA, L. B.; BICALHO, K. V.; HISATUGU, W. H. Prediction of Compression Index of Soft soils from the Brazilian Coast Using Artificial Neural Networks and Empirical Correlations. **SOILS & ROCKS**, v. 43, p. 109-121, 2020.

OLIVEIRA FILHO, A. G. **Análise comparativa da estimativa do índice de compressão de argilas por redes neurais artificiais e correlações empíricas**. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal do Espírito Santo. 2019.

OLIVEIRA, C. M. G. **Carta de risco de colapso de solos para a área urbana do Município de Ilha Solteira (SP)**. Ilha Solteira, 2002, 93p. Dissertação (Mestrado), Faculdade de Engenharia de Ilha Solteira – UNESP.

OLPHEN, H. V. As introduction to Clay colloid chemistry. **Wiley Interscience**, New York, 301p. 1963

PAIVA, S. C. **Estudo do Comportamento Geomecânico dos Solos Expansivos dos Municípios de Cabrobó, Paulista e Ipojuca – PE e de Suas misturas com cal**. Tese de doutorado em Engenharia Civil – UFPE, Recife, Pernambuco, 2016. 191 p.

PARK, H. L. (2011). Study for application of artificial neural networks in geotechnical problems. In: **Hui CLP (ed) Artificial neural networks-application**. In Tech, Croatia, pp 303–336. DOI:10.5772/2052.ISBN 978-953-307-188-6

PEDREGOSA, F. *et al.* **JMLR** 12, pp. 2825-2830, 2011.

PEREIRA, E. M. & PEJON, O. J. (1999). Estudo do Potencial Expansivo dos Sedimentos Argilosos da Formação Guabirotuba na Região do Alto Iguaçu-PR. **In.: 9º Congresso Brasileiro de Geologia de Engenharia**. São Pedro. ABGE.

PEREIRA, R. C.; AMORIM, S. F.; CAVALCANTI, E. J. N. Aplicação de Redes Neurais Artificiais e Geoprocessamento para Previsão da Susceptibilidade de Deslizamentos Superficiais com o Uso de Programas Gratuitos, Recife/PE. In: ANAIS DO XX COBRAMSEG, 2022, Campinas. **Anais [...]**. Campinas, Galoá, 2022.

PEREZ, J. A. *et al.* Artificial Intelligence and Robotics. **EPSRC UK-RAS Network**. 2017.

PTI - Post Tensioning Institute. Design and Construction of Post-tensioned Slabs-on-ground. **Post Tensioning Institute: Phoenix**, 1996.

QI, S.; VANAPALLI, S. K. Hydro-mechanical coupling effect on surficial layer stability of unsaturated expansive soil slopes. **Computers and Geotechnics**, v. 70, p. 68–82, 2015.

RANKKA, K.; ANDERSSON-SKOLD, Y.; HULTÉN C.; LARSSON, R.; LEROUX, V.; DAHLIN, T. (2004). Quick clay in Sweden. Report 65. **Swedish Geotechnical Institute**, Linköping.

REGINATTO, A. R. & FERRERO, J. C. Collapse potential of soils and soil-water chemistry. **In: International Conference on Soil Mechanics and Foundation Engineering**, Moscow, Proceedings, v.2.2, p. 177-183, 1973.

REICHARDT, K. (1985). Processos de Transferência no Sistema Solo - Planta - Atmosfera. **Fund. Cargill**, Campinas, 466p.

RIBEIRO, A. J. A.; SILVA, C. A. U.; BARROSO, S. H. A. Metodologia de baixo custo para mapeamento geotécnico aplicado à pavimentação. **TRANSPORTES**. Volume 26 | Número 2 | 2018. DOI:10.14295/transportes.v26i2.1491

RIPLEY, B. D. (1996). **Pattern recognition and neural networks**. Cambridge University Press, Cambridge.

- ROJAS, I. (1996). **Diseño e implementación de sistemas difusos**. PhD thesis, Depto. De Electronica, Fac. de Ciencias.
- ROMÃO P. A.; GONÇALVES, T. D.; SOUZA, N. M. Uso da cartografia geotécnica na análise de umidade dos solos. Solos não saturados no contexto geotécnico. – São Paulo: **Associação Brasileira de Mecânica dos Solos e Engenharia Geotécnica**, 2015.
- ROSENQUIST, I. T. (1953). Considerations on the sensitivity of Norwegian quick-clays. **Geotechnique** 3:195–200
- SALAHUDEEN, A. B.; SADEEQ, J. A.; BADAMASI, A.; ONYELOWE, K. C. (2020). Prediction of unconfined compressive strength of treated expansive clay using back-propagation artificial neural networks. **Nigerian Journal of Engineering** Vol, 27(1).
- SALEHI, T., SHOKRIAN, M., MODIRROUSTA, A. *et al.* Estimation of the collapse potential of loess soils in Golestan Province using neural networks and neuro-fuzzy systems. **Arab J Geosci** 8, 9557–9567 (2015). <https://doi.org/10.1007/s12517-015-1894-4>
- SANTOS, D. S.; MEDEIROS, C. E. B.; DRUMOND, M. A. M. M.; FERREIRA, S. R. M. Avaliação de Suscetibilidade à Ocorrência de Solos Expansivos e Colapsíveis no estado de Sergipe com base na Pedologia. In: ANAIS DO XX COBRAMSEG, 2022, Campinas. **Anais [...]**. Campinas, Galoá, 2022.
- SANTOS, M. D. D. **Problemas geotécnicos associados a solos expansivos em Lisboa**. Universidade Nova de Lisboa, 2017.
- SAYADI, A.; MONJEZI, M.; TALEBI, N.; KHANDELWAL, M. (2013). A comparative study on the application of various artificial neural networks to simultaneous prediction of rock fragmentation and backbreak. **J Rock Mech Geotech Eng** 5:318–324.
- SCHMIDHUBER, J. Deep Learning in Neural Networks: An Overview. **Neural Networks**, v. 61, p. 85-117, 2015.
- SCHMITZ, R. M. Can the diffuse double layer theory describe changes in hydraulic conductivity of compacted clays? **Geotechnical and Geological Engineering**, 24 (6) (2006), pp. 1835-1844
- SEED, H. B. WOODWARD Jr., R. J.; LUNDGREN, R. (1962). Prediction of Swelling Potential for Compacted Clays. **J. ASCE**, Soil Mechanics and Foundation Division, Vol. 88, No. SM-3, Part 1, pp. 53-87.
- SHAHIN, M. A.; JAKSA, M. B.; MAIER, H. R. (2008). State of the art of artificial neural networks in geotechnical engineering. **Electronic Journal of Geotechnical Engineering**, 8(1), 1-26.
- SHAHIN, M. A.; JAKSA, M. B.; MAIER, H. R. (2001). Artificial neural network applications in geotechnical engineering. **Aust Geomech** 36(1):49–62.
- SHAHRI, A. A. An Optimized Artificial Neural Network Structure to Predict Clay Sensitivity in a High Landslide Prone Area Using Piezocone Penetration Test (CPTu) Data:

A Case Study in Southwest of Sweden. **Geotech Geol Eng** 34, 745–758 (2016).
<https://doi.org/ez371.periodicos.capes.gov.br/10.1007/s10706-016-9976-y>

SHEELA, K. G.; DEEPA, S. N. Review on methods to fix number of hidden neurons in neural networks. **Mathematical Problems in Engineering**, v. 2013.

SILVA FILHO, F. C.; MOURA, A. S.; FUTAI, M. M. Uma solução para fundações em solos colapsíveis no Campus da UFC no Cariri. **In.: XV Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica**, Gramado, 2010. 5 p.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. **Artliber**, 2016.

SILVA, M. J. R.; FERREIRA, S. R. M. Cartas de suscetibilidade ao colapso e à expansão devido á inundação em solos do município de Petrolina-PE. **In.: V Congresso Brasileiro de Geotecnia Ambiental**, 2003, Porto Alegre. V Congresso Brasileiro de Geotecnia Ambiental, 2003. v. 1. p. 1-8.

SIMÕES DE OLIVEIRA, A. G.; JESUS, A. C.; MIRANDA, S. B. Estudo geológico geotécnico dos solos expansivos da região do Recôncavo Baiano. In: SIMPÓSIO BRASILEIRO DE JOVENS GEOTÉCNICOS, 2., 2006, Nova Friburgo-RJ. **Anais[...]**. Nova Friburgo: ABMS, 2006.

SINGH V.K.; SINGH, D.; SINGHA, T.N. Prediction of strength properties of some schistose rocks from petrographic properties using artificial neural networks. **Int. J. Rock Mech. Min. Sci.**, v. 38, p. 269-284, 2001.

SINHA, S. K.; WANG, M. C. (2008). Artificial neural network prediction models for soil compaction and permeability. **Geotech Eng J** 26(1):47–64.

SKEMPTON, A. W. The colloidal “Activity” of clays. **In.: INTERNATIONAL CONFERENCE ON SOIL MECHANICS AND FOUNDATION ENGINEERING**, 3., Zurich, 1953. Proceeding... Zurich, 1953. p.57-61.

SODIKOV, J. Cost estimation of highway projects in developing countries: artificial neural network approach. **Journal of the Eastern Asia Society for Transportation Studies**, v. 6, p. 1036–1047, 2005.

SOLHEIM, A.; BERG, K.; FORSBERG, C.F.; BRYN, P. (2005). The storegga slide complex: repetitive large scale sliding with similar cause and development. **Mar Pet Geol** 22:97–107.

SONMEZ, H.; ERCANOGLU, M.; KALENDER, A.; DAGDELENLER, G.; TUNUSLUOGLY, R. Predicting uniaxial compressive strength and deformation modulus of volcanic bimrocks considering engineering dimension. **Int. J. Rock Mech. Min. Sci.**, v. 86, p. 91-103, 2016.

SRIDHARAN, A.; CHOUDHURY, D. Computation of hydraulic conductivity of montmorillonitic clays by diffuse double layer theory. **International Journal of Geotechnical Engineering**, 2 (1) (2008), pp. 1-10.

SRIDHARAN, A.; NAGARAJ, H. B. Hydraulic conductivity of remoulded fine-grained soils versus index properties. **Geotechnical and Geological Engineering**, 23 (2005), p. 43.

STECK, J. E.; ROKHSAZ, K.; SHUE, S. P. Linear and neural network feedback for flight control decoupling. **IEEE Control Systems**, v. 16, n. 4, 1996, p. 22-30.

STEINBERG, M. Expansive soils and the geomembrane remedy. In: Geo-denver 2000. **ASCE**; 2000.

STONE, M. (1974). Cross-validators choice and assessment of statistical predictions. **Journal of Royal Statistical Society**, B 36, 111-147.

TAMIRU, H.; WAGARI, M.; TADESE, B. An integrated Artificial Intelligence and GIS spatial analyst tools for Delineation of Groundwater Potential Zones in complex terrain: Fincha Catchment, Abay Basi, Ethiopia. **Air, Soil and Water Research**. 2022;15. DOI: 10.1177/11786221211045972

TAVARES, C. M.; OLIVEIRA JUNIOR, A. I.; CALADO, A. C. A.; FERREIRA, S. R. M. Utilização de SIG aplicada aos estudos de solos expansivos e colapsíveis do Estado do Ceará. In: ANAIS DO XX COBRAMSEG, 2022, Campinas. **Anais [...]**. Campinas, Galoá, 2022.

TEIXEIRA, C. (2019). **SOLOS COLAPSÍVEIS: UM PROBLEMA PARA A ENGENHARIA DE FUNDAÇÕES**.

THORNTHWAITE, C. W.; MATHER, J. R. (1955). The Water Balance - Publications in Climatology. **N. Jersey: Centerton**, v. VIII, n. 1.

TORRANCE, J. K. (1983). Towards a general model of quick clay development. **Sedimentology** 30:547-555.

TOTOLA, L. B. **Aplicação de Redes Neurais Artificiais na Estimativa de Curvas de Retenção de Solos Brasileiros**. 2020. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal do Espírito Santo. 2020.

VAREJÃO-SILVA, M. A. (2003). Cartas das Isotermas Médias Anuais da Temperatura Média Diária, Isoietas Médias Anuais e Índices Hidricos Anuais para o estado de Pernambuco. Comunicação pessoal.

VARGAS, M. (1973). Structurally Unstable in Southern Brazil. **In: Proc. VII International Conference on Soil Mechanics and Foundation Engineering**. Moscow. Vol. 2, p. 239-256.

VARGAS, M.; GONÇALVES, H.H.; SANTOS, N.B.; Marinho, F.M.A. (1989). Expansive Soils in Brazil. **Supplementary Contributions by the Brazilian Society for Soil Mechanics**. 12th ICSMFE. Rio de Janeiro.

VASCONCELOS, R. P. R. (2001). Mapas de Suscetibilidade de Solos Colapsíveis e Expansivos do Estado de Pernambuco. Tese de Mestrado, UFPE, 93 p.

VELÁSQUEZ, R. M. G. 2006. **Seleção de variáveis de entrada para previsoeres neurais de carga**. Tese de doutorado, COPPE/UFRJ.

WALTAM, A. C. (1994). **Foundations of Engineering Geology**. Blackie Academic & Professional, London.

WU, M.; CONGRESS, S. S. C.; LIU, L.; CAI, G.; DUAN, W.; CHEN, R. (2021). Prediction of limit pressure and pressure meter modulus using artificial neural network analysis based on CPTU data. **Arabian Journal of Geosciences**, 14(1), 1-18.

YADAV, J. S.; TIWARI, S. K. The impact of end-of-life tires on the mechanical properties of fine-grained soil: a review. **Environment, development and sustainability: a multidisciplinary approach to the theory and practice of sustainable development**, Springer (2017).

YANG, Y.; ROSENBAUM, M.S. (2002). The artificial neural network as a tool for assessing geotechnical properties. **Geotech Eng J** 20(2):149–168.

YILMAZ, I.; YUKSEK, A. G. An example of artificial neural network (ANN) application for indirect estimation of rock parameters. **Rock Mechanics and Rock Engineering**, v. 41, n. 5, p. 781-795, 2008.

ZEGHAL, M.; KHOGALI, W. (2005) Predicting the resilient modulus of unbound granular materials by neural networks. **National Research Council Canada** (NRCC-47704). BCRA 2005, Trondheim, Norway, June 27-29, 2005, pp. 1-9.

ZHOU, Y.; WU, X. (1994). Use of neural networks in the analysis and interpretation of site investigation data. **Comput Geotech** 16:105–122.

ZUQUETTE, L.V.; NAKAZAWA, V.A. Cartas de geologia de engenharia. São Paulo: **ABGE**, 1998. p. 283-300.

APÊNDICE A - REDES NEURAIIS ARTIFICIAIS

Embora os primeiros trabalhos em RNA tenham sido publicados há mais de 50 anos, tal tema começou a ser fortemente pesquisado a partir do início dos anos 1990, após um longo período de descredito da Inteligência Artificial iniciado nos anos 70, quando não havia tecnologia de hardware suficiente o bastante para explorar todo o potencial das redes neurais, trazendo falhas para os sistemas de otimização - com algoritmos de aprendizado não diferenciáveis, isso impossibilitou a obtenção de muitos resultados e a resolução de problemas práticos. Até que a criação dos conceitos de múltiplas camadas nas redes neurais e redes neurais bidirecionais trouxe notoriedade novamente ao tema.

A RNA se trata de uma técnica, dentre as várias metodologias existentes, para a solução de problemas de inteligência artificial, baseada no ajuste de conjuntos de parâmetros (ajuste de pesos), tornando-a capaz de aprender, por meio de experiências (treinamento), e generalizar o comportamento de um determinado problema, ou seja, dar resposta a casos cujo resultado se desconhece previamente (BARROS, 2019).

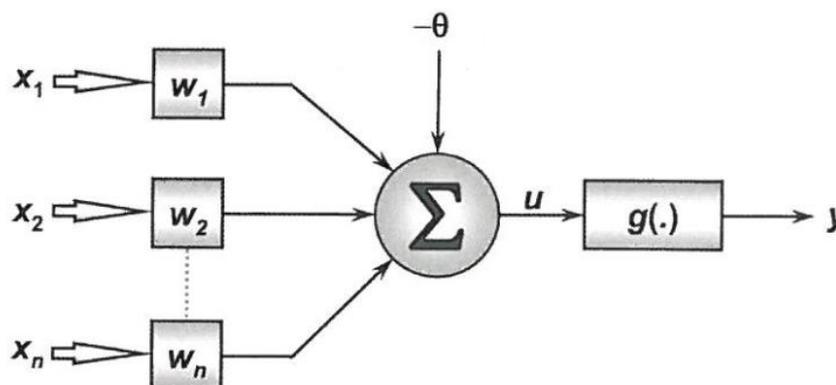
Como não são feitas suposições prévias sobre o modelo estatístico dos dados de entrada, nas RNA as curvas são ajustadas através de dados sem ter uma função predeterminada com parâmetros livres. Assim, por exemplo, em uma hipótese de estimativa de coesão, como no estudo de Khandelwal *et al.* (2018), não é necessário analisar o domínio para obter a forma da função de coesão antes de realizar a estimativa. Semelhante aos métodos paramétricos, as aproximações das Redes Neurais são baseadas em dados de casos passados.

Dos diversos setores do conhecimento humano, as redes neurais podem ser aplicadas a: reconhecimento automático de alvos, reconhecimento de caracteres, robótica, diagnóstico médico, sensoriamento remoto, processamento de sinais, biometria etc. (CABRAL, 2017). A rede neural é dinâmica, não é um programa onde o programador introduz e codifica a tática de solução do problema, ou seja, não possui uma sequência ordenada e definitiva de instruções (caso mais genérico, com treinamento não supervisionado); nem tão pouco que conduz o processo de resolução apenas a partir de um conhecimento explicitamente armazenado, deve-se aprender a tarefa que será executada, a partir da análise dos dados disponíveis (NOLA, 2015). O cálculo é processado através dos neurônios artificiais, que estão interligados para formar uma rede, sendo capazes de armazenar informações, reconhecer e classificar padrões, bem como fazer previsões (MARCY *et al.*, 2014).

A.1. – Neurônio Artificial

A função do neurônio artificial simples, A1, deve ser “coletar os sinais existentes de entrada, agregá-los de acordo com uma função operacional e produzir uma resposta, levando em consideração a sua função de ativação inerente” (SILVA *et al.*, 2016).

Figura A1 - Modelo de neurônio artificial simples.



Fonte: Silva *et al.* (2016).

O conjunto $\{X_1, X_2, \dots, X_n\}$ são os sinais de entrada advindos do meio externo e representam os valores assumidos pelas variáveis condicionantes de uma aplicação específica. Os pesos sinápticos são representados pelo conjunto $\{W_1, W_2, \dots, W_n\}$ e servem para ponderar as entradas da Rede. A relevância de cada uma das entradas $\{X_i\}$ do neurônio é então executada por meio de suas multiplicações pelos respectivos pesos sinápticos $\{W_i\}$, ponderando-se, portanto, todas as informações externas que chegam ao neurônio. Os pesos sinápticos são determinados inicialmente definidos de forma aleatória. A partir do início do processo, os pesos sofrem ajustes de acordo com regras que visam o aprimoramento da Rede Neural na execução de sua tarefa.

A soma ponderada das entradas $\{\Sigma\}$ se transforma na saída do corpo celular artificial que é o potencial de ativação, denotado por $\{u\}$, após a introdução do limiar de ativação / bias $\{\theta\}$ o resultado é obtido, pela diferença do valor produzido entre o combinador linear e o bias. A função de ativação $\{g\}$ tem como objetivo atribuir um padrão de comportamento do sinal de saída $\{y\}$ do neurônio, a depender do tipo de função utilizada. O valor final produzido para os dados X_1, X_2, \dots, X_n – no treinamento, depois será a resposta generalizada para um conjunto x_1, x_2, \dots, x_n para o qual não se tem resposta previa.

O valor final produzido pelo neurônio em relação a um determinado conjunto de sinais de entrada, pode ser também utilizado por outros neurônios que estão sequencialmente interligados. A síntese desses elementos e suas funções está descrita no Quadro A1.

Quadro A 1 - Descrição dos elementos que compõem um neurônio artificial.

Elementos Básicos de um Neurônio Artificial	
Sinais de entrada	São as variáveis assumidas para uma aplicação específica. Cada valor de entrada é um tipo de dado associado à amostra.
Pesos sinápticos	São valores associados a cada variável de entrada a fim de ponderá-las, permitindo quantificar as suas relevâncias. Os pesos se iniciam com valores aleatórios entre 0 e 1, e são ajustados durante o processo de ajuste de cálculo da rede.
Combinador linear	Tem a função de agrupar os produtos dos sinais de entrada e pesos sinápticos.
Limiar de ativação / Bias	Determina um valor que deve ser atingido pelo um combinador linear de forma a produzir um resultado pelo neurônio.
Potencial de ativação	É a diferença de valores entre o combinador linear e o limiar de ativação e determina o comportamento do neurônio (excitatório ou inibitório).
Função de ativação	Normaliza a saída do neurônio para valores que possam ser interpretados.
Sinal de saída	É o resultado do processo de cálculo. Pode ser o produto geral da análise, bem como ser utilizado por outros neurônios da rede.

Fonte: Modificado de Medeiros (2013).

As Equações (2) e (3) sintetizam o resultado produzido pelo neurônio artificial proposto por McCulloch e Pitts (1943). Onde X_i representa a i -ésima entrada do neurônio, W_i é o peso associado à entrada X_i , θ é o limiar de ativação (bias), n é o número de entradas, $g(u)$ é a função de ativação do neurônio e y é a saída dele. Detalhes sobre as diferentes funções de ativação serão apresentadas adiante.

$$u = \sum_{i=1}^n W_i \cdot X_i - \theta \quad \text{Equação (1)}$$

$$y = g(u) \quad \text{Equação (2)}$$

A maioria dos modelos de redes neurais possui alguma regra de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados. Em outras palavras, elas aprendem através de exemplos (padrões). Toda e qualquer rede pode ser descrita de forma matemática, tal como o exemplo das equações 1 e 2, de modo que, quanto mais neurônios e capas, mais complexa a equação fica.

A.2. - Funções de Ativação

As redes neurais podem se classificar quanto às funções de ativação usadas em seu interior, à arquitetura e quanto sua forma de aprendizado. As funções mais utilizadas são: degrau bipolar, logística, tangente hiperbólica, linear ou gaussiana, tais como apresentados na Tabela A1, tendo como objetivo principal a geração das saídas dos neurônios, fazendo uso dos valores dos dados de entrada e dos pesos.

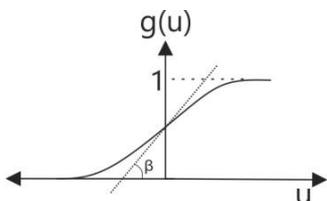
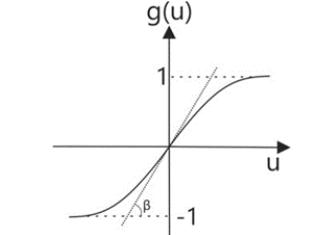
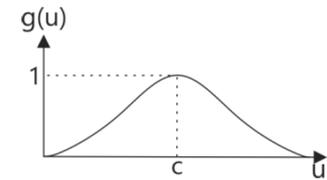
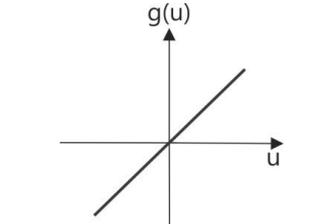
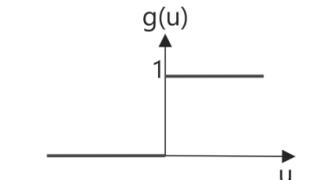
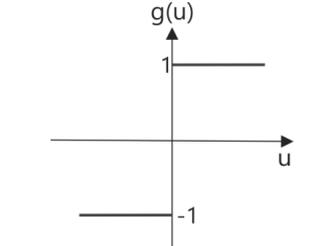
Entre as aplicações usuais das funções de ativação em RNA têm-se:

- Funções lineares, aplicadas a ajuste de funções logaritmo e simulação de classificadores gaussianos;
- Funções tangente-hiperbólicas, aplicadas ao reconhecimento e classificação de padrões;
- Funções gaussianas, aplicadas a previsão de séries temporais, *clustering*, identificação de sistemas complexos, aplicações financeiras e sistemas de informação;
- E, funções sigmóides, aplicadas a aproximação de funções, predição e otimização (SILVA *et al.*, 2016).

De acordo com Steck *et al.* (1996) as RNA podem ser aplicadas também no setor militar no processamento de sinais para identificação de alvos e análise de imagens, sistemas especialistas, processamento de sinais (imagens, sensores, voz, caracteres, visão, compressão de dados, filtragem de sinais), telecomunicações, manufatura, monitoramento de processos e robótica.

As funções de ativação, Tabela A1, podem ser parcialmente deriváveis, propiciando usar algoritmos de treinamento como o de retro propagação do erro quando possuem pontos cujas derivadas de primeira ordem são inexistentes e, totalmente deriváveis, quando as derivadas de segunda ordem são válidas e conhecidas em todo domínio da função (SILVA *et al.*, 2016).

Tabela A 1 - Funções de ativação em Redes Neurais Artificiais

Funções	Domínio	Representação Gráfica
Totalmente deriváveis	Função Logística $g(u) = \frac{1}{1+e^{-\beta \cdot u}}$	
	Função Tangente Hiperbólica $g(u) = \frac{1-e^{-\beta \cdot u}}{1+e^{-\beta \cdot u}}$	
	Função Gaussiana $g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}}$	
	Função Linear $g(u) = u$	
Parcialmente deriváveis	Função Degrau $g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases}$	
	Função Degrau Bipolar $g(u) = \begin{cases} 1, & \text{se } u > 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u < 0 \end{cases}$	

Fonte: Nola (2015).

Onde: $g(\cdot)$: Função de ativação; σ^2 : variância da distribuição;

u: Combinador linear;

c: parâmetro que define o centro da função

 β : Inclinação da curva em relação ao seu ponto de inflexão;

gaussiana.

A.3. - Escolha e Preparação das Variáveis

As variáveis de entrada são geralmente normalizadas - transformadas na mesma ordem de grandeza, de modo a tornar as redes neurais mais estáveis - evitando pesos discrepantes, habilitar taxas de aprendizado mais altas e reduzir o sobre ajuste (*overfitting*). O processo de conversão ocorre transformando um intervalo real de valores em um intervalo padrão, normalmente de -1 a +1 ou de 0 a 1, Equação 3 e 4.

$$y = \frac{x - \min}{\max - \min} \quad \text{Equação (3)}$$

$$y = \frac{x - \mu}{\sigma} \quad \text{Equação (4)}$$

Onde: y: representa o novo valor normalizado;

x: o valor atual;

μ e σ : a média e o desvio padrão da variável;

máx e min: os valores de máximo e mínimo, respectivamente.

Para Bishop (1995), uma das formas mais comuns de processamento consiste em um simples redimensionamento linear das variáveis de entrada. Isso geralmente é útil se diferentes variáveis tiverem valores típicos que diferem significativamente, como variáveis quantitativas e variáveis qualitativas. Dependendo das unidades em que cada um deles é expresso, eles podem ter valores que diferem em várias ordens de grandeza.

- Variáveis Quantitativas: são as características que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos cujas operações matemáticas fazem sentido. Podem ser contínuas ou discretas.
 - Variáveis discretas: características mensuráveis que podem assumir apenas um número finito ou infinito contável de valores e, assim, somente fazem sentido valores inteiros. Geralmente são o resultado de contagens.
 - Variáveis contínuas: características mensuráveis que assumem valores em uma escala contínua (na reta real), para as quais valores fracionais fazem sentido. Usualmente devem ser medidas através de algum instrumento.

- Variáveis Qualitativas (ou categóricas): são as características que não possuem valores quantitativos, mas, ao contrário, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos. Podem ser nominais ou ordinais.
 - Variáveis nominais: não existe ordenação dentre as categorias.
 - Variáveis ordinais: existe uma ordenação entre as categorias.

Identificar as variáveis de entrada mais relevantes é um pré-requisito importante para o sucesso da aplicação da RNA. Pois variáveis que não representam o problema podem estender o tempo de treinamento bem como interferir na obtenção da qualidade dos resultados (VELÁSQUEZ, 2006).

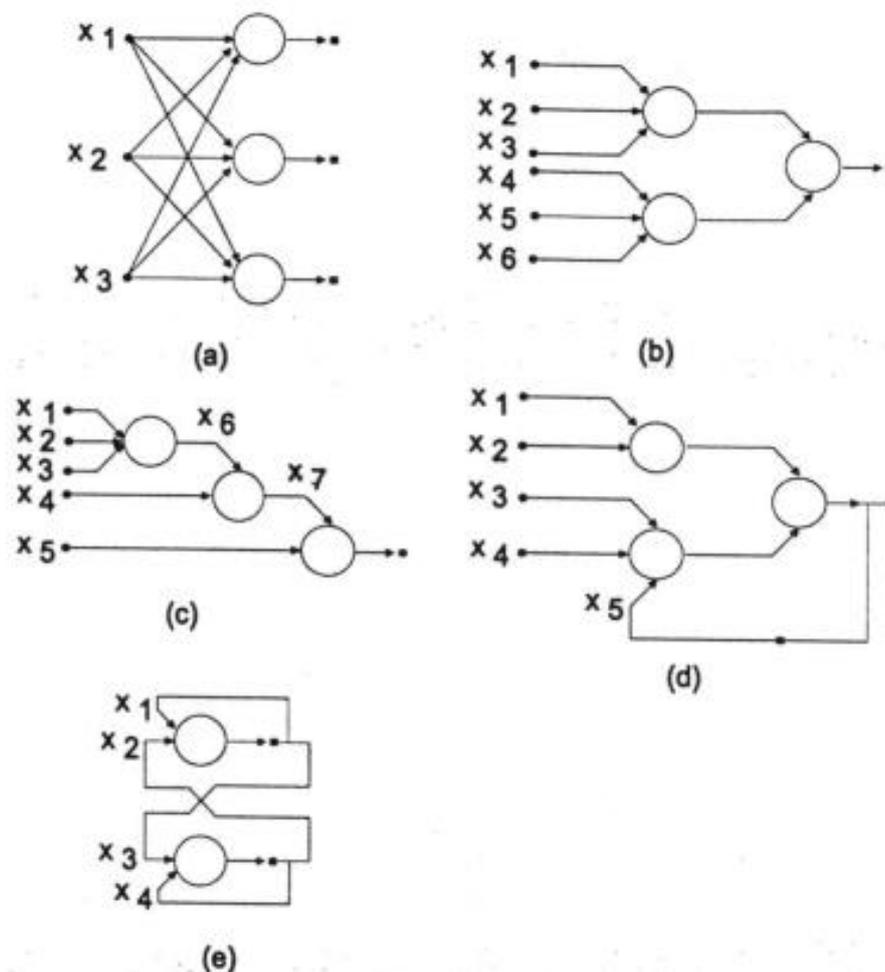
A técnica de Análise de Componentes Principais (PCA, em inglês *Principal Component Analysis*) é uma ferramenta da estatística multivariada, que tem como objetivo descrever a estrutura de variância e covariância de um conjunto de variáveis ou dimensões, por meio de combinações lineares dos membros desse conjunto. Aplicada sobre amostras de medições em um dado sistema, a PCA nos mostra como e com qual importância essas dimensões impactam na variação dos valores medidos, frequentemente explicitando relacionamentos ocultos entre elas.

Além disso, a PCA é uma ferramenta utilizada para mitigar redundâncias e correlações entre variáveis independentes e reduzir a dimensionalidade do conjunto de variáveis utilizadas na observação dos sistemas levando a modelos e consequentemente a redes mais simples e de melhor desempenho.

A.4. - Arquitetura da Rede

A arquitetura da rede corresponde às formas como os neurônios estão arranjados, uns em relação aos outros, onde a rede pode ser composta de uma ou mais camadas de neurônios, independente da função de ativação (Figura A2). Esses arranjos são estruturados de modo a direcionar as conexões sinápticas (BARROS, 2019). A definição da melhor arquitetura da RNA depende das propriedades dos dados, das características dos sistemas envolvidos, do que se espera obter nas saídas, do poder computacional etc. e, um dos fatores principais, da experiência do projetista.

Figura A2 - Exemplos de arquitetura de RNA



Fonte: Braga *et al.* (2000).

Quanto ao número de camadas, pode-se ter:

- Redes de camada única (Figura A2 a, e): só existe um nó entre qualquer entrada e qualquer saída de rede, apenas uma camada alimentada pela frente. As unidades de entrada são neurônios que recebem os sinais externos e, as de saída, os neurônios que apresentam os resultados processados pela rede. Por ter apenas uma camada, esse tipo de rede pode não apresentar um desempenho satisfatório para problemas mais complexos, geralmente são empregadas em problemas de classificação de padrões e filtragem;
- Redes de múltiplas camadas (Figura A2 b, c, d): se distingue da rede de única camada por dispor de camadas intermediárias, denominadas camadas ocultas, em que deve conter ao menos uma função de ativação não linear, apresentando melhores soluções para problemas não lineares (HAYKIN, 2001). A rede surgiu a partir dos agrupamentos

dos neurônios artificiais por camadas (entrada, camada oculta e saída), compostas por quantidades diferentes de neurônios, em função das variáveis envolvidas no problema, e da quantidade de saídas pretendidas. São empregadas em problemas de aproximação de funções, classificação de padrões, identificação de sistemas, otimização, robótica e controle de processos.

Os neurônios podem ter conexões do tipo:

- *Feedforward*, ou acíclica: a saída de um neurônio na i -ésima camada da rede não pode ser usada como entrada em camadas de índice menor ou igual a i (Figura A2 a, b, c)
- *Feedback*, ou cíclica: a saída de algum neurônio na i -ésima camada da rede é usada como entrada em camadas de índice menor ou igual a i (Figura A2 d, e).
 - Na Rede com Recorrência, as saídas dos neurônios são realimentadas como sinais de entrada para outros neurônios (Figura A2 d). São empregadas em problemas de agrupamento, reconhecimento de padrões, otimização de sistemas etc. Redes cuja saída final (única) é ligada às entradas comportam-se como autônomos reconhecedores de cadeias, onde a saída que é realimentada fornece o estado do autônomo.
 - A Rede com Recorrência Auto Associativa (Figura A2 e), é conhecida por Rede de *Hopfield* (1982), quando todas as ligações são cíclicas, a rede é denominada auto associativa. Estas redes associam um padrão de entrada com ele mesmo. A rede não possui entradas externas, e sua operação se dá em função da dinâmica de mudança de estados de neurônios que operam de forma auto associativa. Empregadas em previsões de séries temporais, otimização, identificação de sistemas e controle de processos, e úteis para recuperação ou “regeneração” de um padrão de entrada.

Embora o projeto das camadas de entrada e saída de uma rede neural seja frequentemente direto, pode haver bastante variação nas equações que dão ordem de grandeza ao número de neurônios nas camadas ocultas. Em particular, não é possível resumir esse processo com regras simples. Os mais variados *designs* ajudam as pessoas a obter o comportamento que querem de suas redes.

A.5. - Tipo de Aprendizado

Uma das principais características das RNA é a capacidade de aprender acerca de seu ambiente e assim melhorar o seu desempenho, através de um processo iterativo de ponderações aplicadas aos pesos sinápticos da rede, denominado de treinamento (HAYKIN, 2004).

O tipo de aprendizado é a maneira como a rede é treinada. Existem vários algoritmos de treinamento, e estes são divididos em duas principais classes, aprendizado supervisionado e não supervisionado.

O modelo de aprendizado supervisionado de RNA, faz uso de um bloco para treinar a RNA. Nesse modelo o estado do sistema é representado por um vetor que é aplicado à entrada da rede. Dentre os principais algoritmos de aprendizado supervisionado, um dos mais importantes e utilizados é o de retro propagação de erro - *backpropagation*. A retro propagação de erro busca minimizar a função de erro no espaço dos pesos, usando algumas metodologias, como o método do gradiente descendente. O método calcula o gradiente da função de erro em cada etapa de iteração, portanto, a continuidade e a diferenciação da função de erro devem ser garantidas. Uma vez que isso depende da função composta produzida pelos *perceptrons* interconectados, as funções de ativação devem manter as características (ROJAS, 1996).

O número de entradas e saídas não é restrito, o que é uma vantagem das RNA. Então, define-se o número de camadas ocultas e o número de neurônios ocultos. Um tempo considerável deve ser gasto na determinação do número desses neurônios, que é uma das desvantagens dessa técnica, pois exige alguns processos de tentativa e erro (KIM, AN e KANG, 2004). Entretanto, o suporte de programas como o *Neural Designer* possibilitem a indicação do melhor modelo de arquitetura, com número de camadas e de neurônios por camada, segundo algoritmos genéticos.

A quantidade de neurônios pode variar de acordo com o problema a ser resolvido. Alguns autores sugerem uma formulação para o número ideal de neurônios, conforme compilado por Sheela e Deepa (2013) de acordo com a Tabela A2.

Tabela A 2 - Estimativa do número de neurônios ideal na camada oculta.

Pesquisa	Método	Ano	Número de neurônios na camada oculta
1	Li et al	1995	$N_h = (\sqrt{1 + 8n} - 1)/2$
2	Tamura e Tateish	1997	$N_h = n - 1$
3	Fujita	1998	$N_h = K \log \ P_c Z\ / \log S$
4	Zhang et. al	2003	$N_h = 2^n / n + 1$
5	Jinchuan e Xinzhe	2008	$N_h = (N_{in} + \sqrt{N_p}) / L$
6	Chen e Xu	2008	$N_h = C_f (n / d \log n)^{0,5}$
7	Shibata e Ikeda	2009	$N_h = \sqrt{N_i N_o}$
8	Hunter	2012	$N_h = 2^n - 1$
9	Sheela e Deepa	2013	$N_h = (4n^2 + 3) / (n^2 - 8)$

Fonte: Sheela e Deepa (2013).

Onde: N_h : número de neurônios na camada oculta;

N_{in} , n : número de variáveis de entrada;

N_p : número de projetos;

N_o : número de saídas e;

L : número de camadas ocultas obtidas pela simulação de cada uma das configurações.

Já o processo de aprendizagem não supervisionado, ocorre quando a rede não dispõe das informações sobre as respostas corretas e incorretas, aprendendo com seu próprio agrupamento de padrões, usado para detectar similaridades e anomalias, como por exemplo: recomendação de filmes ou músicas.

Assim, quando a rede neural, em geral, alcança uma solução generalizada para uma determinada classe de problemas entende-se que ocorreu o aprendizado, o que a torna capaz de desempenhar algumas tarefas, tais como a detecção e o diagnóstico de padrões.

A.6. - Modelagem da Rede Supervisionada Tipo MLP com BP

O processo de aprendizagem supervisionada tem a função de modificar os pesos das conexões entre os neurônios (pesos sinápticos) da rede de uma forma ordenada para minimizar os erros entre a saída obtida (y) e a saída esperada (\hat{y}) (HAYKIN, 2001). A partir da informação dos pares de entrada e saídas, contidas em um banco de dados, a RNA deve “aprender” e generalizar o conhecimento contido nas amostras, de forma a responder adequadamente quando entradas não presentes no conjunto de treinamento forem apresentadas. O aprendizado da rede fica então gravado como se sua memória se tratasse, no conjunto final, de pesos que as suas conexões adquiriram.

A generalização de informações aprendidas através de exemplos é a capacidade da rede ‘aprender’ através de um conjunto reduzido de exemplos, cuja resposta é conhecida, dentro do intervalo treinado (interpolação), e posteriormente dar respostas corretas para dados fora do domínio aprendido (extrapolação), é uma demonstração de que a capacidade das RNA vai muito além de simplesmente mapear relações de entrada e saída. As RNA são capazes de extrair informações não apresentadas de forma explícita através dos exemplos. Além disso, as RNA são capazes de atuar como mapeadores universais de funções multivariáveis, com custo computacional que cresce apenas linearmente com o número de variáveis. Outra característica importante é a capacidade de auto-organização e de processamento temporal, que faz da RNA uma ferramenta computacional extremamente poderosa e atrativa para a solução de problemas complexos (BRAGA *et al.*, 2000).

O processo começa com a inserção de variáveis de entrada normalizadas, que servem para treinar a rede, essas informações são repassadas entre neurônios e calculadas a partir de pesos sinápticos e bias gerados aleatoriamente. Ao final, a rede irá prever os dados de saída, calcular o resultado e compará-lo com os dados de treinamento experimental para cada conjunto de variáveis de entrada. A partir desse processo inicial, haverá um erro sobre a diferença entre o valor de saída do neurônio de treinamento e o valor desejado do neurônio de saída, para cada iteração durante o treinamento. Para minimizar o erro, através do cálculo do gradiente iterativo correspondente ao algoritmo de retro propagação, peso sináptico e bias podem ser alterados. Este processo será repetido até que o critério de parada seja alcançado.

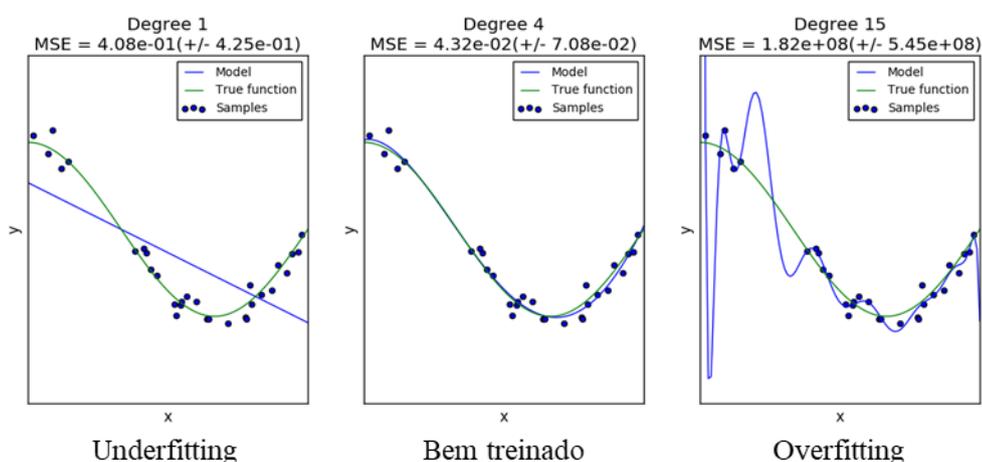
O ajuste do modelo é extremamente importante para que se possa acompanhar e compreender a causa de precisões de modelo insatisfatória, indicando quais medidas corretivas devem ser tomadas. Em relação aos dados utilizados nos modelos, o usual é que parte desses dados sejam

usados para o treinamento da rede e outra parte seja reservado somente para as etapas de seleção e teste, não tendo contato com o treinamento. Dessa forma, é possível avaliar se o modelo neuronal está fazendo *underfitting* ou *overfitting* dos dados de treinamento, Figura A3.

Em um cenário de *underfitting*, o modelo está fazendo o sub ajuste dos dados de treinamento quando o modelo não consegue capturar o relacionamento entre as variáveis de entrada e os valores de destino. O desempenho insatisfatório nos dados de treinamento pode ocorrer porque o modelo é muito simples (as variáveis de entrada não são suficientemente expressivas) para descrever o valor de destino. Para que se possa melhorar o desempenho, aumentando a flexibilidade do modelo, pode ser adicionado novas variáveis específicas de domínio e mais produtos cartesianos de variáveis; alterar os tipos de processamento das variáveis usadas, e reduzir o volume de regularização usado.

Já um cenário de *overfitting* ocorre quando, nos dados de treinamento, o modelo tem um desempenho excelente, porém quando utilizado os dados de teste o resultado é ruim. Pode-se entender que, neste caso, o modelo aprendeu tão bem as relações existentes no treinamento, que acabou apenas decorando o que deveria ser feito, e ao receber as informações das variáveis predictoras nos dados de teste, o modelo tenta aplicar as mesmas regras decoradas, porém com dados diferentes esta regra não tem validade, e o desempenho é afetado. Neste cenário, o modelo treinado tem sua capacidade de generalização comprometida. Devido a essa possibilidade, é interessante que a parada seja antecipada (*early stopping*), definindo durante o treinamento da rede o mínimo de Erro Quadrático Médio (MSW, em inglês *Mean Square Error*) das amostras usadas para o teste (HAYKIN, 2008; SCHMIDHUBER, 2015).

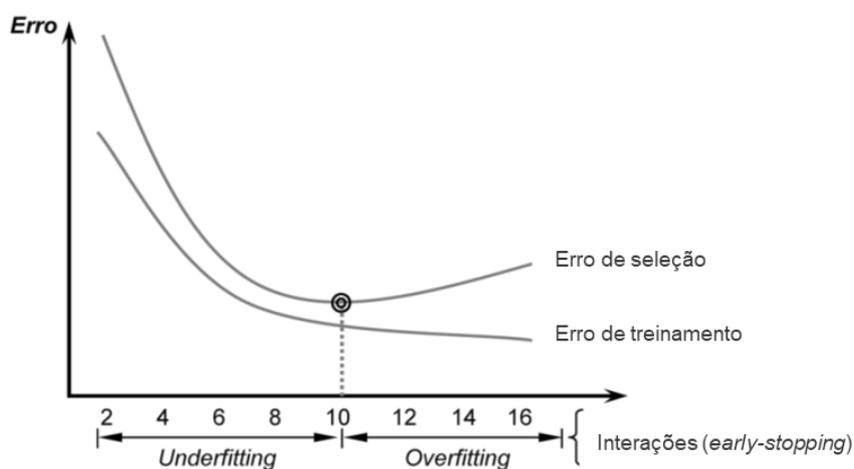
Figura A3 - Aprendizado de máquina – Underfitting, Bem treinado e Overfitting. Onde a linha verde é a função real, os pontos são o conjunto de treinamento com falhas e a linha azul é o modelo gerado



Usando a função de desempenho MSW, a rede obterá pequenos valores de pesos e polarização que causam uma resposta suave e menos propensa a sobreposições, aplicando as técnicas de regularização e parada antecipada. Em suma, esse processo deverá superar o *underfitting* e evitar a condição de *overfitting*, e o método *early stopping*, quando utilizado, deverá ser individualmente aplicado para cada arquitetura da rede candidata (Figura A4).

O acompanhamento do gráfico de erro calculado para conjuntos de treinamento e seleção durante o aprendizado de rede é útil, entre outras coisas, para observar o ponto a partir do qual ocorre sobre ajuste dos dados, e para avaliar a qualidade da divisão dos grupos. Se o erro no conjunto de seleção encontrar um mínimo em uma iteração significativamente anterior do que o erro no conjunto de treinamento, isso pode indicar uma divisão de dados ruim.

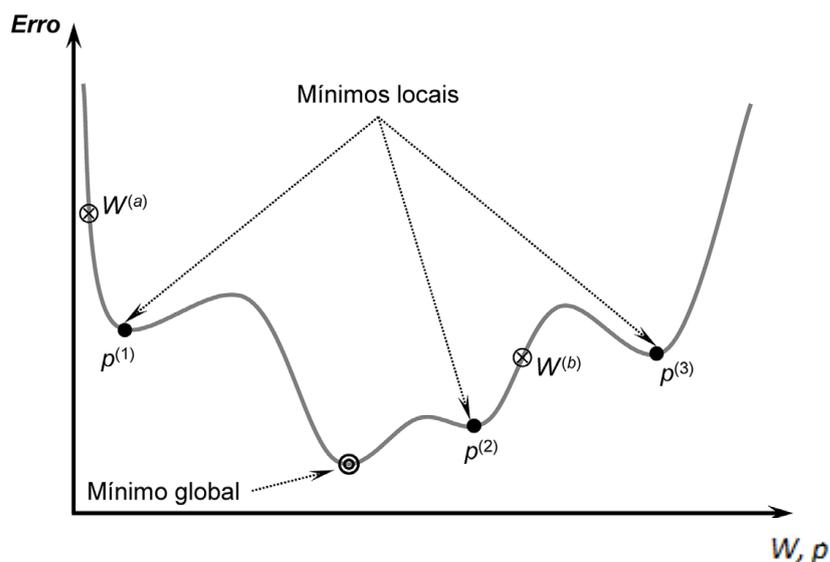
Figura A4 - Rastreamento de erros no conjunto de seleção para interromper o processo de treinamento



Fonte: Silva *et al.* (2016).

A fim de evitar a convergência da rede para pontos de mínimos locais (p) inapropriados, um dos procedimentos práticos a ser adotado consiste em executar o treinamento de cada arquitetura de rede mais de uma vez, com diferentes matrizes de pesos iniciais (geradas aleatoriamente). Assim, dependendo de suas posições espaciais, a rede poderia convergir para pontos de mínimos locais (p) (diferença entre duas soluções em cada etapa), ou até mesmo globais (W) (erro mínimo em qualquer etapa), que possibilitariam uma melhor representação do comportamento do processo (SILVA *et al.*, 2016).

Figura A5 - Pontos de mínimos locais associados à função erro



Fonte: Silva *et al.* (2016).

Após essa etapa o treinamento pode ser interrompido logo que um número fixo de registros de treinamento for apresentado, quando um erro de treinamento atinge um valor suficientemente pequeno ou quando não ocorre nenhuma alteração ligeira no erro de treinamento. No entanto, as técnicas de critérios de parada podem levar o modelo a parar prematuramente ou durante o treinamento. Tais problemas são superados com o uso da técnica de validação cruzada (STONE, 1974).

A técnica da validação cruzada exige que os dados sejam divididos em três conjuntos distintos, a saber, treinamento, seleção e teste/validação, dos quais o conjunto de treinamento é o maior conjunto usado pela rede neural para identificar os padrões presentes nos dados. O objetivo do treinamento é encontrar o conjunto de pesos entre os neurônios que determinem a função global mínima de erro.

A.7. - Testes de Desempenho

Após o treinamento e seleção da melhor arquitetura de rede, inicia-se a fase de testes, em que o objetivo é avaliar o desempenho da rede selecionada, ou seja, a precisão da predição da rede, a qual é apresentado o terceiro conjunto de dados (não utilizados no treinamento e seleção) para avaliar o comportamento da rede já treinada, tais resultados, monitorados nessa fase, indicam a capacidade de generalização desses padrões. Dentre os métodos de teste para redes de classificação tem-se:

- **Matriz de confusão:** Nesta avaliação as colunas representam as classes de dados observados no conjunto de dados (resultados medidos em campo) e as linhas representam as saídas correspondentes a previsão da rede neural, Quadro A2.

Quadro A 2 - Matriz de confusão de duas classes

		Condição Observada	
		Condição positiva	Condição negativa
Condição Prevista	População total		
	Condição prevista positiva	Verdadeiro positivo (TP)	Falso positivo (FP)
	Condição prevista negativa	Falso negativo (FN)	Verdadeiro negativo (TN)

Fonte: A Autora (2022).

- Verdadeiro positivo (TP, em inglês *True Positive*) - a amostra de solo avaliada em campo foi observada como expansiva e prevista pela RNA como expansiva.
- Verdadeiro negativo (TN, em inglês *True Negative*) - a amostra de solo avaliada em campo foi observada como colapsível e prevista pela RNA como colapsível.
- Falso positivo (FP, em inglês *False Positive*) - a amostra de solo avaliada em campo foi observada como colapsível e prevista pela RNA como solo expansivo.
- Falso negativo (FN, em inglês *False Negative*) - a amostra de solo avaliada em campo foi observada como expansivo e prevista pela RNA como solo colapsível.

Observe que o número de classificações corretas é o número de verdadeiros positivos mais o número de verdadeiros negativos, TP + TN; e o número de erros é o número de falsos positivos mais o número de falsos negativos, FP + FN.

- **Métricas de avaliação binária:** Ao obter a matriz de confusão e ser feita a contagem de todos os termos, é possível calcular métricas de avaliação para a classificação, através de um conjunto de parâmetros padrão para testar o desempenho de um modelo de classificação, no caso em questão: uma rede neural. Estes são os parâmetros derivados da matriz de confusão:

- Acurácia: indica uma performance geral do modelo. Dentre todas as classificações, quantas o modelo classificou corretamente;

$$\text{Acurácia} = \frac{(TP + TN)}{(P + N)}$$

Onde: P: Total de amostras positivas (expansivas);

N: Total de amostras negativas (colapsíveis).

Obs.: Essa disposição de amostras ocorre em virtude de o banco de dados ser constituído apenas por amostras de solos colapsíveis e expansivas.

- **Precisão:** dentre todas as classificações de classe Positivo que o modelo fez, quantas estão corretas; pode ser usada em uma situação em que os Falsos Positivos são considerados mais prejudiciais que os Falsos Negativos.

$$\text{Precisão} = \frac{TP}{(TP + FP)}$$

- **Sensibilidade/Recall:** é definida como a habilidade do modelo encontrar as respostas positivas. Pode ser usada em uma situação em que os Falsos Negativos são considerados mais prejudiciais que os Falsos Positivos.

$$\text{Sensibilidade} = \frac{TP}{(TP + FN)}$$

- **Especificidade:** proporção de verdadeiros negativos, a capacidade do sistema em prever corretamente a ausência da condição para casos que realmente não a têm.

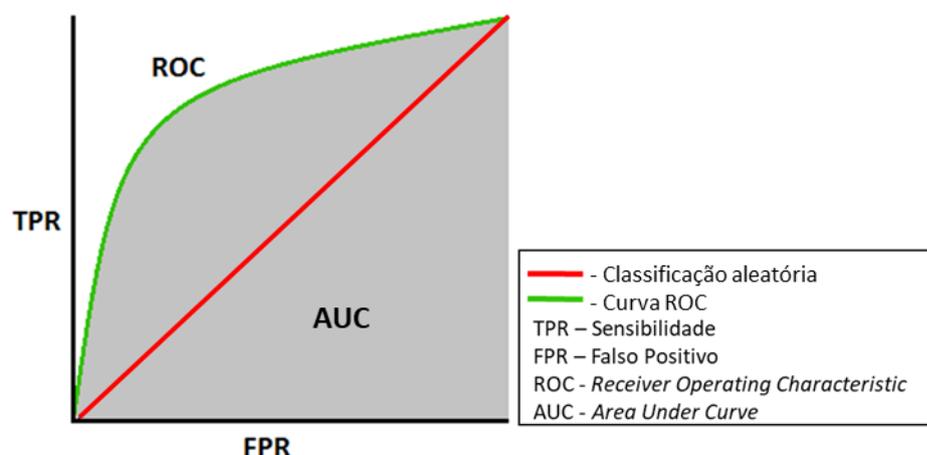
$$\text{Especificidade} = \frac{TN}{(TN + FP)}$$

- **Curva ROC:** Uma alternativa mais abrangente para a avaliação da eficiência da rede neural é a partir do uso de gráficos, que permitem uma melhor visualização da multidimensionalidade do problema. Um método padrão para testar classificadores binários, mesmo que modelados através de uma rede neural.

O gráfico ROC é baseado na probabilidade de detecção, ou taxa de verdadeiros positivos (TPR) ou 1 - especificidade no eixo Y, e na probabilidade de falsos alarmes, ou taxa de falsos positivos (FPR) ou sensibilidade no eixo X, Figura A6.

Um classificador aleatório, representado pela linha de base (diagonal a 45°), tem uma área abaixo da curva (AUC, em inglês *Area Under the Curve*) de 0,5. Quanto mais próximo de 1,00, for essa área e mais próxima a curva estiver do ponto (0, 1), melhor o classificador.

Figura A6 - Exemplo de curva ROC para um classificador.

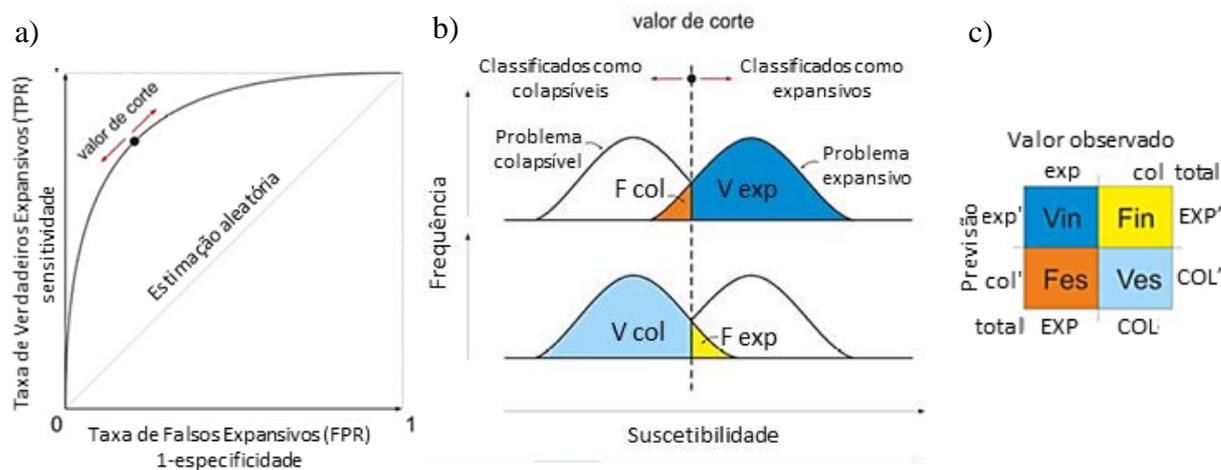


Fonte: A Autora (2022).

Na representação da Curva ROC da Figura A7-a), observamos que cada ponto nela define uma matriz de confusão e conseqüentemente uma relação entre os parâmetros TPR e FPR, se tomarmos um exemplo para avaliação de solos classificados com valor positivo para casos expansivos e negativo para casos colapsíveis, a terminologia TP, FP, TN e FN pode ser substituída respectivamente por Verdadeiro Expansivo, Falso Expansivo, Verdadeiro Colapsível e Falso Colapsível e as respectivas regiões do histograma de frequência das classes de colapsíveis e expansivos. Assim, o modelo aumentará em desempenho à medida que a proporção de verdadeiros expansivos aumenta mais rápido do que a proporção de falsos expansivos.

A curva ROC também é empregada para definir uma regra de decisão baseada em buscar um ponto de corte (Figura A7-b) que resume o aprendizado da rede em uma resposta dicotômica, de forma que amostras com mensurações maiores que o ponto de corte são classificadas como Verdadeiras Expansivas, enquanto amostras com mensurações menores que o ponto de corte são classificadas como Verdadeiras Colapsíveis (Figura A7-c).

Figura A7 - Curva ROC (a), histograma de frequência de célula estável e instável (b), e matriz de confusão para um ponto na curva ROC (c).



Fonte: Amorim, 2012.

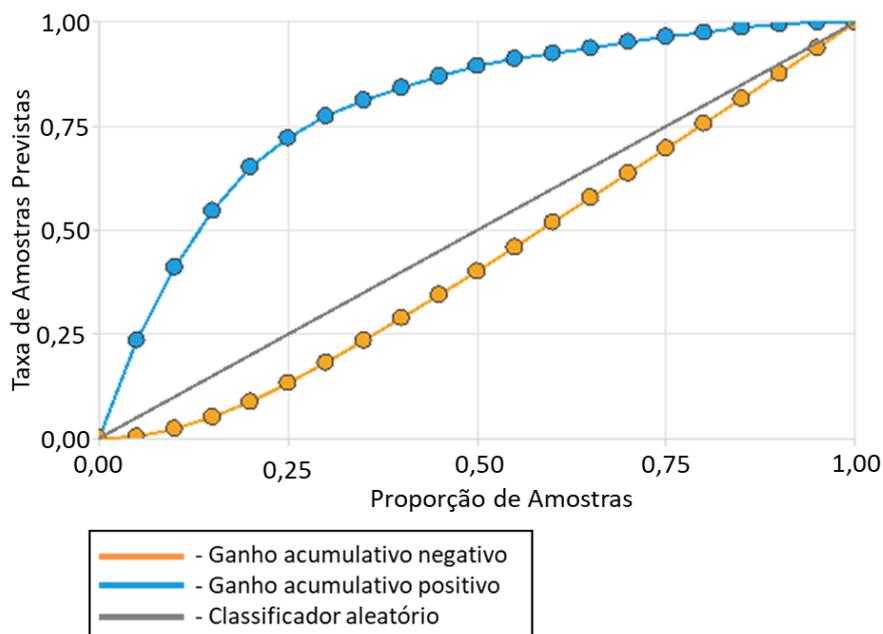
Em suma, a primeira vantagem da análise do classificador a partir da área da curva ROC é poder fazer uma análise independentemente de certas condições, tais como o ponto de corte, as dificuldades relacionadas às classificações errôneas e à distribuição a priori das classes. Essa análise é realizada visualizando-se o comportamento entre taxa de verdadeiro positivo e taxa de falso positivo em um gráfico bidimensional. A segunda vantagem é que a análise ROC pode ser utilizada para a calibração e ajuste de modelos de classificação, quando necessário.

- **Ganho acumulado:** A análise de ganho acumulado é um auxílio visual que representa a vantagem de usar um modelo preditivo em oposição à aleatoriedade, Figura A8. Quanto menor a proporção de amostras necessárias para alcançar uma maior divergência entre a identificação de classes diferentes, melhor; representando uma velocidade de classificação da rede.

O ganho acumulado positivo mostra no eixo y a porcentagem de amostras expansivas previstas em relação à população que está sendo classificada nessa etapa (teste), que é representada no eixo x.

O ganho acumulado negativo mostra a porcentagem de amostras colapsíveis encontradas em relação à população destinada a etapa de teste. A linha de base representa os resultados que seriam obtidos sem o uso de um modelo de previsão, ou com uma classificação aleatória.

Figura A8 - Exemplo de gráfico de ganho acumulativo.



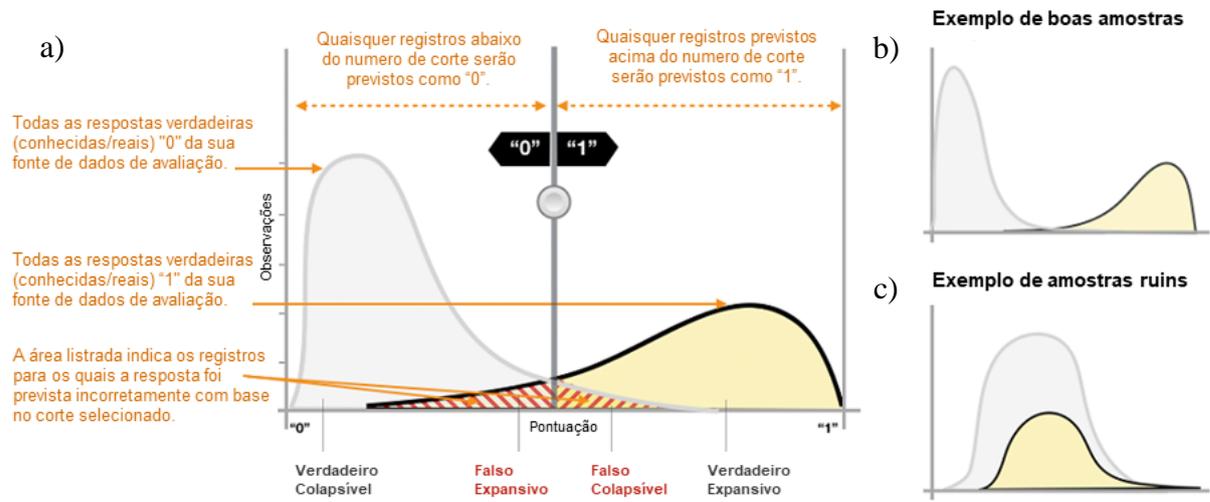
Fonte: A Autora (2022).

O melhor desempenho do modelo é a diferença máxima entre o ganho acumulativo positivo e o ganho acumulativo negativo, ou seja, é o ponto em que a porcentagem de amostras expansivas classificadas é maximizada e a porcentagem de amostras colapsíveis classificadas é minimizada; nesse exemplo o desempenho máximo de previsão acontece na proporção de 25% das amostras.

- **Histogramas Populacionais:** Também conhecido como distribuição de frequências ou diagrama das frequências, é a representação gráfica, em colunas, de um conjunto de dados previamente tabulado e dividido em classes uniformes, Figura A9.

Um modelo classificador com bom desempenho fará a previsão das maiores pontuações até os positivos reais e das menores pontuações até os negativos reais. Um modelo perfeito (Figura A9-b) terá os dois histogramas em duas extremidades diferentes do eixo x, mostrando que todos os positivos reais receberam pontuações altas e os negativos reais receberam pontuações baixas. No entanto, os modelos cometem erros, e um gráfico comum (Figura A9-a) mostrará que os dois histogramas se sobrepõem em determinadas pontuações. Um modelo com desempenho extremamente baixo (Figura A9-c) não conseguirá fazer a distinção entre classes positivas e negativas, e as duas classes terão basicamente histogramas sobrepostos.

Figura A9 - Histogramas Populacionais



Fonte: Adaptado de Amorim (2012).

O bom desempenho no grupo de teste indica que a rede foi treinada com sucesso e que a generalização foi alcançada adequadamente. Dessa forma, pode ser aplicado a rede as informações do problema ao qual busca-se solução.

APÊNDICE B - RESUMO DO DESENVOLVIMENTO DAS RNA

Durante o desenvolvimento da pesquisa, inicialmente alguns modelos de redes neurais, com variáveis de entrada baseadas na literatura de geotecnia, foram treinados e passaram pelas análises de componentes principais, análises de correlação, características dos algoritmos e arquiteturas de redes de modo a se identificar quais as principais variáveis responsáveis pelos melhores resultados de previsão e generalização dos padrões. A redução da quantidade de variáveis de entrada sem prejudicar a qualidade da previsão da rede, foi um dos parâmetros estabelecidos para a pesquisa, a fim de facilitar e atrair um maior uso das redes neurais (fazendo a previsão a partir de variáveis rotineiras do estudo geotécnico) além de reduzir custos e tempo de investigação do solo por outros métodos.

Tais redes iniciais foram resumidas com a descrição de algumas características durante o desenvolvimento e desempenho nos testes (Quadro B1). Após os estudos, 3 redes são apresentadas com detalhes de desenvolvimento, testagem e aplicação de padrões nesse capítulo, pois obtiveram os melhores desempenhos atendendo aos parâmetros antes mencionados. A ordem de apresentação das redes, também é a ordem cronológica de expansão do banco de dados, o que possibilitou a avaliação de outros conjuntos de variáveis de entrada, bem como a aplicação dos padrões dessas redes a amostras desconhecidas em seus treinamentos.

Quadro B 1 - Resumo das redes não detalhadas durante o desenvolvimento da pesquisa.

Rede	Correlação entre Entradas e Saídas	Camadas <i>Perceptron</i>	Nº Neurônios em cada Camada	Parâmetros de Desempenho	Observações
1	% Argila (0,31), % Silte (0,12), % Areia (-0,63), LL (0,37), LP (0,39), Ia (0,17), pd (-0,13), w (-0,66), e (-0,18), Sr (0,02), Gs (-0,12), γ máx. (-0,67).	2	1ª camada: 1 2ª camada: 1	86,5% das amostras classificadas corretamente e 13,5% incorretamente. Sensibilidade: 87,5%, Especificidade: 85,7%. Área da curva ROC: 0,902.	As variáveis ausentes não foram consideradas nessa rede como nulas e sim como zeradas, isso alterou significativamente o aprendizado da RNA.
2	% Argila (0,32), LL (0,41), LP (0,39), Ia (0,20), pd (-0,11), w (0,10), Sr (0,04), Gs (-0,11).	3	1ª camada: 3 2ª camada: 5 3ª camada: 1	83,3% das amostras classificadas corretamente e 13,3% incorretamente. Sensibilidade: 75%, Especificidade: 100%. Área da curva ROC: 0,875.	As variáveis ausentes já caracterizadas de forma correta tiveram sua primeira redução diante das correlações apresentadas.
3	% Argila (0,24), LL (0,25), LP (0,24), Ia (0,23), pd (0,31), w (0,68).	2	1ª camada: 2 2ª camada: 1	69,2 % das amostras classificadas corretamente e 30,8% incorretamente. Área da curva ROC: 0,738	A segunda redução de variáveis ocorreu para que pudéssemos ter cada vez menos variáveis de entrada para a mesma saída.
4	% Areia (-0,71), % Argila (0,81), IP (0,81), Ia (0,42), Pressão de Expansão (1) e Expansão livre (1).	2	1ª camada: 1 2ª camada: 1	92,3% das amostras classificadas corretamente e 7,7% incorretamente. Sensibilidade: 87,5%, Especificidade: 100%. Área da curva ROC: 1,00.	Essa RNA serviu para a sugestão de duas redes futuras com variáveis de entrada: % areia, % argila, IP e IA e com saídas: pressão de expansão e expansão livre.
5	% Areia (-0,71), % Argila (0,81), IP (0,81), Ia (0,42).	2	1ª camada: 1 2ª camada: 1	76,4% das amostras classificadas corretamente e 23,6% incorretamente. Sensibilidade: 75%, Especificidade: 78%. Área da curva ROC: 0,819.	Uma rede com um número mínimo de variáveis de entrada, partindo de ensaios simples foi considerada boa, pelos resultados significativos.
6	% Areia (-0,58), % Argila (0,67), IP (0,67), Ia (0,38), Pedologia (0,47), Climatologia (0,62), Geologia (0,5).	2	1ª camada: 40 2ª camada: 2	87,5 % das amostras classificadas corretamente e 12,5% incorretamente. Sensibilidade: 100%, Especificidade: 83,3%. Área da curva ROC: 1,0.	A inserção de variáveis categóricas aperfeiçoou consideravelmente o aprendizado e classificação da RNA.

Onde: % **Argila** - Percentual do solo com granulometria menor que 0,002 mm; % **Silte** - Percentual do solo com granulometria entre 0,002 e 0,006 mm; % **Areia** - Percentual do solo com granulometria entre 0,006 e 2,0 mm; **LL** - Limite de Liquidez; **LP** - Limite de Plasticidade; **IP** - Índice de Plasticidade; **Ia** - Índice de Atividade; **pd** - Massa específica seca; **e** - Índice de Vazios; **Sr** - Grau de Saturação; γ máx. - Peso específico seco; **Gs** - Densidade Relativa e **w** - Umidade ótima.

Fonte: A Autora (2022).

APÊNDICE C – EQUAÇÕES MATEMÁTICAS DAS RNA

Um modelo matemático descreve as ideias sobre o comportamento de um sistema, de modo a simular o comportamento físico tal como foi percebido mentalmente pelo autor do modelo. Sua funcionalidade depende fundamentalmente do fenômeno que se descreve e da capacidade para simulá-lo. Assim, as equações possuem parâmetros determinados por meio de ensaios de laboratório e de campo.

Para a geomecânica, os modelos construtivos usados baseiam-se nos conceitos da teoria da elasticidade e da plasticidade do solo, e são calibrados a partir de informações experimentais, assim, definem o comportamento do solo usando um conjunto fixo de equações, cuja principal vantagem é a capacidade de generalização, além disso, as expressões matemáticas representadas pelas redes neurais podem ser utilizadas para importa-la em outro programa a partir de diferentes linguagens de programação ou sistemas de computador, no chamado modo de produção.

Equação Matemática para a RNA PE04

```
scaled_Argila = (Argila-30.7829)/20.7821;
scaled_Areia = 2*(Areia-3)/(94-3)-1;
scaled_IP = (IP-18.4687)/14.8603;
scaled_IA = (IA-0.662574)/0.43131;
y_1_1 = Logistic (-4.42867+ (scaled_Argila*11.8894)+ (scaled_Areia*11.6914)+ (scaled_IP*2.26939)+
(scaled_IA*4.45515));
y_1_2 = Logistic (2.30693+ (scaled_Argila*18.0472)+ (scaled_Areia*-10.3963)+ (scaled_IP*-3.93087)+
(scaled_IA*10.8198));
non_probabilistic_EXP = Logistic (-2.57523+ (y_1_1*7.80666)+ (y_1_2*3.10474));
EXP = probability(non_probabilistic_EXP);

logistic(x){
return 1/(1+exp(-x))
}

probability(x){
if x < 0
return 0
else if x > 1
return 1
else
return x
}
```

Equação Matemática para a RNA PE07

```

scaled_Argila = (Argila-31.3102)/20.7336;
scaled_Areia = 2*(Areia-3)/(94-3)-1;
scaled_IP = (IP-17.2545)/15.0367;
scaled_IA = (IA-0.661429)/0.422997;
scaled_BORBOREMA = (BORBOREMA-0.390805)/0.490759;
scaled_COBCENOZOICA = (COBCENOZOICA-0.195402)/0.398809;
scaled_SFRANCISCO = (SFRANCISCO-0.241379)/0.4304;
scaled_MANTIQUEIRA = (MANTIQUEIRA-0.172414)/0.37993;
scaled_LATOSSOLO = (LATOSSOLO-0.195402)/0.398809;
scaled_ARGISSOLO = (ARGISSOLO-0.206897)/0.407429;
scaled_NEOSSOLO = (NEOSSOLO-0.45977)/0.501268;
scaled_PLANOSSOLO = (PLANOSSOLO-0.045977)/0.210649;
scaled_LUVISSOLO = (LUVISSOLO-0.091954)/0.290636;
scaled_UMIDO = (UMIDO-0.241379)/0.4304;
scaled_SUSECO = (SUSECO-0.103448)/0.306309;
scaled_ARIDO = (ARIDO-0.45977)/0.501268;
scaled_SUUMIDO = (SUUMIDO-0.0804598)/0.273581;
scaled_SEMIARIDO = (SEMIARIDO-0.114943)/0.320802;
y_1_1 = Logistic (1.21415+ (scaled_Argila*11.9273)+ (scaled_Areia*6.41713)+
(scaled_IP*2.01725)+ (scaled_IA*3.65757)+ (scaled_BORBOREMA*0.693392)+
(scaled_COBCENOZOICA*-0.674895)+ (scaled_SFRANCISCO*-0.71387)+
(scaled_MANTIQUEIRA*0.850126)+ (scaled_LATOSSOLO*1.83178)+
(scaled_ARGISSOLO*-1.02669)+ (scaled_NEOSSOLO*-0.60598)+
(scaled_PLANOSSOLO*-0.274351)+ (scaled_LUVISSOLO*0.264096)+
(scaled_UMIDO*-0.944913)+ (scaled_SUSECO*1.99221)+
(scaled_ARIDO*0.586873)+ (scaled_SUUMIDO*-1.39018)+ (scaled_SEMIARIDO*-
0.888008));
non_probabilistic_EXP = Logistic (-2.84224+ (y_1_1*10.8581));
EXP = probability(non_probabilistic_EXP);

logistic(x){
return 1/(1+exp(-x))
}

probability(x){
if x < 0
return 0
else if x > 1
return 1
else
return x
}

```

Equação Matemática para a RNA BR03

```

scaled_Argissolo = (Argissolo-0.247588)/0.432307;
scaled_Neossolo = (Neossolo-0.212219)/0.409538;
scaled_Planossolo = (Planossolo-0.0257235)/0.158564;
scaled_Luvisso = (Luvisso-0.0643087)/0.245697;
scaled_Espodossolo = (Espodossolo-0.0192926)/0.137773;
scaled_Latossolo = (Latossolo-0.350482)/0.47789;
scaled_Vertissolo = (Vertissolo-0.0289389)/0.167905;
scaled_Cambissolo = (Cambissolo-0.0321543)/0.176694;
scaled_Plintossolo = (Plintossolo-0.00321543)/0.0567048;
scaled_Nitossolo = (Nitossolo-0.00321543)/0.0567048;
scaled_Organossolo = (Organossolo-0.00643087)/0.0800632;
scaled_Dunas = (Dunas-0.00321543)/0.0567048;
scaled_Gleissolo = (Gleissolo-0.00321543)/0.0567048;
scaled_umido = (umido-0.492308)/0.500583;
scaled_semiarido = (semiarido-0.0948718)/0.293414;
scaled_arido = (arido-0.117949)/0.322962;
scaled_subumolhado = (subumolhado-0.15641)/0.36371;
scaled_subuseco = (subuseco-0.138462)/0.345827;
scaled_Costeira_MargemContinental = (Costeira_MargemContinental-0.124682)/0.330779;
scaled_CobCenozoica = (CobCenozoica-0.165394)/0.37201;
scaled_Borborema = (Borborema-0.127226)/0.333651;
scaled_Reconcavo_Tucano_Jatoba = (Reconcavo_Tucano_Jatoba-0.104326)/0.306072;
scaled_SaoFrancisco = (SaoFrancisco-0.195929)/0.39742;
scaled_Parnaiba = (Parnaiba-0.00508906)/0.0712466;
scaled_Mantiqueira = (Mantiqueira-0.0712468)/0.257565;
scaled_Tocantins = (Tocantins-0.0381679)/0.191846;
scaled_Parana = (Parana-0.147583)/0.355138;
scaled_Amazonas_Solimoes = (Amazonas_Solimoes-0.0178117)/0.132435;
scaled_Amazonia = (Amazonia-0.00254453)/0.0504433;
y_1_1 = Logistic (0.9109+ (scaled_Argissolo*-0.562655)+ (scaled_Neossolo*0.482884)+
(scaled_Planossolo*-0.356395)+ (scaled_Luvisso*0.0661156)+ (scaled_Espodossolo*-
0.661165)+ (scaled_Latossolo*1.03785)+ (scaled_Vertissolo*-0.634753)+
(scaled_Cambissolo*-1.1583)+ (scaled_Plintossolo*-0.201498)+ (scaled_Nitossolo*-
0.248461)+ (scaled_Organossolo*-0.244116)+ (scaled_Dunas*-0.747859)+
(scaled_Gleissolo*-0.0587325)+ (scaled_umido*-0.233544)+
(scaled_semiarido*1.16657)+ (scaled_arido*-1.21157)+ (scaled_subumolhado*-1.1896)+
(scaled_subuseco*0.27908)+ (scaled_Costeira_MargemContinental*0.630182)+
(scaled_CobCenozoica*0.323949)+ (scaled_Borborema*0.406449)+
(scaled_Reconcavo_Tucano_Jatoba*-1.42933)+ (scaled_SaoFrancisco*0.0394679)+
(scaled_Parnaiba*0.236112)+ (scaled_Mantiqueira*-0.544257)+ (scaled_Tocantins*-
0.0216203)+ (scaled_Parana*0.597146)+ (scaled_Amazonas_Solimoes*0.00984775)+
(scaled_Amazonia*0.120275));
y_1_2 = Logistic (-1.12232+ (scaled_Argissolo*-1.1166)+ (scaled_Neossolo*0.192265)+
(scaled_Planossolo*-0.239239)+ (scaled_Luvisso*1.12808)+ (scaled_Espodossolo*-
0.343818)+ (scaled_Latossolo*-0.0328055)+ (scaled_Vertissolo*-0.2834)+
(scaled_Cambissolo*-0.4639)+ (scaled_Plintossolo*0.0197874)+ (scaled_Nitossolo*-
0.0636594)+ (scaled_Organossolo*-0.269052)+ (scaled_Dunas*-0.532296)+

```

(scaled_Gleissolo*-0.471586)+(scaled_umido*-0.354488)+(scaled_semiarido*-0.652879)+(scaled_arido*0.941965)+(scaled_subumolhado*0.317007)+(scaled_subuseco*0.996665)+(scaled_Costeira_MargemContinental*-0.444875)+(scaled_CobCenozoica*-0.502196)+(scaled_Borborema*-1.9477)+(scaled_Reconcavo_Tucano_Jatoba*-0.388874)+(scaled_SaoFrancisco*0.994204)+(scaled_Parnaiba*-0.549999)+(scaled_Mantiqueira*-0.172675)+(scaled_Tocantins*-0.217851)+(scaled_Parana*1.01265)+(scaled_Amazonas_Solimoes*-0.353722)+(scaled_Amazonia*-0.0372485));
 y_1_3 = Logistic (0.324244+(scaled_Argissolo*-0.521932)+(scaled_Neossolo*2.61155)+(scaled_Planossolo*-0.0483673)+(scaled_Luvissolo*-0.183041)+(scaled_Espodossolo*0.53465)+(scaled_Latossolo*-2.03593)+(scaled_Vertissolo*0.478038)+(scaled_Cambissolo*0.353278)+(scaled_Plintossolo*0.212764)+(scaled_Nitossolo*-0.285795)+(scaled_Organossolo*0.19464)+(scaled_Dunas*-0.216839)+(scaled_Gleissolo*-0.130236)+(scaled_umido*-0.946436)+(scaled_semiarido*-0.339532)+(scaled_arido*0.922047)+(scaled_subumolhado*1.14023)+(scaled_subuseco*-0.68745)+(scaled_Costeira_MargemContinental*0.476753)+(scaled_CobCenozoica*-0.695934)+(scaled_Borborema*-0.704799)+(scaled_Reconcavo_Tucano_Jatoba*0.577688)+(scaled_SaoFrancisco*1.2692)+(scaled_Parnaiba*0.423287)+(scaled_Mantiqueira*-0.0877955)+(scaled_Tocantins*-0.113761)+(scaled_Parana*-0.563557)+(scaled_Amazonas_Solimoes*-0.578114)+(scaled_Amazonia*-0.00664421));
 y_1_4 = Logistic (-0.165541+(scaled_Argissolo*0.379186)+(scaled_Neossolo*-0.827383)+(scaled_Planossolo*0.164554)+(scaled_Luvissolo*-0.0997703)+(scaled_Espodossolo*0.460536)+(scaled_Latossolo*-0.617469)+(scaled_Vertissolo*0.399496)+(scaled_Cambissolo*0.590853)+(scaled_Plintossolo*-0.0852574)+(scaled_Nitossolo*0.0480216)+(scaled_Organossolo*0.526408)+(scaled_Dunas*0.840671)+(scaled_Gleissolo*0.221989)+(scaled_umido*-0.512505)+(scaled_semiarido*-0.535813)+(scaled_arido*0.630293)+(scaled_subumolhado*1.09267)+(scaled_subuseco*-1.25203)+(scaled_Costeira_MargemContinental*-0.499371)+(scaled_CobCenozoica*-0.783655)+(scaled_Borborema*-0.589389)+(scaled_Reconcavo_Tucano_Jatoba*-0.185003)+(scaled_SaoFrancisco*0.478974)+(scaled_Parnaiba*-0.0670511)+(scaled_Mantiqueira*0.415492)+(scaled_Tocantins*-0.337987)+(scaled_Parana*1.52136)+(scaled_Amazonas_Solimoes*0.413631)+(scaled_Amazonia*0.345901));
 y_1_5 = Logistic (-1.46252+(scaled_Argissolo*-1.3724)+(scaled_Neossolo*1.42123)+(scaled_Planossolo*-0.615933)+(scaled_Luvissolo*0.471772)+(scaled_Espodossolo*0.142004)+(scaled_Latossolo*-0.630427)+(scaled_Vertissolo*-0.251903)+(scaled_Cambissolo*0.42597)+(scaled_Plintossolo*0.194722)+(scaled_Nitossolo*0.277943)+(scaled_Organossolo*0.530631)+(scaled_Dunas*-0.331389)+(scaled_Gleissolo*-0.212066)+(scaled_umido*0.91335)+(scaled_semiarido*-0.467074)+(scaled_arido*-1.81328)+(scaled_subumolhado*0.0415083)+(scaled_subuseco*-0.497855)+(scaled_Costeira_MargemContinental*0.151202)+(scaled_CobCenozoica*-2.52719)+(scaled_Borborema*0.624016)+(scaled_Reconcavo_Tucano_Jatoba*0.736243)+(scaled_SaoFrancisco*0.291576)+(scaled_Parnaiba*0.167652)+(scaled_Mantiqueira*0.8936)+(scaled_Tocantins*-0.0741716)+(scaled_Parana*-0.103656)+(scaled_Amazonas_Solimoes*-0.0763654)+(scaled_Amazonia*-0.226106));
 y_1_6 = Logistic (-0.858593+(scaled_Argissolo*-0.632104)+(scaled_Neossolo*-2.18425)+(scaled_Planossolo*0.0332736)+(scaled_Luvissolo*1.89109)+

```

(scaled_Espodossolo*0.283699)+(scaled_Latossolo*1.45802)+
(scaled_Vertissolo*0.556234)+(scaled_Cambissolo*0.413183)+
(scaled_Plintossolo*0.320496)+(scaled_Nitossolo*0.398595)+
(scaled_Organossolo*0.0056888)+(scaled_Dunas*0.0939892)+(scaled_Gleissolo*-
0.109608)+(scaled_umido*1.44894)+(scaled_semiarido*0.697119)+
(scaled_arido*0.330905)+(scaled_subumolhado*-1.61737)+
(scaled_subuseco*0.105482)+(scaled_Costeira_MargemContinental*-0.074858)+
(scaled_CobCenozoica*-0.109746)+(scaled_Borborema*-1.2546)+
(scaled_Reconcavo_Tucano_Jatoba*1.23973)+(scaled_SaoFrancisco*0.718232)+
(scaled_Parnaiba*0.201)+(scaled_Mantiqueira*0.276044)+(scaled_Tocantins*-
0.598171)+(scaled_Parana*-0.675438)+(scaled_Amazonas_Solimoes*0.0760764)+
(scaled_Amazonia*-0.0664998));
y_1_7 = Logistic (-0.84205+(scaled_Argissolo*0.0821398)+
(scaled_Neossolo*0.248346)+(scaled_Planossolo*-0.420246)+(scaled_Luvisso*-
0.542639)+(scaled_Espodossolo*0.143378)+(scaled_Latossolo*1.38258)+
(scaled_Vertissolo*-0.412911)+(scaled_Cambissolo*-0.122575)+
(scaled_Plintossolo*0.110495)+(scaled_Nitossolo*-0.279257)+(scaled_Organossolo*-
0.229677)+(scaled_Dunas*0.00229444)+(scaled_Gleissolo*0.251174)+
(scaled_umido*0.513045)+(scaled_semiarido*-0.724441)+(scaled_arido*-0.0385427)+
(scaled_subumolhado*0.471535)+(scaled_subuseco*-0.797236)+
(scaled_Costeira_MargemContinental*-0.918775)+(scaled_CobCenozoica*-1.93773)+
(scaled_Borborema*0.606531)+(scaled_Reconcavo_Tucano_Jatoba*0.981186)+
(scaled_SaoFrancisco*0.753791)+(scaled_Parnaiba*0.164525)+
(scaled_Mantiqueira*1.62598)+(scaled_Tocantins*-0.280789)+(scaled_Parana*-
1.08234)+(scaled_Amazonas_Solimoes*0.00792791)+(scaled_Amazonia*-0.288122));
non_probabilistic_SOLO_EXP_COL_ = Logistic (-0.400954+(y_1_1*-3.94238)+
(y_1_2*-4.34826)+(y_1_3*3.93771)+(y_1_4*2.96398)+(y_1_5*4.82757)+
(y_1_6*4.12318)+(y_1_7*-4.66562));
SOLO_EXP_COL_ = probability(non_probabilistic_SOLO_EXP_COL_);

logistic(x){
return 1/(1+exp(-x))
}

probability(x){
if x < 0
return 0
else if x > 1
return 1
else
return x
}

```

APÊNDICE D – INTERFACE DO PROGRAMA SECBRAPP

A interface do programa desenvolvido com a equação da rede BR03 nomeada como SECBRApp é apresentada na Figura D1, na qual inicialmente o usuário deverá inserir como dados de entrada as coordenadas da localidade a ser investigada no canto superior esquerdo (item 1), as coordenadas devem ser inseridas no campo de busca no formato Graus Decimais e estarem georreferenciadas no Datum SIRGAS 2000, assim como os mapas temáticos fundamentais. Nos casos em que a coordenada não esteja no formato necessário, o ícone de ajuda poderá ser acionado (item 2), apresentando um link de acesso ao conversor de coordenadas do INPE (Instituto Nacional de Pesquisas Espaciais).

O sistema de referência SIRGAS 2000 tem a finalidade de estabelecer um sistema de referência geocêntrico para a América do Sul. E em 2005 a Resolução do IBGE N°1/2005 estabeleceu o Sistema de Referência Geocêntrico para as Américas (SIRGAS), em sua realização do ano de 2000 (SIRGAS 2000), como novo SGR para o Sistema Geodésico Brasileiro (SGB) (MONICO, 2008).

Figura D1 - Interface do programa de classificação de solos colapsáveis e expansivos

The interface is titled 'JNeuraSoil v1.0' and features logos for 'UNIVERSIDADE FEDERAL DE PERNAMBUCO' and 'GN SAT'. It includes a help icon (item 2) in the top right corner.

ENTRADA:

Item 1: Coordenada X (-38,010) and Coordenada Y (-8,316).
 Item 3: Tipo de Resultado (radio buttons for Expansão and Colapso).
 Item 4: Executar button.

RESULTADO:

Item 5: Progress bar for 'Processando Pedologia...'.
 Item 6: Output fields for Climatologia (Semiarido), Geologia (Borborema), Pedologia, Probabilidade, and Indicativo De:.

HISTÓRICO:

Item 7: Table with 9 columns: Log Id, Data, Coord X, Coord Y, Climatologia, Geologia, Pedologia, Probab., and Indicativo De:.

Item 8: Exportar para CSV button.

Log Id	Data	Coord X	Coord Y	Climatologia	Geologia	Pedologia	Probab.	Indicativo De:
8	12/07/2022	-38,010	-8,316	Semiarido	Borborema	Planossolo	0,999595	Expansivo
9	12/07/2022	-49,025	-23,710	Subumido	Cobertura Cen...	Latossolo	0,629603	Expansivo
10	13/08/2022	-38,010	-8,316	Semiarido	Borborema	Planossolo	0,999595	Expansivo
11	13/08/2022	-38,010	-8,316	Semiarido	Borborema	Planossolo	0,999595	Expansivo
12	13/08/2022	-38,010	-8,316	Semiarido	Borborema	Planossolo	0,999595	Expansivo

Fonte: A Autora (2022).

Posteriormente deverá ser indicada uma previsão de tipo de resultado a ser encontrado (colapso ou expansão) selecionando a opção desejada (item 3), seguindo da execução do programa (item 4), e aguardar a busca das características para cada classe (clima, geologia e pedologia) (item 5).

Os resultados de cada característica das variáveis de entrada, bem como da probabilidade e da característica indicada pela classificação do programa serão descritos no centro da tela (item 6). E o histórico das classificações feitas fica registrado na tabela ao final da tela, contendo a identificação numérica da amostra classificada, as coordenadas usadas na classificação e os resultados: características verificadas nos mapas fundamentais, resultado indicado pela previsão e probabilidade da classificação que indica o grau de suscetibilidade (item 7). Esse resultado ou conjunto de resultados poderá ser exportado para planilhas em formato .csv (item 8).