
Desenvolvimento de um Modelo Constitutivo para Solos Arenosos Reforçados com Fibras Utilizando Redes Neurais Artificiais

João Paulo de Araújo Rodrigues



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA CIVIL
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

Uberlândia
2026

João Paulo de Araújo Rodrigues

**Desenvolvimento de um Modelo Constitutivo
para Solos Arenosos Reforçados com Fibras
Utilizando Redes Neurais Artificiais**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Engenharia Civil da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Engenharia Civil.

Área de concentração: Construção civil, estruturas e geotecnia.

Orientador: Prof. Dr. Michael Andrade Maedo

Uberlândia

2026

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

R696
2026 Rodrigues, João Paulo de Araujo, 1985-
Desenvolvimento de um Modelo Constitutivo para Solos Arenosos Reforçados com Fibras Utilizando Redes Neurais Artificiais [recurso eletrônico] / João Paulo de Araujo Rodrigues. - 2026.

Orientador: Michael Andrade Maedo.

Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Pós-graduação em Engenharia Civil.

Modo de acesso: Internet.

DOI <http://doi.org/10.14393/ufu.di.2026.592>

Inclui bibliografia.

Inclui ilustrações.

1. Engenharia civil. I. Maedo, Michael Andrade, 1987-, (Orient.).
II. Universidade Federal de Uberlândia. Pós-graduação em
Engenharia Civil. III. Título.

CDU: 624

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:

Gizele Cristine Nunes do Couto - CRB6/2091

Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Engenharia Civil
 Av. João Naves de Ávila, 2121, Bloco 1Y - Bairro Santa Monica, Uberlândia-MG, CEP 38400-902
 Telefone: 34 3239-4137 - www.feciv.ufu.br/ppgec - posgradcivil@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Civil				
Defesa de:	Dissertação de Mestrado Acadêmico, 336, PPGE				
Data:	29 de julho de 2026	Hora de início:	9:02	Hora de encerramento:	10:30
Matrícula do Discente:	12422ECV010				
Nome do Discente:	João Paulo de Araújo Rodrigues				
Título do Trabalho:	Desenvolvimento de um Modelo Constitutivo para Solos Arenosos Reforçados com Fibras Utilizando Redes Neurais Artificiais				
Área de concentração:	Construção Civil, Estruturas e Geotecnia				
Linha de pesquisa:	Geotecnia				
Projeto de Pesquisa de vinculação:	Modelagem e simulação computacional do comportamento mecânico de solos reforçados com fibras - CNPQ - 134533/2025-9				

Reuniu-se, em sessão pública por webconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Civil, assim composta pelos Professores Doutores: Caio Gorla Nogueira - UNESP, Leonardo José do Nascimento Guimarães - UFPE e Michael Andrade Maedo, orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Michael Andrade Maedo, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Michael Andrade Maedo, Professor(a) do Magistério Superior**, em 29/07/2026, às 10:27, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Leonardo José do Nascimento Guimarães, Usuário Externo**, em 29/07/2026, às 10:39, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **CAIO GORLA NOGUEIRA, Usuário Externo**, em 29/07/2026, às 10:45, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **7530751** e o código CRC **A6CD3E5A**.

Agradecimentos

À Saira, minha esposa, dedico o mais profundo sentimento de gratidão. Por sua paciência, incentivo e por ser meu refúgio seguro em todos os desafios que esta pesquisa impôs. Sua presença foi a força motriz para a conclusão desta etapa.

Ao meu orientador, Professor Doutor Michael, expresso minha sincera gratidão pela confiança depositada em meu potencial, mesmo com minha formação externa à área de pesquisa. Seu acolhimento e sua orientação foram fundamentais para a concretização deste trabalho.

Agradeço, por fim, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pelo fomento concedido por meio do Processo nº 134533/2025-9. O suporte financeiro foi indispensável para a realização desta pesquisa e para o avanço dos estudos aqui desenvolvidos.

Resumo

O reforço de solos arenosos com fibras distribuídas aleatoriamente melhora significativamente suas propriedades mecânicas, impondo desafios à modelagem constitutiva devido às complexas interações não-lineares entre a matriz e o reforço. Este trabalho desenvolve um modelo constitutivo baseado em Redes Neurais Artificiais (RNA) para prever o comportamento tensão-deformação de uma areia uniforme reforçada com fibras poliméricas. A metodologia emprega uma abordagem híbrida em duas fases, combinando dados experimentais de ensaios triaxiais com dados sintéticos gerados por meio de uma adaptação do modelo Cam-Clay modificado. A base experimental, obtida na Universidade Federal da Bahia, abrange teores de fibra de 0%, 0,5% e 1,0%, comprimentos de fibra de 12,5 mm, 25 mm e 51 mm, e tensões de confinamento de 50, 100, 200 e 300 kPa, e dela se utilizaram os 24 ensaios de compressão triaxial drenada convencional, totalizando 1.207 pontos. A geração de dados sintéticos produziu cerca de 19,5 mil pontos, distribuídos em 42 cenários, expandindo o domínio experimental para incluir níveis intermediários de confinamento. A arquitetura adotada é uma rede neural do tipo perceptron multicamadas (*Multilayer Perceptron*, MLP) que mapeia diretamente as condições de ensaio e o nível de deformação axial nas variáveis de estado de interesse, a saber, a tensão desviadora, a tensão média efetiva e a deformação volumétrica. O treinamento combina pré-treinamento com os dados sintéticos fisicamente consistentes e ajuste fino, por aprendizado por transferência, com os dados experimentais. A validação, conduzida por validação cruzada exaustiva do tipo *leave-one-out*, resultou em coeficiente de determinação agregado de 0,977 para a tensão desviadora, 0,991 para a tensão média efetiva e 0,914 para a deformação volumétrica. O modelo reproduz a mobilização progressiva das fibras, o pico de resistência e a transição para o comportamento dilatante, oferecendo uma alternativa computacionalmente eficiente aos modelos analíticos tradicionais de difícil calibração.

Palavras-chave: Solo Reforçado com Fibras. Redes Neurais Artificiais. Modelo Constitutivo. Ensaio Triaxial. Cam-Clay modificado. Aprendizado por Transferência.

Abstract

The reinforcement of sandy soils with randomly distributed fibers significantly improves their mechanical properties, posing challenges to constitutive modeling due to complex nonlinear interactions between the matrix and the reinforcement. This work develops a constitutive model based on Artificial Neural Networks (ANN) to predict the stress-strain behavior of a uniform sand reinforced with polymeric fibers. The methodology employs a two-phase hybrid approach, combining experimental triaxial test data with synthetic data generated through an adapted Modified Cam-Clay model. The experimental database, obtained at the Federal University of Bahia, encompasses fiber contents of 0%, 0.5%, and 1.0%, fiber lengths of 12.5, 25, and 51 mm, and confining stresses of 50, 100, 200, and 300 kPa; from it, the 24 conventional drained triaxial compression tests were used, totaling 1,207 points. The synthetic data generation produced about 19.5 thousand points across 42 scenarios, expanding the experimental domain to include intermediate confinement levels. The adopted architecture is a multilayer perceptron (MLP) neural network that directly maps the test conditions and the axial strain level onto the state variables of interest, namely the deviatoric stress, the mean effective stress, and the volumetric strain. Training combines pre-training with the physically consistent synthetic data and fine-tuning, through transfer learning, with the experimental data. Validation, carried out by exhaustive leave-one-out cross-validation, yielded an aggregate coefficient of determination of 0.977 for the deviatoric stress, 0.991 for the mean effective stress, and 0.914 for the volumetric strain. The model reproduces the progressive mobilization of the fibers, the peak strength, and the transition to dilatant behavior, offering a computationally efficient alternative to traditional analytical models that are difficult to calibrate.

Keywords: Fiber-Reinforced Soil. Artificial Neural Networks. Constitutive Model. Triaxial Test. Modified Cam-Clay. Transfer Learning.

Lista de ilustrações

Figura 1 – Estrutura de um perceptron multicamadas: neurônios organizados em uma camada de entrada, camadas ocultas e uma camada de saída, com conexões totais entre camadas sucessivas e fluxo de informação em sentido único	20
Figura 2 – Curva granulométrica da areia uniforme	27
Figura 3 – Exemplos de trajetórias tensão-deformação sintéticas geradas pelo modelo analítico	34
Figura 4 – Esquema da validação cruzada <i>leave-one-out</i> : o pré-treinamento, executado uma única vez, fornece o ponto de partida comum às 24 rodadas de ajuste fino; em cada rodada, um ensaio é reservado para teste, e as previsões dos 24 ensaios reservados são reunidas ao final para o cálculo das métricas agregadas	39
Figura 5 – Visão de conjunto da rede neural: da entrada às variáveis de estado de saída, com o detalhe de um neurônio e o ciclo de treinamento, no qual as previsões da rede ($\hat{y}$) são comparadas aos valores observados (y) pela função de perda	41
Figura 6 – Arquitetura da rede neural de propagação direta adotada	42
Figura 7 – Mapeamento direto das variáveis de entrada nas variáveis de estado de saída	44
Figura 8 – Função de ativação ELU adotada nas camadas ocultas, comparada às funções tanh, ReLU e Leaky ReLU	46
Figura 9 – Estratégia híbrida de treinamento em duas fases: pré-treinamento com dados sintéticos e ajuste fino com dados experimentais	48
Figura 10 – Liberação progressiva das camadas no ajuste fino, da camada de saída em direção à de entrada	50
Figura 11 – Agregação das previsões da validação cruzada <i>leave-one-out</i> para o cálculo das métricas globais	53

Figura 12 – Dispersão entre valores previstos e observados na validação cruzada, para a tensão desviadora (q) e a tensão média efetiva (p')	57
Figura 13 – Distribuição dos coeficientes de determinação por ensaio, segundo a classe de teor de fibra	59
Figura 14 – Curvas tensão-deformação e de variação volumétrica previstas e observadas, para os ensaios de pior, mediano e melhor ajuste de q , com a previsão do modelo analítico de referência (Seção 4.6)	61
Figura 15 – Evolução da função de perda no pré-treinamento e no ajuste fino . . .	62
Figura 16 – Ensaios 1 a 3	79
Figura 17 – Ensaios 4 a 6	80
Figura 18 – Ensaios 7 a 9	81
Figura 19 – Ensaios 10 a 12	82
Figura 20 – Ensaios 13 a 15	83
Figura 21 – Ensaios 16 a 18	84
Figura 22 – Ensaios 19 a 21	85
Figura 23 – Ensaios 22 a 24	86

Lista de tabelas

Tabela 1	–	Propriedades índice da areia uniforme	27
Tabela 2	–	Propriedades das fibras poliméricas	28
Tabela 3	–	Composição da base experimental utilizada, em ensaios triaxiais drenados convencionais	29
Tabela 4	–	Equações constitutivas do modelo Cam-Clay modificado adaptado para solos reforçados com fibras	30
Tabela 5	–	Parâmetros do modelo Cam-Clay modificado adaptado	31
Tabela 6	–	Parâmetros do modelo adotados na geração dos dados sintéticos	33
Tabela 7	–	Resultados da validação cruzada <i>leave-one-out</i> por ensaio	58
Tabela 8	–	Desempenho do modelo analítico de referência e da rede sobre os 1.207 pontos experimentais	64
Tabela 9	–	Estudo de ablação: desempenho com e sem o pré-treinamento sintético, em média sobre cinco repetições independentes do experimento	65

Sumário

1	INTRODUÇÃO	10
1.1	Objetivos	11
1.1.1	Objetivo Geral	11
1.1.2	Objetivos Específicos	11
1.2	Justificativa	12
2	REVISÃO BIBLIOGRÁFICA	14
2.1	Solos Reforçados com Fibras	14
2.2	Modelagem Constitutiva de Solos	16
2.2.1	Modelos clássicos: capacidades e limitações	16
2.2.2	Desafios na modelagem de solos reforçados	18
2.3	Redes Neurais Artificiais em Geotecnia	19
2.3.1	Fundamentos das redes neurais artificiais	19
2.3.2	Aplicações estabelecidas em geotecnia	21
2.3.3	Modelos constitutivos baseados em RNA	21
2.3.4	Vantagens da abordagem baseada em dados e a oportunidade para os solos reforçados com fibras	22
3	MATERIAIS E MÉTODOS	25
3.1	Base de Dados Experimentais	25
3.1.1	Origem e descrição dos dados	25
3.1.2	Características dos materiais ensaiados	26
3.1.3	Seleção e caracterização da base utilizada	28
3.2	Geração de Dados Sintéticos	30
3.2.1	Formulação do modelo constitutivo	30
3.2.2	Calibração e geração das trajetórias sintéticas	31
3.3	Preparação e Tratamento dos Dados	34
3.3.1	Estruturação e limpeza	34

3.3.2	A variável indicadora de reforço	35
3.3.3	Normalização	36
3.3.4	Estratégia de validação	37
3.4	Arquitetura da Rede Neural Artificial	40
3.4.1	Tipo e estrutura da rede	41
3.4.2	Variáveis de entrada e saída	43
3.4.3	Função de ativação	44
3.4.4	Função de perda	46
3.5	Estratégia de Treinamento	47
3.5.1	Pré-treinamento	48
3.5.2	Ajuste fino	50
3.6	Métricas de Avaliação	51
3.7	Ferramentas Computacionais	53
4	RESULTADOS E DISCUSSÃO	55
4.1	Desempenho Global do Modelo	55
4.2	Distribuição do Desempenho por Ensaio	58
4.3	Efeito do Reforço e do Confinamento	59
4.4	Curvas Tensão-Deformação Representativas	60
4.5	Comportamento do Treinamento	62
4.6	Comparação com o Modelo Analítico de Referência	63
4.7	Contribuição do Pré-Treinamento Sintético	64
4.8	Discussão da Métrica sob Baixo Confinamento	67
5	CONCLUSÕES	68
5.1	Limitações	69
5.2	Sugestões para Trabalhos Futuros	70
	REFERÊNCIAS	71

APÊNDICES 77

APÊNDICE A	– CURVAS TENSÃO-DEFORMAÇÃO DE TODOS OS ENSAIOS	78
-------------------	---	-----------

Introdução

O desenvolvimento urbano contemporâneo e a crescente demanda por infraestrutura têm levado à necessidade de construção em áreas com solos de propriedades geotécnicas desfavoráveis. Neste contexto, o reforço de solos com fibras distribuídas aleatoriamente emerge como uma técnica promissora para o melhoramento das características mecânicas de materiais geotécnicos, oferecendo vantagens como a facilidade de execução, a melhoria isotrópica das propriedades e a possibilidade de utilização de materiais sustentáveis (HEJAZI et al., 2012).

Os solos arenosos, amplamente encontrados no território brasileiro, apresentam desafios específicos na prática geotécnica, incluindo baixa coesão, suscetibilidade à liquefação e comportamento dilatante sob cisalhamento (SANTAMARINA; KLEIN; FAM, 2001). A adição de fibras a esses materiais modifica significativamente seu comportamento mecânico, resultando em aumento de resistência, maior ductilidade e redução da fragilidade (CONSOLI et al., 2009).

A modelagem constitutiva adequada desses geocompósitos representa um desafio fundamental para aplicações práticas em engenharia. Os modelos constitutivos clássicos, desenvolvidos para solos homogêneos, falham em capturar características essenciais dos solos reforçados, como a mobilização progressiva das fibras, a evolução da anisotropia e as interações complexas entre matriz e reforço (DIAMBRA et al., 2013). Tentativas de adaptar modelos existentes frequentemente resultam em formulações com um número excessivo de parâmetros, cuja calibração requer extensos programas experimentais (MACHADO et al., 2024).

As Redes Neurais Artificiais (RNA) oferecem uma abordagem alternativa para superar essas limitações. Sua capacidade de aprender relações complexas e não-lineares diretamente de dados experimentais, sem a necessidade de formulações matemáticas explícitas, torna-as particularmente adequadas para a modelagem de materiais com comportamento complexo (GHABOUSSI; JR; WU, 1991). Aplicações bem-sucedidas em diversos problemas geotécnicos demonstram o potencial das RNA para capturar comportamentos que desafiam a modelagem tradicional (SHAHIN; JAKSA; MAIER, 2001).

Esta pesquisa beneficia-se de um banco de dados experimental obtido na Universidade Federal da Bahia (UFBA) por Conceição, Carvalho e Machado (2022), que investigou sistematicamente o comportamento tensão-deformação de uma areia uniforme reforçada com fibra polimérica por meio de ensaios de compressão triaxial. Esse mesmo conjunto experimental subsidiou o desenvolvimento do modelo constitutivo analítico de Machado et al. (2024), mas ainda não havia sido explorado por meio de técnicas de aprendizado de máquina, o que representa uma oportunidade para o desenvolvimento de modelos constitutivos baseados em inteligência artificial.

A estratégia adotada combina o pré-treinamento de uma rede neural com dados sintéticos, gerados pelo modelo constitutivo de Machado et al. (2024), baseado na teoria do estado crítico e adaptado para solos reforçados, e um posterior ajuste fino com os dados experimentais reais. A lógica é a do aprendizado por transferência (PAN; YANG, 2010): a rede primeiro aprende, a partir do grande volume de dados sintéticos, o comportamento físico descrito pelo modelo analítico; em seguida, o ajuste fino com o conjunto experimental, mais escasso, corrige e especializa essa resposta para o material real. Desse modo, o conhecimento físico é incorporado à rede pelos próprios dados de pré-treinamento, antes do contato com os dados de laboratório.

1.1 Objetivos

1.1.1 Objetivo Geral

Desenvolver um modelo constitutivo baseado em Redes Neurais Artificiais para prever o comportamento tensão-deformação de solos arenosos reforçados com fibras poliméricas. O modelo mapeia diretamente as condições de ensaio e o nível de deformação imposto nas variáveis de estado de interesse, a saber, a tensão desviadora, a tensão média efetiva e a deformação volumétrica, por meio de uma estratégia de treinamento em duas fases que combina pré-treinamento com dados sintéticos e ajuste fino com dados experimentais.

1.1.2 Objetivos Específicos

- ❑ Estruturar e processar a base de dados experimentais de ensaios triaxiais drenados convencionais, aplicando rotinas de extração, limpeza e padronização;
- ❑ Gerar uma base de dados sintética fisicamente consistente por meio do modelo Cam-Clay modificado adaptado para solos reforçados, destinada ao pré-treinamento da rede;
- ❑ Desenvolver e ajustar uma arquitetura de rede neural do tipo perceptron multicamadas (*Multilayer Perceptron*, MLP) adequada ao mapeamento constitutivo proposto;

- ❑ Implementar uma estratégia de treinamento em duas fases, com pré-treinamento nos dados sintéticos e posterior ajuste fino, por aprendizado por transferência, nos dados experimentais;
- ❑ Validar o modelo por meio de validação cruzada exaustiva do tipo *leave-one-out* e de métricas estatísticas apropriadas à heterogeneidade de escala entre os ensaios;
- ❑ Avaliar criticamente a consistência física das previsões e discutir as limitações do modelo.

1.2 Justificativa

O reforço de solos arenosos com fibras é uma técnica de melhoramento de baixo custo, execução simples e ampla aplicabilidade, empregada na estabilização de taludes e aterros, no suporte de fundações superficiais, em camadas de pavimento, no controle de erosão e em estruturas de contenção em solo reforçado (HEJAZI et al., 2012; CONSOLI et al., 2009). Em um país como o Brasil, onde os solos arenosos são abundantes, o domínio do comportamento mecânico desses geocompósitos tem relevância direta para a engenharia geotécnica.

Apesar dos benefícios demonstrados, a adoção dessa técnica em projeto depende da capacidade de prever a resposta tensão-deformação do compósito no interior de análises numéricas, como as de elementos finitos. É justamente nesse ponto que reside o principal entrave: os modelos constitutivos disponíveis ou são simples demais para representar a mobilização progressiva do reforço, ou exigem um grande número de parâmetros de difícil determinação experimental (DIAMBRA; IBRAIM, 2015; MACHADO et al., 2024). A ausência de uma descrição constitutiva ao mesmo tempo precisa e prática limita o pleno aproveitamento da técnica.

As Redes Neurais Artificiais oferecem um caminho para contornar esse impasse com uma vantagem específica sobre as adaptações de modelos analíticos: novas variáveis que governam o reforço, como o teor e o comprimento das fibras, podem ser incorporadas pela simples ampliação da camada de entrada, sem a reformulação matemática e a multiplicação de parâmetros que as abordagens tradicionais exigem. Essa flexibilidade favorece o desenvolvimento incremental do modelo à medida que novos dados se tornam disponíveis (JAVADI; REZANIA, 2009).

Do ponto de vista metodológico, a principal justificativa está na estratégia de pré-treinamento com dados sintéticos. A obtenção de dados experimentais em geotecnia é onerosa e demorada, o que restringe o treinamento de redes neurais. Ao gerar um grande volume de exemplos fisicamente consistentes a partir de um modelo analítico e empregá-los no pré-treinamento, busca-se contornar essa restrição, incorporando conhecimento físico

estabelecido antes do ajuste fino com o conjunto experimental, mais limitado (MASI et al., 2021).

O uso do mesmo arcabouço analítico de Machado et al. (2024) como gerador dos dados sintéticos confere à pesquisa uma vantagem adicional: permite contrastar de forma direta o comportamento herdado da teoria com aquele corrigido pelos dados experimentais, contribuindo para a compreensão das relações de troca entre a modelagem baseada em física e a modelagem baseada em dados. Soma-se a isso a eficiência computacional do modelo neural, que, uma vez treinado, fornece previsões a custo desprezível (STOFFEL; BAMER; MARKERT, 2018), compatível com as demandas iterativas de projeto e de otimização.

Por fim, até onde se tem conhecimento, não há na literatura modelo constitutivo baseado em aprendizado de máquina para a areia reforçada com fibras aqui estudada. Ao preencher essa lacuna, o trabalho não apenas avança a modelagem de geocompósitos, como também demonstra uma metodologia replicável, que combina pré-treinamento sintético, aprendizado por transferência e validação cruzada rigorosa, passível de extensão a outros materiais geotécnicos de comportamento complexo, contribuindo para o avanço da mecânica computacional dos solos.

Revisão Bibliográfica

Este capítulo apresenta os fundamentos teóricos necessários para o desenvolvimento do modelo constitutivo baseado em redes neurais artificiais para solos reforçados com fibras. A revisão está organizada em três eixos principais: (i) o comportamento mecânico de solos reforçados com fibras; (ii) os modelos constitutivos tradicionais e suas limitações; e (iii) a aplicação de redes neurais artificiais em geotecnia, da qual se examinam os fundamentos, os modelos constitutivos existentes, as vantagens da abordagem baseada em dados e a oportunidade de contribuição para a modelagem de geocompósitos.

2.1 Solos Reforçados com Fibras

O reforço de solos com fibras distribuídas aleatoriamente constitui uma técnica eficaz para o melhoramento das propriedades mecânicas de materiais geotécnicos, particularmente no aumento da resistência ao cisalhamento e da ductilidade (HEJAZI et al., 2012). Diferentemente de reforços planares como geogrelhas, as fibras criam uma rede tridimensional que se mobiliza progressivamente com o desenvolvimento das deformações, proporcionando um reforço isotrópico ao material (GRAY; OHASHI, 1983; HEJAZI et al., 2012).

A eficácia do reforço depende fundamentalmente de três parâmetros principais: teor de fibras, comprimento das fibras e nível de confinamento. Michalowski e Čermák (2003) demonstraram experimentalmente que o ganho de resistência cresce com o teor e com a razão de aspecto das fibras, tendo registrado acréscimos de tensão de ruptura de até 70% para o teor volumétrico de 2%. O ganho de resistência cresce de forma aproximadamente linear com o teor de fibras até teores moderados e, em seguida, aproxima-se de um limite assintótico, a partir do qual o contato fibra-fibra e o emaranhamento das fibras comprometem a distribuição uniforme na mistura e reduzem o atrito mobilizado entre os grãos (HEJAZI et al., 2012). Esse efeito é coerente com a redução da eficiência de compactação observada para teores elevados.

O comprimento das fibras influencia diretamente os mecanismos de mobilização. Michalowski e Čermák (2003) observaram que o efeito do reforço depende da relação entre o comprimento da fibra e o tamanho dos grãos: para um mesmo comprimento de fibra e baixos teores, o ganho de resistência foi maior na areia fina do que na areia grossa, o que indica a necessidade de comprimentos muito superiores ao diâmetro médio das partículas (D_{50}) para mobilizar eficientemente a interação fibra-grão. Ensaios triaxiais com diferentes comprimentos mostram que a contribuição das fibras se manifesta sobretudo a partir de um certo nível de deformação cisalhante, sendo tanto maior quanto mais longa a fibra, em razão da maior área de contato disponível para a mobilização das tensões de tração (DIAMBRA; IBRAIM, 2015).

Em ensaios de compressão triaxial em condição drenada, o comportamento tensão-deformação de solos reforçados exibe características distintivas: (i) aumento expressivo da tensão de ruptura, acompanhado de redução da rigidez inicial e de aumento da deformação na ruptura (MICHALOWSKI; ČERMÁK, 2003); (ii) transição de comportamento frágil para dúctil (CONSOLI et al., 2009); (iii) modificação marcante do comportamento tensão-dilatância do compósito (CONSOLI et al., 2009); e (iv) manutenção ou aumento da resistência última, sem perda de resistência mesmo a deslocamentos muito grandes, aspecto fundamental para a estabilidade pós-ruptura (HEINECK; COOP; CONSOLI, 2005).

A mobilização das fibras ocorre por meio de mecanismos progressivos que dependem do nível de deformação. Para deformações pequenas ($\varepsilon_a < 1\%$), o comportamento é dominado pela matriz do solo, e a inclusão das fibras praticamente não altera a rigidez inicial do material (HEINECK; COOP; CONSOLI, 2005). Com o aumento das deformações, desenvolve-se a transferência de carga da matriz para as fibras através de tensões cisalhantes na interface, resultando em um enrijecimento aparente do material. Este processo continua até a mobilização plena do sistema solo-fibra, que só ocorre a deformações mais elevadas (MICHALOWSKI; ČERMÁK, 2003; LI; ZORNBERG, 2013).

O mecanismo de reforço em solos granulares é governado essencialmente pela interação por atrito entre os grãos e as fibras. Sob carregamento, o solo tende a se deformar lateralmente por tração, e é nessa tendência de movimento que as fibras são solicitadas, absorvendo parte dos esforços de tração e gerando tensões cisalhantes na interface solo-fibra (DIAMBRA; IBRAIM, 2015).

Uma interpretação amplamente adotada para esse mecanismo, no contexto do ensaio triaxial, descreve a ação das fibras como um acréscimo de tensão confinante efetiva. Ao serem tracionadas por uma força T_f , as fibras reagem comprimindo a matriz, o que se traduz em uma tensão de confinamento adicional $\sigma_R = \Delta\sigma_3$ entre as partículas do solo. Dessa forma, para uma mesma tensão confinante externa aplicada, o estado de tensão efetiva no interior do compósito é superior ao do solo sem reforço, o que ajuda a explicar o ganho de resistência observado (CONCEIÇÃO, 2021).

Esse mecanismo depende diretamente do nível de confinamento. Tensões de confinamento mais elevadas favorecem a ancoragem das fibras e a mobilização das forças de tração, ao passo que, sob confinamento muito baixo, a ancoragem é limitada e o efeito do reforço é reduzido (CONCEIÇÃO, 2021). Essa dependência tem implicação direta na interpretação dos resultados do presente trabalho, conforme discutido no Capítulo 4.

2.2 Modelagem Constitutiva de Solos

2.2.1 Modelos clássicos: capacidades e limitações

A modelagem constitutiva estabelece as relações matemáticas entre tensões e deformações que governam o comportamento mecânico dos solos. Os modelos clássicos, desenvolvidos com base em princípios da mecânica do contínuo, incorporam conceitos fundamentais como elasticidade, plasticidade e estado crítico (WOOD, 2004).

O modelo de Mohr-Coulomb, amplamente utilizado na prática geotécnica, baseia-se em um critério de ruptura linear que relaciona a tensão cisalhante e a tensão normal por meio de dois parâmetros: coesão (c) e ângulo de atrito (ϕ). Apesar de sua simplicidade e facilidade de calibração, o modelo apresenta limitações significativas, incluindo a previsão de comportamento elástico e perfeitamente plástico, singularidades numéricas nos vértices da superfície de plastificação hexagonal, e a incapacidade de representar endurecimento ou amolecimento (POTTS; ZDRAVKOVIC, 1999).

O modelo Cam-Clay e sua versão modificada representam um avanço conceitual significativo ao incorporar o conceito de estado crítico. A formulação utiliza uma superfície de plastificação elíptica que evolui com deformações volumétricas plásticas:

$$F = \left(\frac{q}{M} \right)^2 + p'(p' - p'_c) = 0, \quad (1)$$

em que M é a inclinação da linha de estado crítico, p'_c a tensão de pré-adensamento, p' a tensão média efetiva e q a tensão desviadora. O modelo captura características essenciais do comportamento das argilas, incluindo a transição entre o comportamento contrátil e dilatante, com apenas cinco parâmetros de significado físico claro (ROSCOE; BURLAND, 1968).

Desenvolvimentos posteriores estenderam a estrutura do estado crítico para abranger comportamentos mais complexos. O Clay-and-Sand Model (CASM), proposto por Yu (1998), representa um marco importante na unificação da modelagem constitutiva de solos. O modelo incorpora o conceito de parâmetro de estado (ψ), definido como:

$$\psi = e - e_c, \quad (2)$$

em que e é o índice de vazios atual e e_c é o índice de vazios na linha de estado crítico para a mesma tensão média. A superfície de plastificação do CASM é expressa por:

$$f = \left(\frac{q}{Mp'} \right)^n + \frac{1}{\ln r} \ln \left(\frac{p'}{p'_0} \right) = 0, \quad (3)$$

em que p'_0 é a tensão de pré-adensamento, que define o tamanho da superfície, n é o parâmetro que controla a forma da superfície e r é a razão de espaçamento (*spacing ratio*), que posiciona a linha de estado crítico em relação à superfície de plastificação. Esta formulação permite tanto a expansão quanto a contração da superfície de plastificação, possibilitando representar o comportamento de argilas normalmente adensadas, sobreadensadas e areias dentro de uma estrutura única. A incorporação do parâmetro de estado (ψ) à formulação captura a influência da densidade relativa no comportamento do material, permitindo que o modelo preveja adequadamente a transição entre o comportamento contrátil e dilatante observada em solos granulares compactos e argilas sobreadensadas.

O modelo NorSand, desenvolvido por Jefferies (1993) e posteriormente refinado por Jefferies e Been (2015), constitui uma contribuição significativa à modelagem do estado crítico, com foco particular em areias. O modelo fundamenta-se na premissa de que o comportamento mecânico de solos granulares é controlado primariamente pelo parâmetro de estado (ψ), definido no espaço e - $\ln p'$ como:

$$\psi = e - e_{cs} = e - (\Gamma - \lambda \ln p'), \quad (4)$$

em que Γ é o índice de vazios na linha de estado crítico para $p' = 1$ kPa e λ é a inclinação da linha de estado crítico. Diferentemente do Cam-Clay modificado, a superfície de plastificação do NorSand deriva da formulação original do Cam-Clay, e o modelo abandona a hipótese de que os estados do solo repousem sobre uma única superfície de fronteira, admitindo uma infinidade de linhas de adensamento normal. A superfície de plastificação evolui em torno de um estado-imagem próprio, e a lei de endurecimento, controlada pelo parâmetro de estado, impõe um limite à expansão da superfície. É esse limite que permite capturar características essenciais do comportamento de areias, como o pico de resistência e o amolecimento pós-pico em areias compactas ($\psi < 0$) e a tendência à liquefação em areias fofas ($\psi > 0$) sob carregamento não-drenado. A capacidade do modelo de reproduzir comportamentos sob diferentes trajetórias de tensão e condições de drenagem, com um conjunto reduzido de parâmetros, tem levado à sua adoção crescente em projetos geotécnicos complexos, particularmente em análises de estabilidade de barragens de rejeitos e na avaliação do potencial de liquefação.

Modelos mais avançados, como o *Hardenig Soil* e o *Hypoplastic*, introduzem superfícies de plastificação múltiplas e leis de endurecimento complexas para melhor representar o comportamento observado. Entretanto, o aumento na sofisticação vem acompanhado de um maior número de parâmetros (tipicamente de 10 a 15) e de dificuldades crescentes

de calibração e implementação numérica (SCHANZ; VERMEER; BONNIER, 1999). A proliferação de parâmetros torna-se ainda mais problemática quando se busca estender esses modelos para materiais compósitos, como solos reforçados com fibras, nos quais a heterogeneidade e a anisotropia induzidas pelo reforço introduzem complexidades adicionais que desafiam a estrutura conceitual dos modelos tradicionais.

2.2.2 Desafios na modelagem de solos reforçados

A aplicação de modelos constitutivos clássicos a solos reforçados com fibras revela limitações fundamentais que comprometem a qualidade das previsões. A heterogeneidade inerente ao material compósito, com propriedades mecânicas drasticamente diferentes entre a matriz e as fibras, não pode ser adequadamente representada por modelos desenvolvidos para meios homogêneos (DIAMBRA; IBRAIM, 2015).

A mobilização progressiva das fibras introduz não-linearidades que transcendem aquelas capturadas por modelos elastoplásticos convencionais. Enquanto modelos tradicionais assumem comportamento isotrópico ou leis de evolução simples, solos reforçados exibem resistência anisotrópica, dependente da orientação e da distribuição das fibras no interior da massa de solo (MICHALOWSKI, 2008).

Tentativas de adaptar modelos clássicos para incorporar o efeito das fibras resultam em formulações com um número excessivo de parâmetros. No contexto de resíduos sólidos urbanos, cuja matriz fibrosa é análoga ao reforço, Machado, Carvalho e Vilar (2002) propuseram uma formulação do tipo Cam-Clay modificado para a fase pasta, acoplada a uma função de mobilização das fibras, ao custo de um número expressivo de parâmetros adicionais. Semelhantemente, formulações que acoplam a contribuição das fibras a modelos da matriz de solo, como a de Diambra et al. (2010), fundada no modelo Mohr-Coulomb, exigem a caracterização da distribuição de orientação das fibras e a definição de parâmetros adicionais de interação, de difícil determinação experimental.

A proliferação de parâmetros cria problemas práticos significativos. A calibração requer programas experimentais extensivos, cobrindo múltiplas trajetórias de tensão e configurações de reforço, com custos e tempos frequentemente limitativos. Mesmo quando dados adequados estão disponíveis, a não-unicidade na determinação de parâmetros compromete a confiabilidade das previsões. Diferentes combinações de parâmetros podem produzir ajustes de qualidade comparável aos dados de calibração, como demonstrado em estudos de identificação de parâmetros por análise inversa (CALVELLO; FINNO, 2004).

A modelagem de solos reforçados com fibras tem sido abordada por meio de diferentes estratégias analíticas, cada uma com suas vantagens e limitações específicas. Modelos de homogeneização buscam derivar propriedades efetivas do compósito por meio de técnicas de média volumétrica, considerando a distribuição e orientação das fibras (MICHALOWSKI; ZHAO, 1996). Embora teoricamente rigorosos, esses modelos requerem

informações detalhadas sobre a microestrutura, as quais são difíceis de obter experimentalmente.

O modelo analítico recente de Machado et al. (2024) representa um avanço significativo ao propor uma abordagem bifásica com a modificação do Cam-Clay modificado. O modelo trata o solo reforçado como um material composto de duas fases, a pasta (matriz de solo) e as fibras, cada qual com seu próprio comportamento constitutivo. Para a pasta, utiliza-se a superfície de plastificação:

$$f_p = q_p - M [p'^n (p_o - p')]^{\frac{1}{1+n}} n^{\frac{1}{1+n}} = 0, \quad (5)$$

em que q_p é a tensão desviadora na pasta, M é a inclinação da linha de estados críticos, p' é a tensão média efetiva, p_o é a tensão de pré-adensamento e n controla a forma da superfície.

Os quatro parâmetros adicionais aparecem em equações complementares do modelo:

- a_{ff} , na equação do módulo elástico das fibras: $E_f = E_u + a_{ff} \cdot \sigma_r$;
- n_{g2} , na lei de fluxo não-associada para controle da dilatação;
- a_d , no parâmetro que controla a transição entre comportamentos contrátil e dilatante;
- f_m , função de mobilização que relaciona as deformações das fibras com as da pasta.

A tensão desviadora total é obtida pela soma ponderada das contribuições de cada fase, incorporando explicitamente o efeito do teor de fibras, comprimento e confinamento, por meio de funções empiricamente calibradas. O modelo oferece um balanço favorável entre precisão e aplicabilidade prática, com parâmetros de clara interpretação física, permitindo a reprodução do comportamento de areias reforçadas com diferentes teores e comprimentos de fibra. Embora represente um progresso significativo, o modelo ainda apresenta limitações na captura de comportamentos complexos, como a anisotropia evolutiva. É precisamente esse modelo que se adota, nesta pesquisa, como gerador dos dados sintéticos de pré-treinamento, conforme detalhado no capítulo seguinte.

2.3 Redes Neurais Artificiais em Geotecnia

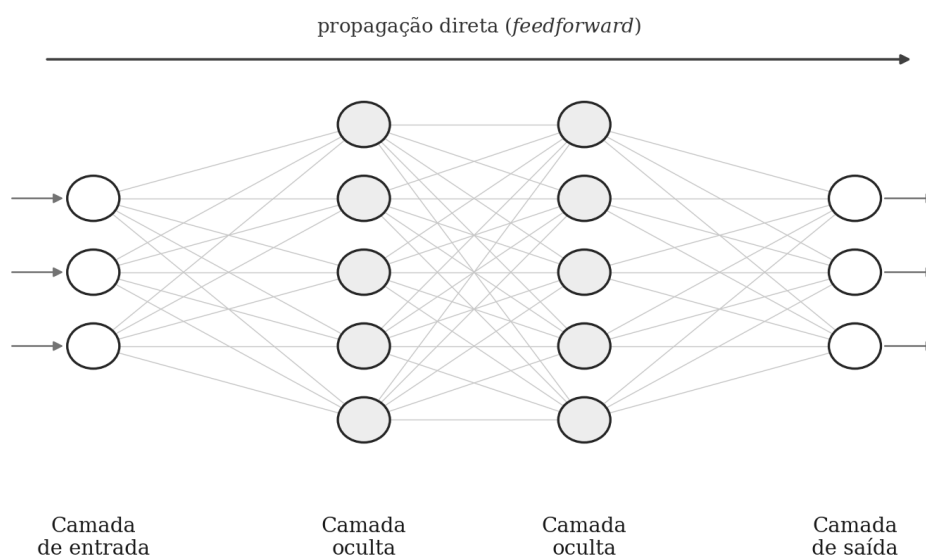
2.3.1 Fundamentos das redes neurais artificiais

Antes de discutir as aplicações das redes neurais à geotecnia, convém apresentar de forma acessível os elementos que as constituem. As Redes Neurais Artificiais (RNA) são sistemas computacionais inspirados na organização do cérebro humano, formados por unidades de processamento simples, os neurônios artificiais, conectadas entre si (HAYKIN,

2009). Cada neurônio recebe um conjunto de valores de entrada, multiplica cada um deles por um peso, soma os produtos, acrescenta um termo constante denominado viés e submete o resultado a uma função de ativação, que define o valor de saída do neurônio. Os pesos quantificam a importância relativa de cada entrada, e é no ajuste desses pesos que a rede armazena o conhecimento extraído dos dados.

Os neurônios organizam-se em camadas. Uma camada de entrada recebe as variáveis do problema, uma ou mais camadas ocultas processam a informação de forma intermediária, e uma camada de saída fornece a resposta. Quando os neurônios de cada camada se conectam a todos os da camada seguinte e a informação flui em sentido único, da entrada para a saída, a arquitetura é denominada perceptron multicamadas (*Multilayer Perceptron*, MLP), do tipo de propagação direta (*feedforward*). O número de camadas ocultas e de neurônios em cada uma define a capacidade da rede de representar relações complexas (HAYKIN, 2009). A Figura 1 ilustra essa organização em camadas.

Figura 1 – Estrutura de um perceptron multicamadas: neurônios organizados em uma camada de entrada, camadas ocultas e uma camada de saída, com conexões totais entre camadas sucessivas e fluxo de informação em sentido único



Fonte: Elaborado pelo autor.

A função de ativação cumpre papel decisivo: é ela que introduz a não-linearidade na rede. Sem uma função de ativação não-linear, a composição de várias camadas equivaleria a uma única transformação linear, incapaz de representar os fenômenos não-lineares que caracterizam o comportamento mecânico dos solos. É a presença dessas funções que confere à rede a capacidade de aproximar relações tensão-deformação complexas, com endurecimento, amolecimento e mudança de regime volumétrico.

O conhecimento de uma rede neural não é programado de forma explícita, mas aprendido a partir de exemplos. No aprendizado supervisionado, apresentam-se à rede pares de entrada e de saída conhecida; a rede produz uma previsão, e uma função de perda mede

a discrepância entre a previsão e o valor esperado. O treinamento consiste em ajustar iterativamente os pesos de modo a reduzir essa perda. O ajuste é conduzido pelo método do gradiente descendente (método do máximo declive), no qual cada peso é modificado na direção que mais reduz o erro, e o cálculo eficiente desses gradientes em redes de múltiplas camadas é realizado pelo algoritmo de retropropagação (*backpropagation*) (HAYKIN, 2009).

A justificativa teórica para o emprego de redes neurais na representação de relações arbitrariamente complexas repousa no teorema da aproximação universal, segundo o qual uma rede de propagação direta com uma única camada oculta, contendo um número finito de neurônios, é capaz de aproximar qualquer função contínua com a precisão desejada, resultado demonstrado originalmente para funções de ativação sigmóides (HORNIK; STINCHCOMBE; WHITE, 1989) e posteriormente generalizado para qualquer função de ativação não-polinomial (LESHNO et al., 1993). Essa propriedade assegura que, em princípio, uma RNA pode representar qualquer relação constitutiva, por mais intrincada que seja, o que a torna uma candidata natural à modelagem do comportamento dos solos.

2.3.2 Aplicações estabelecidas em geotecnia

Na geotecnia, as RNA têm demonstrado notável capacidade de modelar comportamentos não-lineares complexos que desafiam as abordagens analíticas tradicionais. Aplicações consolidadas incluem a previsão da capacidade de carga de fundações, a análise da estabilidade de taludes, a caracterização de solos e a interpretação de ensaios de campo. A revisão de Shahin, Jaksa e Maier (2001) documenta o desempenho superior das RNA em relação aos métodos tradicionais na previsão da capacidade de carga de estacas, atribuído à capacidade desses modelos de incorporar implicitamente, a partir dos dados, fatores de difícil quantificação analítica, como as condições de instalação.

Na estimativa de propriedades e comportamentos geotécnicos, os métodos de aprendizado de máquina têm superado de forma consistente as correlações empíricas univariadas. Na avaliação do potencial de liquefação a partir de registros de casos históricos de ensaios de penetração de cone (*Cone Penetration Test*, CPT), por exemplo, esses modelos alcançaram elevadas taxas de classificação correta, ao capturarem interações entre múltiplos parâmetros que as correlações tradicionais não conseguem representar (GOH; GOH, 2007). Esse histórico de sucesso em problemas de natureza preditiva fundamenta a extensão da abordagem ao problema, mais exigente, da modelagem constitutiva.

2.3.3 Modelos constitutivos baseados em RNA

O emprego de redes neurais na modelagem constitutiva propriamente dita foi inaugurado pelo trabalho de Ghaboussi, Jr e Wu (1991), que demonstrou ser possível aprender relações constitutivas diretamente de dados experimentais, sem a necessidade de postular

leis matemáticas explícitas ou de identificar mecanismos de deformação específicos. Com uma rede de propagação direta de formulação incremental, que recebe o estado corrente e um incremento de deformação e devolve o incremento de tensão correspondente, os autores reproduziram o comportamento do concreto simples sob carregamento biaxial, monotônico e cíclico, capturando a não-linearidade da resposta e a sua dependência da trajetória de tensões, sem qualquer informação prévia sobre a forma da curva tensão-deformação.

Desenvolvimentos posteriores ampliaram as capacidades dos modelos neurais constitutivos e os integraram a análises numéricas. Hashash, Jung e Ghaboussi (2004) implementaram redes neurais em análises de elementos finitos, propondo uma metodologia para a extração automática da matriz tangente consistente por meio da diferenciação analítica da rede, etapa indispensável à solução de problemas de valor de contorno. A abordagem foi aplicada pelo grupo à simulação de escavações profundas, com o modelo neural em substituição ao modelo constitutivo convencional na análise por elementos finitos.

Para materiais geotécnicos específicos, os modelos neurais têm-se mostrado capazes de capturar comportamentos de difícil representação analítica. Penumadu e Zhao (1999) desenvolveram redes neurais para areias e pedregulhos sob carregamento triaxial que reproduzem transições suaves entre os regimes contrátil e dilatante. A possibilidade de incorporar múltiplas variáveis de estado sem ampliar drasticamente a complexidade matemática constitui uma vantagem fundamental dessas abordagens em relação aos modelos fenomenológicos.

Avanços mais recentes buscam incorporar princípios físicos diretamente à arquitetura ou ao treinamento da rede. Masi et al. (2021) propuseram redes neurais artificiais baseadas na termodinâmica (*Thermodynamics-based Artificial Neural Networks*, TANNs), que asseguram a consistência termodinâmica ao estruturar a rede para aprender potenciais termodinâmicos dos quais derivam as relações constitutivas, com ganho na capacidade de extrapolação. Essas formulações ilustram uma tendência mais ampla de aporte de conhecimento físico aos modelos baseados em dados, da qual a estratégia de pré-treinamento adotada neste trabalho constitui uma alternativa, conforme se discute na sequência.

2.3.4 Vantagens da abordagem baseada em dados e a oportunidade para os solos reforçados com fibras

A modelagem constitutiva baseada em dados oferece vantagens específicas sobre as abordagens analíticas, particularmente relevantes para os geocompósitos. A primeira é a flexibilidade na incorporação de novas variáveis. Enquanto a inclusão de um novo fator em um modelo analítico exige a reformulação das equações governantes e a introdução de parâmetros adicionais, uma RNA pode incorporar variáveis como o teor, o comprimento e a orientação das fibras pela simples ampliação da camada de entrada, sem alteração de sua estrutura matemática. Essa adaptabilidade favorece o desenvolvimento incremental

do modelo à medida que novos dados se tornam disponíveis (JAVADI; REZANIA, 2009).

A segunda vantagem é a capacidade de capturar relações não-lineares arbitrariamente complexas sem a especificação prévia de uma forma funcional. A rede aprende os padrões diretamente dos dados, incluindo interações sutis que os modelos fenomenológicos, presos a uma forma matemática previamente fixada, dificilmente representam (ZHANG et al., 2021).

A terceira vantagem é a eficiência computacional após o treinamento. Uma vez treinada, a avaliação de uma RNA reduz-se a operações matriciais simples, de custo desprezível, o que viabiliza análises iterativas e estudos de otimização que seriam proibitivos com simulações numéricas detalhadas (STOFFEL; BAMER; MARKERT, 2018).

Soma-se às anteriores uma quarta vantagem, central para esta pesquisa: a combinação entre o pré-treinamento e o aprendizado por transferência (*transfer learning*), que permite desenvolver modelos eficientes mesmo quando os dados experimentais são escassos. O aprendizado por transferência consiste em aproveitar o conhecimento adquirido por uma rede em uma tarefa, ou em um conjunto de dados, para acelerar e aprimorar o aprendizado em outra (PAN; YANG, 2010). Esse princípio responde a um obstáculo característico da geotecnia: a obtenção de dados experimentais de qualidade é onerosa e demorada, pois exige programas extensos de ensaios triaxiais, o que restringe severamente o número de exemplos disponíveis para o treinamento de redes neurais, que tipicamente demandam grandes volumes de dados.

Reside precisamente nesse ponto a lacuna metodológica que o presente trabalho explora. A estratégia consiste em treinar a rede, em uma primeira fase, com um grande volume de dados sintéticos gerados por um modelo analítico, fisicamente consistentes e de obtenção barata, e, em uma segunda fase, refiná-la, por ajuste fino (*fine-tuning*), com os dados experimentais de laboratório, mais escassos, porém mais precisos. O pré-treinamento sintético incorpora à rede o conhecimento físico estabelecido antes do contato com o conjunto experimental restrito, de modo que o ajuste fino parte de um estado já fisicamente informado, e não de uma inicialização aleatória. Essa estratégia compartilha a motivação das redes neurais informadas pela física, como as de Masi et al. (2021), mas aporta o conhecimento físico por uma via distinta, pelos dados de pré-treinamento, em vez de impô-lo às equações da função de perda.

A despeito do sucesso das RNA em diversos problemas geotécnicos, sua aplicação a solos reforçados com fibras tem-se concentrado na previsão de parâmetros mecânicos isolados, como a resistência ao cisalhamento e o módulo de elasticidade do compósito (CHOU; YANG; LIN, 2016; OWUSU-ANSAH et al., 2022), e não na construção de um modelo constitutivo capaz de reproduzir a resposta tensão-deformação completa. A complexidade inerente ao comportamento desses geocompósitos, com a mobilização progressiva do reforço, a anisotropia induzida e a interação em múltiplas escalas, torna-os candidatos particularmente adequados às abordagens baseadas em dados, capazes de capturar

padrões complexos sem o recurso a formulações explícitas.

A disponibilidade de um conjunto experimental sistemático e de alta qualidade, como o obtido por Conceição, Carvalho e Machado (2022) para uma areia uniforme reforçada com fibra polimérica, cria uma oportunidade concreta para o desenvolvimento de modelos neurais robustos. Sua conjugação com um modelo analítico de referência calibrado para o mesmo material (MACHADO et al., 2024) viabiliza, ao mesmo tempo, o ajuste fino com dados experimentais confiáveis e a geração controlada de dados sintéticos de pré-treinamento, o que assegura que ambas as fases do treinamento descrevam rigorosamente o mesmo solo e a mesma fibra.

É justamente essa lacuna, a ausência de um modelo constitutivo neural capaz de prever a resposta tensão-deformação completa da areia reforçada com fibras aqui estudada, treinado sob a estratégia híbrida de pré-treinamento sintético e ajuste fino experimental, que a presente pesquisa busca preencher. O capítulo seguinte detalha a metodologia adotada para tanto, da geração dos dados sintéticos à arquitetura da rede e à estratégia de treinamento, bem como as métricas empregadas em sua validação.

Materiais e Métodos

Este capítulo apresenta a metodologia empregada no desenvolvimento do modelo constitutivo baseado em Redes Neurais Artificiais para solos arenosos reforçados com fibras. A abordagem fundamenta-se em uma estratégia híbrida de aprendizado, na qual a rede é primeiramente pré-treinada com um grande volume de dados sintéticos fisicamente consistentes, gerados por um modelo analítico, e posteriormente ajustada, por aprendizado por transferência, a um conjunto de dados experimentais de alta qualidade. O capítulo descreve a base experimental utilizada, o processo de geração dos dados sintéticos, o tratamento dos dados, a arquitetura da rede, a estratégia de treinamento e as métricas de avaliação.

3.1 Base de Dados Experimentais

3.1.1 Origem e descrição dos dados

Os dados experimentais utilizados nesta pesquisa originam-se do programa de ensaios triaxiais desenvolvido por Conceição, Carvalho e Machado (2022) no Laboratório de Geotecnia Ambiental (GEOAMB) da Universidade Federal da Bahia (UFBA), descrito em detalhe em Conceição (2021). Esse mesmo conjunto experimental subsidiou a calibração do modelo constitutivo analítico de Machado et al. (2024), o que torna a base particularmente adequada à estratégia híbrida adotada, pois permite que os dados sintéticos de pré-treinamento e os dados experimentais de ajuste fino descrevam rigorosamente o mesmo material.

A investigação experimental estudou uma areia de granulometria uniforme coletada no bairro de Stella Maris, na região metropolitana de Salvador, na Bahia. A escolha de uma areia uniforme deve-se ao interesse em isolar o efeito do reforço fibroso em uma matriz granular de comportamento bem definido.

O programa experimental foi integralmente planejado e executado por Conceição, Carvalho e Machado (2022), e cabe frisar que a presente pesquisa não envolveu a realização

de ensaios de laboratório: todo o trabalho experimental descrito nesta seção é de autoria daqueles pesquisadores. O programa compreendeu ensaios de compressão triaxial consolidados isotropicamente drenados (CID), executados em corpos de prova de dimensões superiores às convencionais, com 100 mm de diâmetro e 200 mm de altura, buscando resultados mais representativos de um maciço de solo reforçado (CONCEIÇÃO; CARVALHO; MACHADO, 2022). O escopo abrangeu sistematicamente quatro fatores de ensaio:

- ❑ teores de fibra de 0%, 0,5% e 1,0% em relação ao peso seco do solo;
- ❑ comprimentos de fibra de 12,5 mm, 25 mm e 51 mm;
- ❑ tensões de confinamento de 50, 100, 200 e 300 kPa;
- ❑ trajetórias de tensão com razões dq/dp iguais a 3 (trajetória convencional), 1,5, 10 e -3 .

As trajetórias não convencionais foram executadas apenas para a tensão de adensamento de 100 kPa. Cada combinação foi ensaiada em duplicata, totalizando 60 ensaios triaxiais. Cabe antecipar que, das quatro razões dq/dp contempladas pelo programa, a presente pesquisa emprega exclusivamente a trajetória convencional ($dq/dp = 3$); a justificativa metodológica dessa restrição, ligada à coerência com os dados sintéticos de pré-treinamento, é detalhada na Seção 3.1.3.

Os corpos de prova foram moldados em molde bipartido, com aplicação de sucção para a retirada do molde, e os procedimentos gerais seguiram as práticas consagradas para o ensaio triaxial descritas em detalhe por Conceição (2021). A qualidade da moldagem foi avaliada pelo coeficiente de permeabilidade, medido após a saturação por fluxo ascendente. Em seguida, procedeu-se à saturação por contrapressão, com verificação do parâmetro B de Skempton superior a 0,90, ao adensamento isotrópico sob a tensão efetiva desejada e ao cisalhamento drenado a uma taxa de carregamento de 0,667 mm/min (CONCEIÇÃO; CARVALHO; MACHADO, 2022).

3.1.2 Características dos materiais ensaiados

3.1.2.1 Propriedades da areia

A areia foi caracterizada por meio de ensaios geotécnicos padronizados, incluindo análise granulométrica (NBR 7181), massa específica dos sólidos (NBR 6458) e determinação dos índices de vazios máximo (NBR 12004) e mínimo (NBR 12051). As propriedades são resumidas na Tabela 1. Segundo o Sistema Unificado de Classificação de Solos (SUCS), a areia é classificada como SP, areia mal graduada, com predominância da fração média e baixo coeficiente de uniformidade, características de um material muito uniforme (CONCEIÇÃO, 2021).

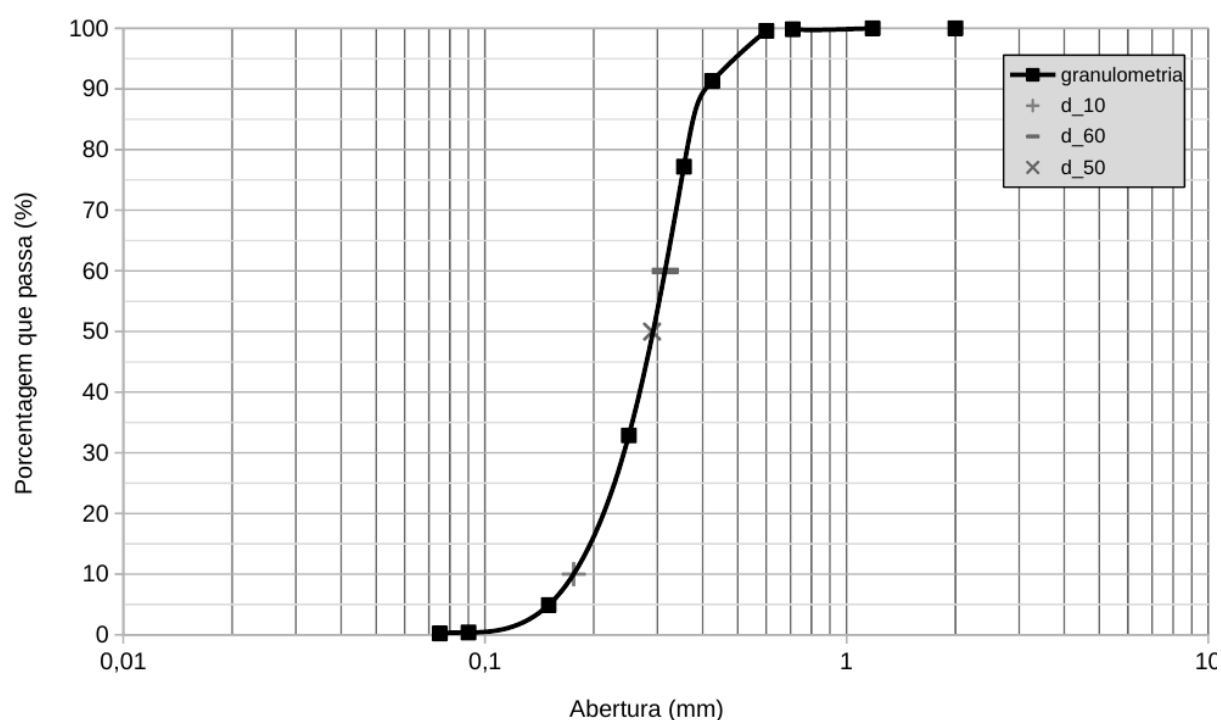
Tabela 1 – Propriedades índice da areia uniforme

Propriedade	Valor
Classificação SUCS	SP
Massa específica dos sólidos, ρ_s (g/cm ³)	2,657
D_{10} (mm)	0,176
D_{50} (mm)	0,29
Coeficiente de uniformidade, C_U	1,79
Coeficiente de curvatura, C_C	1,23
Índice de vazios máximo, e_{max}	0,748
Índice de vazios mínimo, e_{min}	0,523
Densidade relativa dos ensaios, D_r (%)	60
Índice de vazios de moldagem, e_0	0,613
Massa específica seca de moldagem, ρ_d (g/cm ³)	1,647

Fonte: Conceição (2021).

A curva granulométrica da areia é apresentada na Figura 2, obtida do peneiramento de uma amostra de 30 kg. Trata-se de uma areia média, com pequena fração fina, sem material de granulometria grossa, o que confirma a sua uniformidade.

Figura 2 – Curva granulométrica da areia uniforme



Fonte: Conceição (2021).

A densidade relativa de 60% foi mantida constante em todos os corpos de prova, o que corresponde a um índice de vazios de moldagem $e_0 = 0,613$ e a uma massa específica seca de 1,647 g/cm³. Por se tratar de uma areia de grãos uniformes e arredondados, a faixa entre os índices de vazios máximo e mínimo é estreita, de modo que pequenas variações de compactidade, ocorridas durante a moldagem ou durante o próprio ensaio, podem alterar

de forma sensível a resposta mecânica observada. O material apresentou comportamento típico de areia compacta, com dilatação durante o cisalhamento (CONCEIÇÃO, 2021).

3.1.2.2 Características das fibras de reforço

As fibras utilizadas consistem em monofilamentos de copolímero de polipropileno e polietileno, quimicamente inertes, fornecidas comercialmente. As propriedades físicas e mecânicas reportadas são apresentadas na Tabela 2. Foram empregados três comprimentos de fibra (12,5 mm, 25 mm e 51 mm), com teores de 0,5% e 1,0% em relação ao peso seco do solo, além da condição de referência sem reforço.

Tabela 2 – Propriedades das fibras poliméricas

Propriedade	Valor
Composição	Copolímero PP/PE
Massa específica, ρ_f (g/cm ³)	0,91
Módulo de elasticidade, E_f (GPa)	5,0
Resistência à tração, σ_{tf} (MPa)	600
Diâmetro, d_f (mm)	0,51
Comprimentos utilizados, L (mm)	12,5; 25; 51
Índice de aspecto, L/d_f	24,5; 49; 100

Fonte: Conceição (2021).

As propriedades indicadas na Tabela 2 são as informadas pelo fabricante. A contribuição mecânica efetiva do reforço no comportamento do compósito é governada pelos parâmetros de interação fibra-solo calibrados do modelo analítico, apresentados na Seção 3.2.

Na preparação dos compósitos, as fibras foram tratadas como parte da fase sólida, calculando-se a massa específica dos sólidos ponderada para cada teor de fibra, de modo a manter o índice de vazios de moldagem constante em $e_0 = 0,613$, seguindo o procedimento de Diambra et al. (2010).

3.1.3 Seleção e caracterização da base utilizada

Do conjunto completo de ensaios, esta pesquisa utilizou exclusivamente os ensaios de compressão triaxial drenada convencional, isto é, aqueles conduzidos sob a trajetória de tensões $dq/dp = 3$. Essa escolha é metodologicamente necessária, pois o pré-treinamento sintético precisa ser coerente com os dados experimentais de ajuste fino. Como o gerador de dados sintéticos reproduz a trajetória drenada convencional, a inclusão de ensaios sob trajetórias não convencionais, sem contraparte sintética equivalente, introduziria um descasamento de domínio entre as duas fases de treinamento. As trajetórias não convencionais (dq/dp iguais a 1,5, 10 e -3) foram, portanto, reservadas como proposta de trabalho futuro.

A base experimental efetivamente empregada compreende 24 ensaios triaxiais drenados convencionais, distribuídos em três classes de reforço: 4 ensaios em areia sem fibra, como referência; 8 ensaios com teor de fibra de 0,5% (comprimentos de 25 mm e 51 mm); e 12 ensaios com teor de fibra de 1,0% (comprimentos de 12,5 mm, 25 mm e 51 mm). Cada ensaio cobre as quatro tensões de confinamento (50, 100, 200 e 300 kPa). O conjunto soma 1.207 pontos experimentais de relações tensão-deformação-volume, com índice de vazios de moldagem constante ($e_0 = 0,613$). Como cada combinação foi ensaiada em duplicata no programa experimental, utilizou-se a curva representativa de cada uma das 24 combinações, conforme disponibilizada por Conceição (2021). A opção por essa base, e não por bases de outras areias, como a areia aluvionar mais graduada de Pinto (2021), inicialmente cogitada e cuja efetividade de reforço foi comparada à da presente areia por Conceição et al. (2023), deve-se à coerência com o modelo analítico de referência: os ensaios de Conceição (2021), Conceição, Carvalho e Machado (2022) referem-se ao mesmo material para o qual o modelo de Machado et al. (2024) foi calibrado, o que assegura que as fases de pré-treinamento e de ajuste fino descrevam rigorosamente o mesmo solo e a mesma fibra. A composição dessa base está resumida na Tabela 3.

Tabela 3 – Composição da base experimental utilizada, em ensaios triaxiais drenados convencionais

Teor de fibra	Comprimento de fibra	Confinamentos (kPa)	Ensaio
0% (referência)	-	50, 100, 200, 300	4
0,5%	25 mm	50, 100, 200, 300	4
0,5%	51 mm	50, 100, 200, 300	4
1,0%	12,5 mm	50, 100, 200, 300	4
1,0%	25 mm	50, 100, 200, 300	4
1,0%	51 mm	50, 100, 200, 300	4
Total			24

Fonte: Elaborado pelo autor, com dados de Conceição (2021).

Para cada ponto registrado ao longo do cisalhamento, a base contém a tensão média efetiva (p'), a tensão desviadora (q), a deformação axial (ε_a) e a deformação volumétrica (ε_v), além dos parâmetros de configuração do ensaio (tensão de confinamento, teor de fibra e comprimento de fibra). A tensão média efetiva em ensaio drenado obedece à relação $p' = \sigma_3 + q/3$, e o índice de vazios instantâneo é obtido por $e = e_0 - \varepsilon_v (1 + e_0)$, adotando-se a convenção clássica da mecânica dos solos, em que deformações volumétricas positivas correspondem a compressão e negativas a dilatação, de forma consistente com a redução do índice de vazios sob compressão expressa pela relação anterior.

3.2 Geração de Dados Sintéticos

A geração de dados sintéticos constitui um elemento central da metodologia, pois fornece o grande volume de exemplos fisicamente consistentes necessário ao pré-treinamento da rede. O modelo adotado para essa finalidade é a adaptação do Cam-Clay modificado proposta por Machado et al. (2024) para solos arenosos reforçados com fibras.

Nessa formulação, o compósito é tratado como a superposição de duas fases distintas: a matriz de solo, constituída pela areia, e o reforço, constituído pelas fibras, cada qual com o seu próprio comportamento constitutivo. O termo fase é aqui empregado no sentido da mecânica dos materiais compósitos: as fibras formam um constituinte distinto da matriz, com rigidez e resposta mecânica próprias, e ocupam uma fração do volume total do conjunto.

A tensão do compósito resulta da combinação das contribuições das duas fases, ponderadas pelas respectivas frações volumétricas, conforme a partição de tensões apresentada na Tabela 4. Essa decomposição permite capturar a evolução da interação fibra-matriz ao longo do processo de deformação.

3.2.1 Formulação do modelo constitutivo

A formulação do modelo é apresentada de forma sintética na Tabela 4, com a descrição de seus parâmetros na Tabela 5. O modelo incorpora tanto aspectos elastoplásticos clássicos quanto mecanismos específicos de mobilização progressiva do reforço fibroso.

Tabela 4 – Equações constitutivas do modelo Cam-Clay modificado adaptado para solos reforçados com fibras

Componente	Formulação	Eq.
Superfície de plastificação	$f_p = q_p - M [p'^n(p_o - p')]^{\frac{1}{1+n}} n^{\frac{1}{1+n}} = 0$	(1)
	Para $n = 1$: $f = q^2 + M^2 p'(p' - p_o) = 0$	(2)
Inclinação da LEC	$M = \frac{6 \sin \phi}{3 - \sin \phi}$	(3)
Partição de tensões	$dq = 3G_f de_s^f V_f + 3f_m de_s^s G_s V_s$	(4)
Função de mobilização	$f_m = \frac{2}{\pi(1 + P_f(\%)/G_{sf}(\sigma_{atm}/p)^{0.5})} \tan^{-1} \left[\left(\frac{q}{p} \right)^2 \right]$	(5)
Lei de endurecimento	$\lambda = \lambda^* f(\eta_s, M_{pk}, \dot{p}_o) + \bar{\lambda} g(\eta_s, \eta_{smax})$	(6)
Parâmetro de estado	$\psi_{cr} = e_c - e_{cr}$	(7)

Fonte: Adaptado de Machado et al. (2024).

As equações que sintetizam o modelo estão reunidas na Tabela 4; as referências entre parênteses a seguir remetem à numeração própria dessa tabela, distinta da numeração

Tabela 5 – Parâmetros do modelo Cam-Clay modificado adaptado

Símbolo	Descrição	Tipo
<i>Parâmetros da areia</i>		
ϕ	Ângulo de atrito crítico	Solo
M	Inclinação da linha de estados críticos	Solo
λ^*	Inclinação da linha de compressão isotrópica	Solo
κ	Inclinação da linha de descarregamento	Solo
a_d	Parâmetro de dilatância	Solo
<i>Parâmetros do reforço</i>		
P_f	Teor de fibras (% em peso seco)	Fibra
L	Comprimento das fibras	Fibra
a_{ff}	Parâmetro de interação fibra-solo	Fibra
E_u	Módulo elástico inicial das fibras	Fibra
G_{sf}	Massa específica relativa das fibras	Fibra
<i>Variáveis de estado</i>		
p'	Tensão média efetiva	Estado
q	Tensão desviadora	Estado
q_p	Tensão desviadora na pasta (matriz)	Estado
p_o	Tensão de pré-adensamento	Estado
e_c	Índice de vazios atual	Estado
e_{cr}	Índice de vazios crítico	Estado
ψ_{cr}	Parâmetro de estado	Estado

Fonte: Adaptado de Machado et al. (2024).

sequencial das equações do corpo do texto. O elemento central é a superfície de plastificação (Equação 1), que define o limite entre os comportamentos elástico e plástico. Para o solo reforçado, após a calibração, o parâmetro n assume valor unitário, simplificando a formulação para a forma elíptica clássica do Cam-Clay modificado (Equação 2). A mobilização progressiva das fibras é capturada pela função f_m (Equação 5), que controla a partição das tensões entre a matriz e o reforço (Equação 4) e incorpora explicitamente o efeito do teor de fibras, da tensão de confinamento e da razão de tensões. Para deformações pequenas, o comportamento é dominado pela matriz; para deformações maiores, as fibras passam a contribuir significativamente para a resistência. O comportamento dilatante das areias compactas reforçadas é incorporado por meio do parâmetro de estado ψ_{cr} (Equação 7).

3.2.2 Calibração e geração das trajetórias sintéticas

Os parâmetros do modelo foram adotados integralmente da calibração conduzida por Machado et al. (2024) para esse mesmo material, por meio das planilhas de cálculo dos autores que implementam a formulação, e não foram recalibrados no âmbito desta pesquisa; estão reunidos na Tabela 6. Naquele trabalho, os parâmetros foram calibrados a partir do próprio programa experimental de Conceição, Carvalho e Machado (2022).

Os parâmetros da matriz arenosa decorrem da caracterização de estado crítico da areia, determinada experimentalmente por Conceição (2021), da qual se obtêm a reta de estado crítico, os ângulos de atrito e as inclinações λ^* e κ . Já os parâmetros efetivos de interação fibra-solo, a_{ff} , E_u e a resistência à tração efetiva, correspondem aos valores que ajustam a resposta do modelo de duas fases aos ensaios triaxiais drenados dos compósitos. A opção por adotar esses parâmetros, e não por recalibrá-los, é coerente com o papel atribuído ao modelo analítico nesta pesquisa, o de gerar dados sintéticos fisicamente consistentes para o pré-treinamento da rede.

A reta de estado crítico da areia foi determinada experimentalmente por Conceição (2021). No plano de tensões, a referida reta é dada por $q = M_{cv} p'$, com $M_{cv} = 1,22$, correspondente a um ângulo de atrito a volume constante $\phi_{cv} = 30,5^\circ$; no plano de compressão, o volume específico crítico segue

$$v_{cr} = N (p')^{-\lambda^*}, \quad (6)$$

com volume específico de referência $N = 1,83$ e $\lambda^* = 0,0164$. Esses valores são coerentes com a caracterização experimental da areia, cujo ângulo de atrito de pico situou-se em torno de 33° (CONCEIÇÃO, 2021), o que confere consistência física à base sintética.

O parâmetro de interação fibra-solo, a_{ff} , cresce com o comprimento da fibra, refletindo a maior eficiência de ancoragem das fibras mais longas. Cabe observar que o módulo elástico, a resistência à tração e o parâmetro de interação que governam a contribuição do reforço no modelo são valores efetivos, mobilizados no compósito, e, portanto, distintos das propriedades nominais da fibra isolada apresentadas na Tabela 2.

Com o modelo calibrado, foram geradas trajetórias sintéticas de compressão triaxial drenada convencional para 42 cenários distintos, resultantes da combinação de seis tensões de confinamento (50, 100, 150, 200, 250 e 300 kPa) com sete configurações de reforço (a areia sem fibra e os teores de 0,5% e 1,0% nos três comprimentos de fibra). A inclusão das tensões intermediárias de 150 e 250 kPa, ausentes do programa experimental, visa adensar a cobertura do espaço de estados e favorecer a interpolação da rede. Todas as trajetórias sintéticas reproduzem a mesma trajetória de tensões dos ensaios experimentais selecionados, a da compressão triaxial drenada convencional: a tensão de confinamento permanece constante e a tensão axial é aumentada de forma monotônica, fazendo a tensão média efetiva p' e a tensão desviadora q evoluírem segundo a razão $dq/dp' = 3$. Como o ajuste fino emprega exclusivamente ensaios sob essa trajetória (Seção 3.1.3), a geração sintética restringe-se a ela, de modo a manter a correspondência entre as duas fases de treinamento.

Um aspecto relevante da geração diz respeito à extensão das trajetórias simuladas. Em uma versão preliminar do procedimento de geração dos dados sintéticos, o parâmetro de tensão final da planilha era mantido fixo em 650 kPa para todos os cenários. Como consequência, os cenários de baixo confinamento eram truncados prematuramente, atingindo

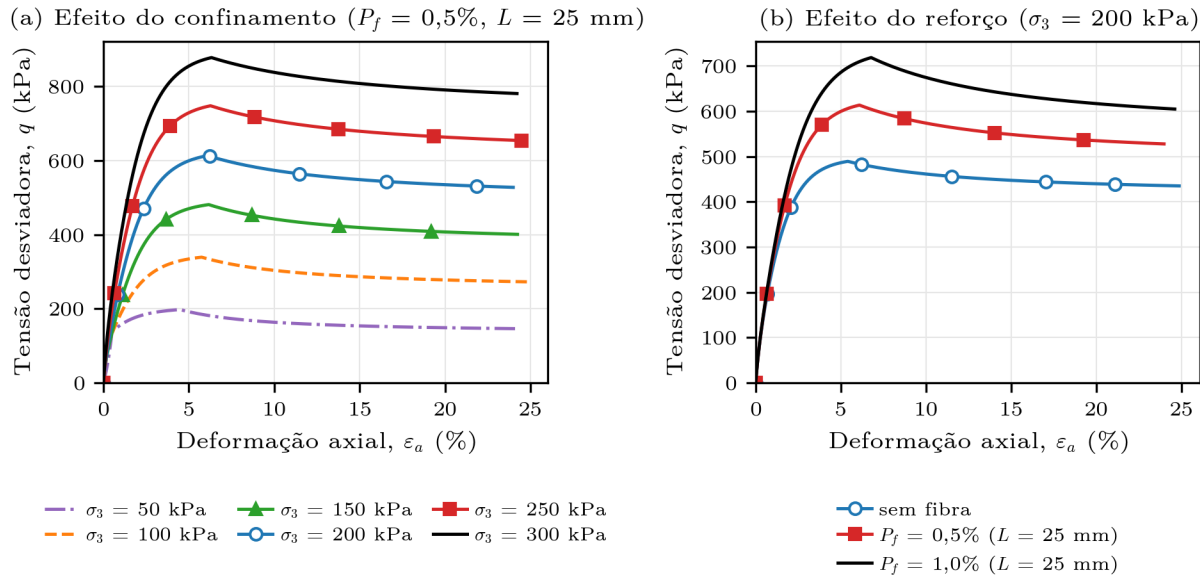
Tabela 6 – Parâmetros do modelo adotados na geração dos dados sintéticos

Parâmetro	Valor
<i>Matriz (areia)</i>	
Inclinação da reta de compressão isotrópica, λ^*	0,0164
Inclinação da reta de descarregamento, κ	0,0060
Ângulo de atrito de pico, ϕ_{pk} ($^\circ$)	33,1
Ângulo de atrito a volume constante, ϕ_{cv} ($^\circ$)	30,5
Inclinação da LEC no pico, M_{pk}	1,337
Inclinação da LEC a volume constante, M_{cv}	1,222
Volume específico de referência, N	1,83
Coefficiente de Poisson da matriz, ν_p	0,15
Parâmetro de dilatância, a_d	6,0
Expoente da lei de fluxo, n_{g2}	1,0
<i>Reforço (valores efetivos do modelo)</i>	
Interação fibra-solo, a_{ff} ($L = 12,5$ mm)	0,85
Interação fibra-solo, a_{ff} ($L = 25$ mm)	1,40
Interação fibra-solo, a_{ff} ($L = 51$ mm)	2,10
Módulo elástico inicial das fibras, E_u (MPa)	910
Resistência à tração efetiva (MPa)	225
Coefficiente de Poisson das fibras, ν_f	0,25

Fonte: Machado et al. (2024).

deformações axiais máximas da ordem de 2% a 3%, em vez dos cerca de 25% alcançados nos cenários de maior confinamento, o que produzia uma cobertura desigual do espaço de deformações. Esse parâmetro foi corrigido para uma proporção fixa da tensão de confinamento, igual a 2,17 vezes o valor de σ_3 , de modo a estender substancialmente a faixa de deformação axial coberta pelos cenários, sobretudo os de baixo confinamento, antes truncados em 2% a 3%. A correção praticamente dobrou o número de pontos sintéticos úteis, que passou a totalizar 19.555 pontos, distribuídos no intervalo de deformação axial de 0% a 25%. Esse conjunto, após a consolidação descrita na Seção 3.3, compõe a base de pré-treinamento da rede. A Figura 3 apresenta exemplos das trajetórias sintéticas geradas, ilustrando o efeito da tensão de confinamento e do reforço sobre a resposta tensão-deformação.

Figura 3 – Exemplos de trajetórias tensão-deformação sintéticas geradas pelo modelo analítico



3.3 Preparação e Tratamento dos Dados

A preparação dos dados é uma etapa central da metodologia, pois dela depende a qualidade da informação apresentada à rede em ambas as fases de treinamento. Os dados experimentais e os dados sintéticos descrevem o mesmo tipo de ensaio e as mesmas grandezas, o que permite consolidá-los em uma representação comum, na qual cada exemplo associa uma configuração de ensaio e um nível de deformação às variáveis de estado correspondentes; a definição completa das variáveis de entrada e de saída é apresentada na Seção 3.4.2. Esta seção descreve, na ordem, a estruturação e a limpeza dos dados, a variável que distingue a areia pura do compósito, a normalização das variáveis e a estratégia de validação adotada.

3.3.1 Estruturação e limpeza

Os dados experimentais e os dados sintéticos foram consolidados em uma estrutura tabular única, em que cada linha corresponde a um ponto de uma curva tensão-deformação. Cada registro reúne três grupos de informação: as variáveis de configuração do ensaio, isto é, a tensão de confinamento σ_3 , o teor de fibra P_f e o comprimento de fibra L ; as variáveis de estado, a saber, a tensão desviadora q , a tensão média efetiva p' , a deformação volumétrica ε_v , a deformação axial ε_a e o índice de vazios instantâneo e ; e os incrementos dessas grandezas entre pontos sucessivos da curva. Os incrementos não são empregados como entrada da rede, que adota um mapeamento direto (Seção 3.4.2), e servem exclusivamente ao procedimento de limpeza descrito a seguir.

Cabe precisar o sentido em que esses dois termos são empregados neste trabalho. As variáveis de configuração do ensaio são as condições nele impostas, definidas antes do

cisalhamento e mantidas constantes ao longo de toda a curva, e identificam, portanto, qual ensaio está sendo descrito. As variáveis de estado, por sua vez, são as grandezas que descrevem a condição mecânica corrente do corpo de prova em cada instante do carregamento: o nível de deformação já atingido e a resposta em tensão e em volume que lhe corresponde. A expressão é aqui utilizada no sentido corrente da mecânica dos solos, em que tensões, deformações e índice de vazios caracterizam o estado do material em um dado ponto da trajetória, e não no sentido estrito da termodinâmica do contínuo, em que as variáveis de estado são grandezas como a pressão, a temperatura e a entropia.

A limpeza foi conduzida segundo um critério de plausibilidade física, de natureza essencialmente de engenharia, e deliberadamente conservadora, de modo a preservar integralmente os comportamentos extremos, mas genuínos, do material. Nos dados sintéticos, foram removidos, em primeiro lugar, os pontos de preenchimento, identificados pela ausência simultânea de variação de deformação axial e de tensão desviadora, resíduos do preenchimento das planilhas de cálculo; em segundo lugar, restringiu-se a deformação axial à faixa de 0% a 25%, que corresponde ao intervalo fisicamente representado pelo ensaio. Dos 19.555 pontos gerados, 19.513 são retidos após essa limpeza. Nos dados experimentais, foram descartados apenas os poucos registros com retrocesso anômalo da deformação axial, identificados por decréscimos superiores a 5% entre pontos sucessivos, fisicamente incompatíveis com a compressão monotônica do ensaio. Dos 1.231 pontos brutos, 1.207 são retidos, distribuídos pelos 24 ensaios. Atentou-se ainda, na extração das curvas a partir das planilhas de ensaio, à preservação integral de cada curva, incluindo o seu último ponto registrado, cuja perda foi corrigida em relação a uma versão preliminar do processamento.

O critério de remoção foi, portanto, físico, e não estatístico. Optou-se deliberadamente por não empregar um truncamento puramente estatístico, como a exclusão de valores situados fora do intervalo interquartil, como regra de descarte. A razão é que diversas respostas mecânicas legítimas do solo reforçado, como os picos de resistência pronunciados e as transições abruptas entre contração e dilatação, afastam-se estatisticamente do conjunto sem deixar de ser fisicamente válidas, de modo que removê-las por um critério estatístico descaracterizaria justamente os comportamentos extremos que o modelo precisa aprender a reproduzir. A remoção restringiu-se, assim, aos registros fisicamente impossíveis ou desprovidos de significado mecânico, preservando-se os valores que, embora extremos, são factíveis para o material.

3.3.2 A variável indicadora de reforço

Às variáveis descritivas de cada ponto foi acrescentada uma variável binária, denominada *tem_fibra*, que assume valor igual a 1 quando há fibra na amostra e igual a 0 para a areia sem reforço. À primeira vista, poderia parecer suficiente representar a areia

sem reforço apenas pelo teor de fibra nulo ($P_f = 0$), sem qualquer variável adicional. A introdução dessa variável justifica-se, contudo, pelo fato de a areia pura e o compósito solo-fibra constituírem regimes mecânicos categoricamente distintos, e não apenas pontos extremos de uma mesma variação quantitativa de teor. Sem essa indicação explícita, a condição de teor de fibra nulo se confundiria numericamente com o limite inferior do compósito, dificultando a distinção pela rede. A variável binária atua, assim, como um interruptor de regime: quando seu valor é zero, a rede aprende a desconsiderar as variáveis de teor e de comprimento de fibra, irrelevantes para a areia pura. Essa necessidade não se manifestava em uma versão preliminar do modelo, que não dispunha de ensaios experimentais de areia sem reforço para validação.

3.3.3 Normalização

A normalização, também chamada padronização, é uma transformação aplicada às variáveis antes do treinamento para colocá-las todas em uma mesma escala numérica, prática consagrada no pré-processamento de dados para redes neurais (GOODFELLOW; BENGIO; COURVILLE, 2016). Adotou-se a padronização por escore-padrão (*z-score*), na qual, de cada valor de uma variável, subtrai-se a média dessa variável e divide-se o resultado pelo seu desvio padrão:

$$z = \frac{x - \mu}{\sigma}, \quad (7)$$

em que x é o valor original, μ a média da variável e σ o seu desvio padrão. Cada variável transformada passa a ter média nula e desvio padrão unitário, e o valor z exprime a quantos desvios padrão o dado original se encontra da média. A padronização foi aplicada tanto às variáveis de entrada quanto às de saída.

Esse passo é necessário porque as grandezas envolvidas possuem ordens de magnitude e unidades muito distintas: as tensões são da ordem de centenas de quilopascal, ao passo que as deformações são adimensionais e da ordem de centésimos. Sem a padronização, as variáveis de maior magnitude dominariam tanto o cálculo da função de perda quanto a atualização dos pesos, prejudicando o aprendizado das demais. A recomendação de padronizar as variáveis de entrada para média nula e desvio padrão unitário, de modo a acelerar e a estabilizar a convergência do treinamento baseado em gradiente, é justificada em detalhe por LeCun et al. (2012).

Um cuidado metodológico essencial diz respeito ao ajuste dos parâmetros de normalização. Os parâmetros de média e desvio padrão foram estimados exclusivamente sobre os dados sintéticos, e os mesmos parâmetros foram aplicados, sem reajuste, aos dados experimentais. Esse procedimento garante que a normalização seja idêntica em todas as rodadas de validação e elimina qualquer possibilidade de vazamento de informação dos dados experimentais de validação para o processo de treinamento, o que poderia inflar artificialmente as métricas de desempenho.

3.3.4 Estratégia de validação

A avaliação de um modelo preditivo exige medir o seu desempenho sobre dados que ele não utilizou durante o treinamento, pois somente assim se afere a sua capacidade de generalização, isto é, de prever a resposta em condições não vistas, e não a mera memorização dos exemplos apresentados (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). A maneira de separar os dados entre treinamento e avaliação não é, portanto, um detalhe operacional, mas uma decisão que condiciona a confiabilidade de toda a aferição de desempenho. No presente trabalho, duas características da base experimental tornam essa decisão particularmente crítica: o número reduzido de ensaios, apenas 24, e a heterogeneidade entre eles, pois cobrem quatro tensões de confinamento, três teores e três comprimentos de fibra. Dessas características decorrem três requisitos para a partição adotada: preservar o máximo de informação para o treinamento; produzir uma estimativa de desempenho que não dependa do acaso de um sorteio particular; e garantir que a avaliação percorra todas as condições experimentais. As alternativas usuais de partição são examinadas a seguir à luz desses requisitos.

O procedimento mais simples, a partição única reservada (*holdout*), separa de uma só vez uma fração fixa dos dados para teste e treina o modelo no restante. Aplicada à base aqui disponível, essa estratégia não atende a nenhum dos três requisitos. Primeiro, com apenas 24 ensaios, reservar uma fração para teste subtrai uma parcela valiosa de uma base já reduzida, justamente a que mais carece de exemplos para o ajuste. Segundo, a estimativa de desempenho resultante depende fortemente de quais ensaios são sorteados para o conjunto de teste: uma única partição de um conjunto tão pequeno produz uma medida instável, que pode parecer otimista ou pessimista por simples acaso da divisão. Terceiro, e mais decisivo, a heterogeneidade dos ensaios faz com que uma partição única possa deixar regimes inteiros de comportamento ausentes do treinamento ou da avaliação, comprometendo tanto o aprendizado quanto a representatividade da medida.

Uma alternativa intermediária é a validação cruzada em k partições (k -fold), na qual os dados são divididos em k blocos e cada bloco serve uma vez como teste, enquanto os demais compõem o treinamento. A divisão clássica em torno de 80% para treinamento e 20% para teste corresponde, nesse esquema, a aproximadamente cinco partições. Esse arranjo melhora a cobertura da avaliação, pois todos os ensaios são previstos ao longo dos blocos, mas atende apenas parcialmente aos demais requisitos: cada repetição ainda exclui do treinamento um bloco de cerca de cinco ensaios, o que mantém o risco de deixar regimes inteiros de comportamento fora do ajuste, e a composição dos blocos permanece dependente de sorteio, com estimativas relativamente ruidosas dado o tamanho reduzido da base.

A validação cruzada exaustiva do tipo *leave-one-out* (LOOCV), expressão que significa deixar um de fora, é o caso-limite da validação em k partições em que k iguala o número

de ensaios, sendo particularmente apropriada quando esse número é pequeno (KOHAVI, 1995; HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Sua adoção responde diretamente aos três requisitos: cada repetição treina com todos os ensaios menos um, preservando o máximo de informação para o ajuste; a estimativa não depende de sorteio algum, pois o esquema enumera todas as partições possíveis em que um único ensaio é reservado; e a avaliação percorre todas as condições experimentais, pois cada um dos 24 ensaios é previsto exatamente uma vez, sempre na condição de dado não visto. Obtém-se ainda, como subproduto, uma estimativa de desempenho para cada ensaio individualmente. A contrapartida é o custo computacional, pois o ajuste é repetido tantas vezes quantos forem os ensaios, o que, no presente caso, permanece modesto.

Definido o protocolo, cabe especificar como ele se articula com a estratégia de treinamento deste trabalho, detalhada na Seção 3.5. O treinamento divide-se em duas fases sequenciais: na primeira, o pré-treinamento, a rede aprende o comportamento constitutivo geral a partir do grande volume de dados sintéticos; na segunda, o ajuste fino, esse conhecimento é refinado com os dados experimentais, mais escassos, porém mais precisos. Ambas as fases são etapas de treinamento, nas quais os pesos da rede são efetivamente ajustados, mas apenas a segunda participa da repetição imposta pela validação cruzada: o pré-treinamento é executado uma única vez, antes do início da validação, e o estado pré-treinado resultante serve de ponto de partida comum a todas as repetições. Cada repetição, aqui denominada rodada, parte desse estado: a rede é reajustada com 23 dos 24 ensaios experimentais, e o seu erro de previsão é medido sobre o ensaio restante, reservado exclusivamente para teste e sem qualquer participação no ajuste dos pesos. Os ensaios experimentais desempenham, assim, papel duplo, mas nunca simultâneo: em cada rodada, 23 deles servem ao treinamento e um serve exclusivamente à avaliação.

Em síntese, cada rodada emprega os dados da seguinte forma:

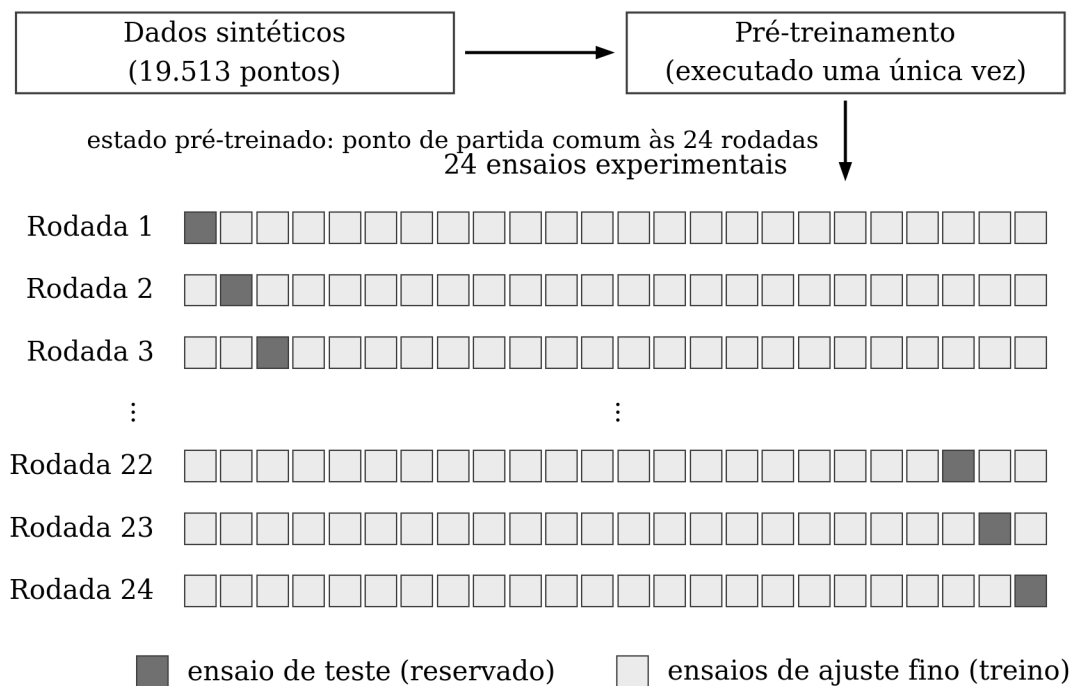
- ❑ os dados sintéticos, 19.513 pontos, alimentam o pré-treinamento, executado uma única vez e reaproveitado por todas as rodadas, e constituem, portanto, dados de treino;
- ❑ os 23 ensaios experimentais que compõem o ajuste fino, de 1.147 a 1.166 pontos conforme a rodada, são igualmente dados de treino;
- ❑ o ensaio experimental restante, de 41 a 60 pontos, é apenas previsto, e constitui o único dado de teste.

O procedimento completo compreende, portanto, 24 rodadas, uma para cada ensaio reservado. Cabe esclarecer a natureza dessas rodadas, uma vez que a denominação validação cruzada pode induzir à interpretação equivocada de um treinamento único que se aprimora sucessivamente. As 24 rodadas são treinamentos independentes entre si: em

cada uma, a rede retorna ao mesmo estado pré-treinado e o ajuste fino é reiniciado, empregando os 23 ensaios daquela rodada, sem que os pesos ajustados em qualquer rodada anterior sejam herdados. Em consequência, o desempenho obtido independe da ordem de execução das rodadas, e os únicos elementos comuns a todas elas são o protocolo de treinamento e o estado pré-treinado. Cada rodada constitui, assim, um experimento completo e autônomo, destinado exclusivamente a produzir a previsão do ensaio nela reservado.

A Figura 4 sintetiza o esquema completo: a montante, o pré-treinamento sintético, executado uma única vez; ao centro, as 24 rodadas independentes de ajuste fino, que partem de cópias idênticas do estado pré-treinado; e, a jusante, a reunião das previsões dos 24 ensaios reservados, da qual se extraem as métricas de desempenho.

Figura 4 – Esquema da validação cruzada *leave-one-out*: o pré-treinamento, executado uma única vez, fornece o ponto de partida comum às 24 rodadas de ajuste fino; em cada rodada, um ensaio é reservado para teste, e as previsões dos 24 ensaios reservados são reunidas ao final para o cálculo das métricas agregadas



Fonte: Elaborado pelo autor.

Nenhum gradiente é calculado sobre o ensaio reservado de cada rodada, de modo que os seus pontos jamais contribuem para o ajuste dos pesos da rede que o avalia. Ao final das 24 rodadas, as previsões dos ensaios reservados são reunidas em um único conjunto, que reconstrói a previsão do modelo sobre a totalidade dos 1.207 pontos experimentais. Convém frisar que é sobre esse conjunto concatenado, e não sobre a média dos 24 resultados individuais, que se calcula a métrica principal de desempenho: a média dos coeficientes de determinação por ensaio seria dominada pelos poucos ensaios de baixo confinamento em que a métrica se torna instável, pelas razões discutidas na Seção 3.6. A distribuição

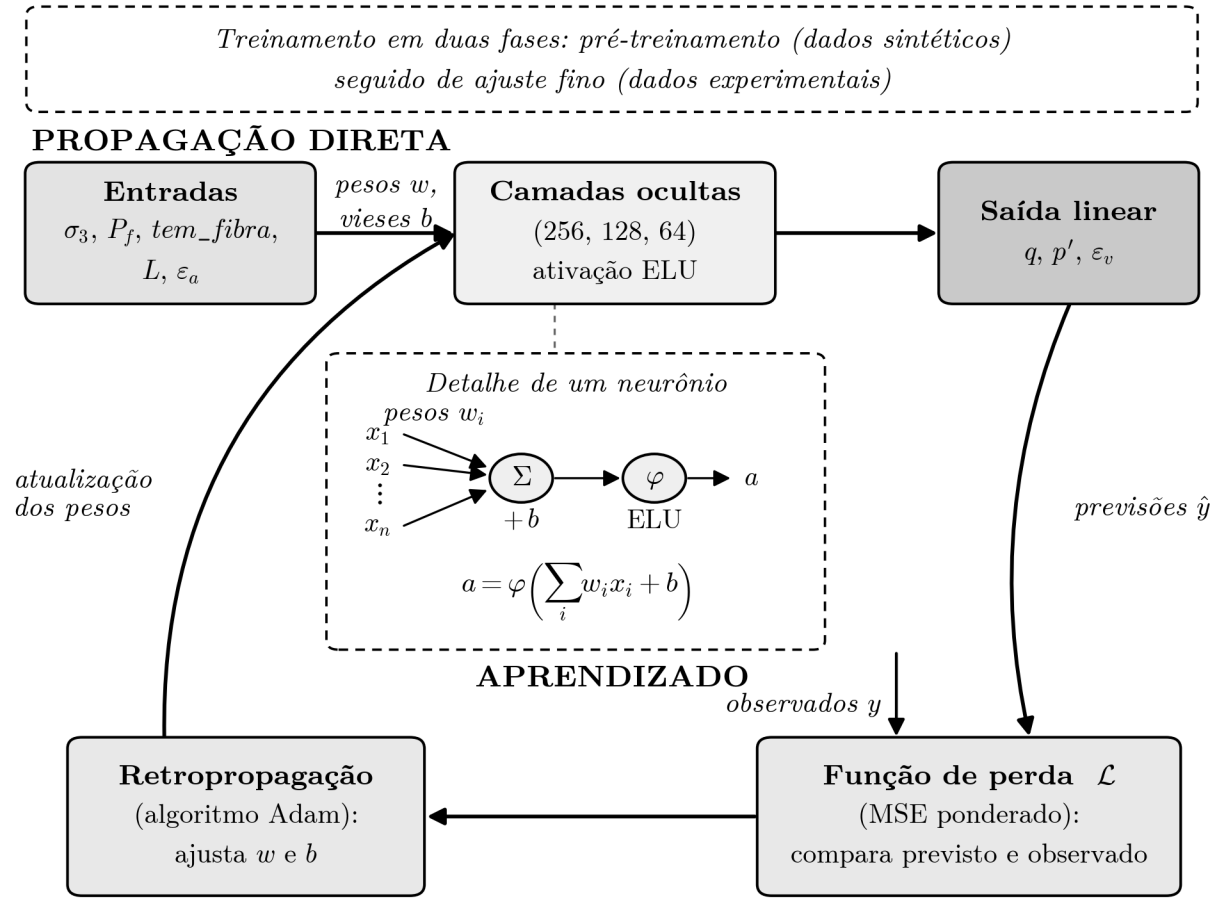
do desempenho entre os ensaios é reportada à parte, por medidas robustas de posição e de dispersão.

Cabe registrar, por transparência metodológica, o papel desempenhado pelos dados experimentais nos critérios de parada do treinamento. No pré-treinamento, a perda de validação que rege a interrupção antecipada é medida sobre o conjunto experimental completo (Seção 3.5); como o estado pré-treinado resultante é compartilhado identicamente por todas as rodadas, essa influência é simétrica entre elas. No ajuste fino de cada rodada, a interrupção antecipada e a seleção do melhor estado monitoram a perda sobre o próprio ensaio reservado. Em ambos os casos, os pesos não são ajustados sobre esses dados, mas a escolha do ponto de parada é por eles informada, o que constitui uma fonte conhecida de leve otimismo nas métricas reportadas (CAWLEY; TALBOT, 2010). A alternativa de reservar um segundo ensaio como validação interna reduziria o já escasso conjunto de ajuste fino e tornaria a parada dependente de um único ensaio, de condição experimental possivelmente muito distinta da do ensaio avaliado. Diante disso, optou-se pelo procedimento descrito, e a ressalva correspondente é retomada nas limitações do trabalho.

3.4 Arquitetura da Rede Neural Artificial

Esta seção apresenta a arquitetura da rede neural adotada, detalhando o tipo de rede, a sua estrutura em camadas, as variáveis de entrada e de saída, a função de ativação e a função de perda. Convém, de início, uma visão de conjunto do modelo e de seu funcionamento, sintetizada na Figura 5. A rede recebe, na camada de entrada, as variáveis que descrevem a configuração do ensaio e o nível de deformação; propaga essa informação pelas camadas ocultas, nas quais cada neurônio calcula uma soma de suas entradas ponderada por pesos, acrescida de um viés, e submete o resultado a uma função de ativação não-linear; e devolve, na camada de saída, as três variáveis de estado de interesse. No treinamento, uma função de perda mede a discrepância entre as previsões e os valores observados, e o algoritmo de retropropagação ajusta os pesos e os vieses no sentido que reduz essa discrepância, em um ciclo repetido por muitas iterações. Esse ajuste ocorre em duas fases, o pré-treinamento com dados sintéticos e o ajuste fino com dados experimentais. Cumpre distinguir o ajuste fino da validação: o ajuste fino é uma etapa de treinamento, na qual os pesos da rede são atualizados com os ensaios experimentais de ajuste, ao passo que a validação é realizada exclusivamente sobre o ensaio reservado de cada rodada, que não participa de nenhuma das duas fases de treinamento (Seção 3.3.4). As subseções seguintes detalham cada elemento dessa visão de conjunto, e a Seção 3.5 trata da estratégia de treinamento.

Figura 5 – Visão de conjunto da rede neural: da entrada às variáveis de estado de saída, com o detalhe de um neurônio e o ciclo de treinamento, no qual as previsões da rede ($\hat{y}$) são comparadas aos valores observados (y) pela função de perda



Fonte: Elaborado pelo autor.

3.4.1 Tipo e estrutura da rede

A arquitetura adotada é uma rede neural de propagação direta, do tipo perceptron multicamadas (*Multilayer Perceptron*, MLP) (HAYKIN, 2009). Essa escolha fundamenta-se no teorema da aproximação universal (HORNIK; STINCHCOMBE; WHITE, 1989), que assegura capacidade teórica para representar relações constitutivas complexas, e no histórico consolidado de sucesso de arquiteturas de propagação direta na modelagem constitutiva de solos (GHABOUSSI; JR; WU, 1991; HASHASH; JUNG; GHABOUSSI, 2004; ZHANG et al., 2021).

Em termos acessíveis, uma rede dessa natureza é uma função matemática flexível, organizada em camadas de unidades de processamento chamadas neurônios. Cada neurônio recebe valores numéricos, calcula uma soma ponderada dessas entradas, à qual acrescenta uma constante, e aplica ao resultado uma transformação não-linear, repassando-o adiante.

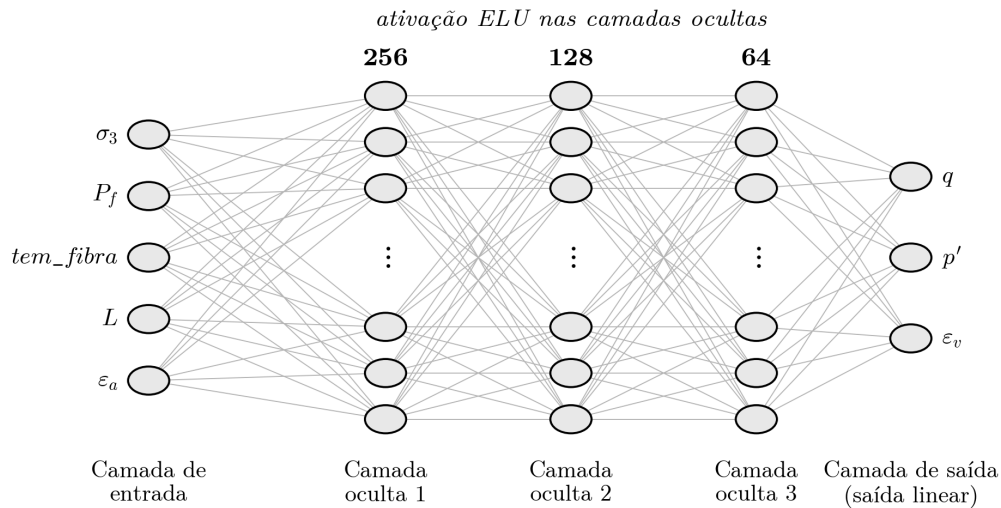
Formalmente, a saída a_j de um neurônio j é

$$a_j = \varphi \left(\sum_i w_{ji} x_i + b_j \right), \quad (8)$$

em que o subíndice i indexa as entradas do neurônio e o subíndice j identifica o neurônio considerado; x_i é a i -ésima entrada, w_{ji} o peso que a pondera, b_j o viés, a constante aditiva, e φ a função de ativação (GOODFELLOW; BENGIO; COURVILLE, 2016). Os pesos e os vieses constituem os parâmetros ajustáveis da rede, e é no seu ajuste, durante o treinamento, que reside o próprio aprendizado, conforme apresentado no Capítulo 2. A expressão propagação direta (*feedforward*) indica que a informação flui em um único sentido, da camada de entrada para a de saída, sem realimentação.

A estrutura da rede é ilustrada na Figura 6. Ela é composta por uma camada de entrada com cinco neurônios, um para cada variável de entrada; três camadas ocultas com, respectivamente, 256, 128 e 64 neurônios; e uma camada de saída com três neurônios, um para cada variável prevista. As camadas são totalmente conectadas, isto é, cada neurônio de uma camada recebe as saídas de todos os neurônios da camada anterior, multiplicadas por pesos ajustáveis, aos quais se soma um viés. A contagem total dos pesos e vieses decorre das conexões entre camadas sucessivas: são 1.536 parâmetros entre a camada de entrada e a primeira oculta (os 5×256 pesos mais 256 vieses), 32.896 entre a primeira e a segunda camada oculta, 8.256 entre a segunda e a terceira, e 195 entre a terceira oculta e a camada de saída, perfazendo 42.883 parâmetros treináveis.

Figura 6 – Arquitetura da rede neural de propagação direta adotada



Entradas:

σ_3 (tensão de confinamento) L (comprimento da fibra)
 P_f (teor de fibra) ε_a (deformação axial)
 tem_fibra (indicador binário)

Saídas:

q (tensão desviadora)
 p' (tensão média efetiva)
 ε_v (deformação volumétrica)

Fonte: Elaborado pelo autor.

A redução progressiva do número de neurônios entre as camadas ocultas, de 256 para 128 e depois para 64, configura uma estrutura em funil. A primeira camada oculta, mais larga, expande as cinco variáveis de entrada em uma representação interna rica, capaz de distinguir um grande número de padrões do comportamento mecânico; as camadas seguintes, mais estreitas, sintetizam progressivamente essa representação até as três variáveis de saída. Essa estrutura em funil resultou de uma escolha de projeto, detalhada a seguir, e não de uma regra geral de topologia.

A definição dessa topologia, isto é, do número de camadas e de neurônios por camada, não resultou de uma busca automatizada de hiperparâmetros, mas de um refinamento empírico. Existem técnicas sistemáticas para essa busca, como a busca em grade (*grid search*), a busca aleatória (*random search*) e a otimização bayesiana, esta última implementada em ferramentas como o Optuna (AKIBA et al., 2019), que exploram automaticamente um espaço de configurações candidatas em busca da de melhor desempenho. Optou-se, porém, por um procedimento manual e incremental, adequado ao porte do problema: partiu-se de uma rede mais simples, associada a uma versão inicial do trabalho, ampliada de forma gradual ao longo do desenvolvimento. O critério para interromper essa ampliação foi prático: aumentou-se a capacidade da rede até que ela reproduzisse adequadamente a resposta do compósito, sem atingir um porte que comprometesse a estabilidade do treinamento ou favorecesse o sobreajuste. O desempenho obtido na validação cruzada, apresentado no Capítulo 4, confirma a adequação da topologia escolhida.

3.4.2 Variáveis de entrada e saída

A formulação adotada para o modelo é a de um mapeamento direto, no qual a rede recebe a configuração do ensaio e o nível de deformação imposto e devolve as variáveis de estado correspondentes a esse nível de deformação, sem qualquer integração ao longo da trajetória. O vetor de entrada é composto por cinco variáveis:

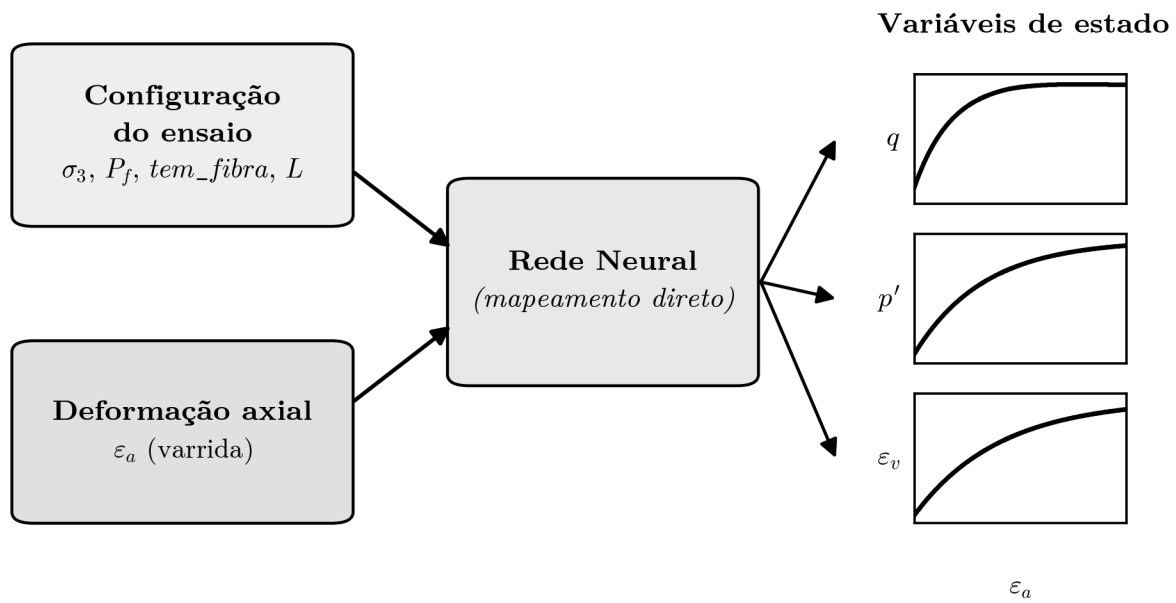
- tensão de confinamento, σ_3 ;
- teor de fibra, P_f ;
- variável indicadora de reforço, tem_fibra ;
- comprimento de fibra, L ;
- deformação axial, ε_a .

As três primeiras variáveis de entrada, a tensão de confinamento, o teor e o comprimento de fibra, definem a configuração do ensaio e permanecem constantes ao longo de uma mesma curva; a variável indicadora de reforço distingue a areia pura do compósito,

conforme discutido na Seção 3.3.2; e a deformação axial é a variável independente, percorrida ponto a ponto para descrever o ensaio. O vetor de saída é composto por três variáveis: a tensão desviadora (q), a tensão média efetiva (p') e a deformação volumétrica (ε_v).

A Figura 7 sintetiza esse mapeamento. Fixada uma configuração de ensaio e percorrida a deformação axial, a rede reconstrói ponto a ponto as três curvas que caracterizam a resposta do material: a curva tensão-deformação, isto é, q em função de ε_a ; a trajetória de tensões, por meio de p' ; e a curva de variação volumétrica, ε_v em função de ε_a . O modelo, portanto, não prevê um único valor, mas a resposta completa do ensaio.

Figura 7 – Mapeamento direto das variáveis de entrada nas variáveis de estado de saída



Fonte: Elaborado pelo autor.

Cabe registrar que essa formulação difere daquela de uma versão preliminar deste trabalho, de natureza incremental, na qual a rede recebia o estado corrente de tensão e deformação acrescido de incrementos de deformação e devolvia os incrementos de tensão correspondentes. A formulação de mapeamento direto foi preferida por sua maior estabilidade e por dispensar a integração sucessiva de incrementos, que tende a acumular erros ao longo da trajetória.

3.4.3 Função de ativação

Em cada neurônio das camadas ocultas, após a combinação linear ponderada das entradas, o resultado é transformado por uma função de ativação não-linear. É essa não-linearidade que confere à rede a capacidade de representar relações complexas: sem ela, a composição de várias camadas equivaleria a uma única transformação linear, incapaz de descrever o comportamento tensão-deformação dos solos.

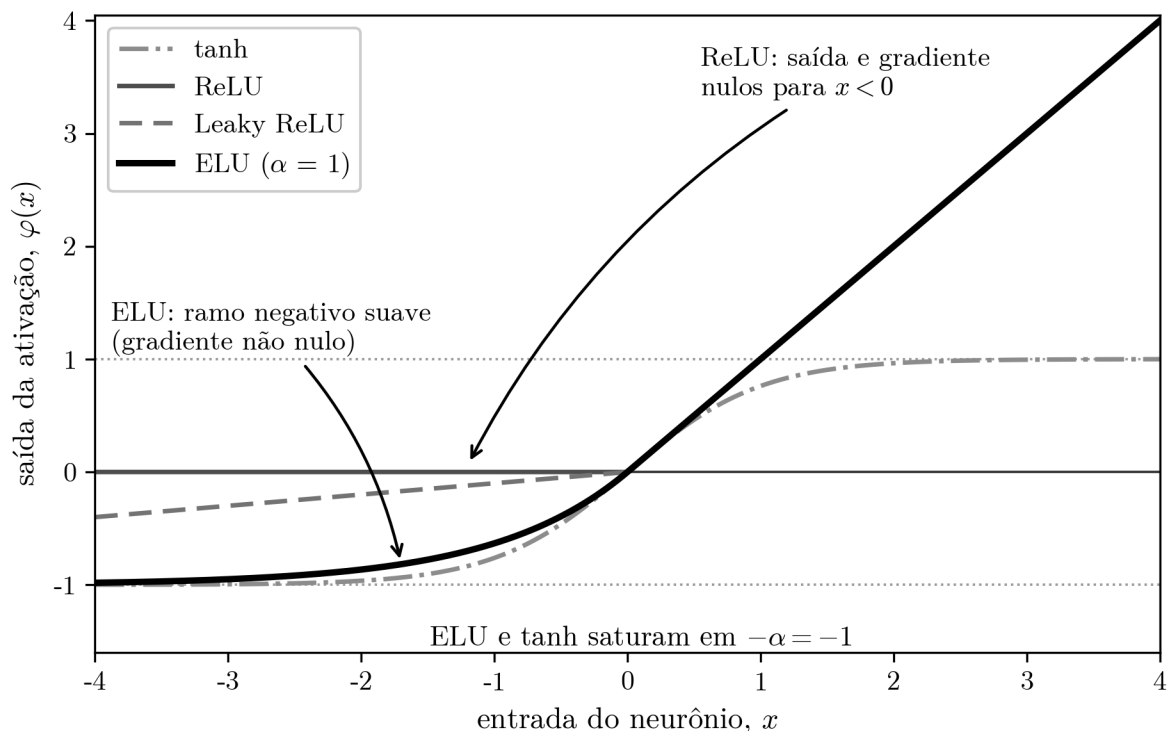
Diversas funções de ativação são empregadas na prática, cada uma com vantagens e limitações. As funções sigmoide (logística) e tangente hiperbólica ($\tanh$), de formato em S, predominaram nas primeiras redes, mas saturam para entradas de grande módulo, naquela região em que sua derivada se aproxima de zero; em redes profundas, essa saturação conduz ao desaparecimento do gradiente, fenômeno que estagna o aprendizado das camadas iniciais (HAYKIN, 2009; GOODFELLOW; BENGIO; COURVILLE, 2016). A função *Rectified Linear Unit* (ReLU), definida como o máximo entre zero e a entrada, tornou-se o padrão em aprendizado profundo por ser computacionalmente simples e por não saturar para entradas positivas, o que mitiga o desaparecimento do gradiente (NAIR; HINTON, 2010; GLOROT; BORDES; BENGIO, 2011). A ReLU, contudo, anula a saída e o gradiente em toda a região negativa, o que pode provocar o problema dos neurônios inativos, o chamado neurônio morto, no qual uma unidade fica permanentemente em zero e deixa de aprender. Variantes como a *Leaky ReLU* (ReLU com vazamento) atribuem uma pequena inclinação à região negativa, justamente para contornar esse problema (MAAS; HANNUN; NG, 2013).

Nas camadas ocultas deste trabalho foi empregada a função *Exponential Linear Unit* (ELU), definida como

$$\text{ELU}(x) = \begin{cases} x & \text{se } x > 0 \\ \alpha(e^x - 1) & \text{se } x \leq 0 \end{cases} \quad (9)$$

em que o parâmetro α , adotado como unitário, controla a saturação para entradas negativas. O comportamento da função, comparado ao das alternativas discutidas, é ilustrado na Figura 8. Para entradas positivas, a ELU coincide com a própria entrada, transmitindo o sinal sem distorção e com gradiente constante, à semelhança da ReLU; para entradas negativas, em vez de anulá-las, ela decresce de forma suave até saturar em $-\alpha$, mantendo uma saída e um gradiente não nulos.

Figura 8 – Função de ativação ELU adotada nas camadas ocultas, comparada às funções tanh, ReLU e Leaky ReLU



Fonte: Elaborado pelo autor.

A escolha da ELU justifica-se por reunir, entre as funções usuais, as propriedades mais convenientes para o problema. A primeira é que o ramo negativo suave mantém a média das ativações próxima de zero, o que acelera e estabiliza a convergência do treinamento. A segunda é que, por preservar gradientes não nulos também para entradas negativas, a ELU evita o problema do neurônio morto que afeta a ReLU (CLEVERT; UNTERTHINER; HOCHREITER, 2016). A terceira é a suavidade da função, contínua e diferenciável em toda a sua extensão, característica particularmente adequada à regressão de curvas constitutivas suaves, ao contrário da ReLU e da Leaky ReLU, que apresentam um vértice anguloso em zero. Por reunir essas três características, a ELU foi preferida às demais funções mostradas na Figura 8.

Na camada de saída foi empregada a função de ativação linear (identidade), adequada a um problema de regressão em que as variáveis de saída não possuem limites de intervalo pré-definidos.

3.4.4 Função de perda

A função de perda é a medida que o treinamento procura minimizar: uma função que quantifica, em um único número, a discrepância entre as previsões da rede e os valores observados (GOODFELLOW; BENGIO; COURVILLE, 2016; BISHOP, 2006). É ela que define o objetivo do aprendizado, pois, a cada iteração, o seu valor é calculado sobre

um conjunto de exemplos e o seu gradiente em relação aos pesos indica a direção em que estes devem ser ajustados para reduzir o erro, conforme detalhado na Seção 3.5. Para problemas de regressão, como o aqui tratado, a escolha usual é o erro quadrático médio (*Mean Squared Error*, MSE), que penaliza os desvios de forma proporcional ao seu quadrado e corresponde à estimativa de máxima verossimilhança sob a hipótese de erros gaussianos (GOODFELLOW; BENGIO; COURVILLE, 2016). Adotou-se o erro quadrático médio em uma versão ponderada, na qual o erro de cada variável de saída é multiplicado por um peso antes do cálculo da média. Para um conjunto de n exemplos e três variáveis de saída, a função de perda é

$$\mathcal{L} = \frac{1}{3} \sum_{k=1}^3 w_k \left[\frac{1}{n} \sum_{i=1}^n (\hat{y}_{ik} - y_{ik})^2 \right], \quad (10)$$

em que y_{ik} e $\hat{y}_{ik}$ são, respectivamente, o valor observado e o previsto da k -ésima variável de saída no i -ésimo exemplo, e w_k é o peso atribuído a essa variável. O termo entre colchetes é o erro quadrático médio de uma variável, e a ponderação faz a rede priorizar as variáveis de maior peso. Foram adotados os pesos $w_q = 2,0$ para a tensão desviadora, $w_{p'} = 1,0$ para a tensão média efetiva e $w_{\varepsilon_v} = 3,0$ para a deformação volumétrica.

Como as variáveis são padronizadas antes do treinamento (Seção 3.3), todas compartilham uma mesma escala na função de perda, de modo que os pesos não corrigem diferenças de magnitude, mas exprimem a importância e a dificuldade relativas de cada variável. A deformação volumétrica, que é a grandeza de previsão mais difícil e a responsável por capturar a dilatação, recebe o maior peso; a tensão desviadora, central para a resistência, recebe peso intermediário; e a tensão média efetiva, que em ensaio drenado é parcialmente determinada pela tensão desviadora pela relação $p' = \sigma_3 + q/3$ e é, portanto, de previsão mais simples, recebe o menor peso. A ponderação evita que as variáveis mais fáceis dominem a otimização em detrimento da deformação volumétrica. Cabe registrar que o critério de interrupção antecipada do treinamento monitorou o erro quadrático médio não ponderado, para que a seleção do melhor modelo não fosse enviesada pelos pesos.

3.5 Estratégia de Treinamento

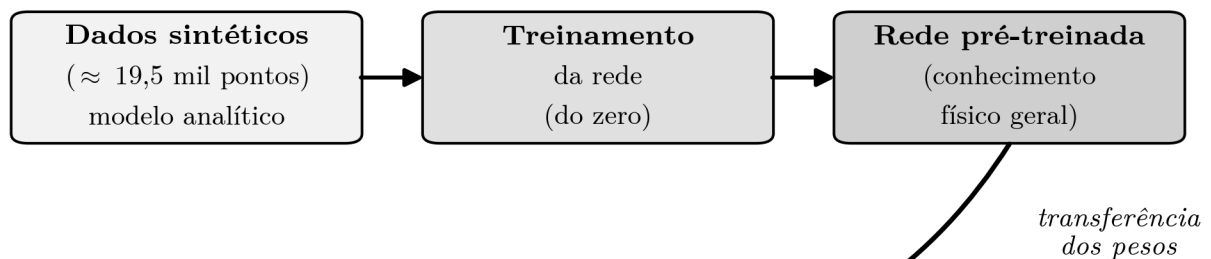
A estratégia de treinamento materializa o aprendizado por transferência discutido no Capítulo 2 e antecipado na Seção 3.3.4. Em vez de treinar a rede diretamente com os escassos dados experimentais, o que seria insuficiente para ajustar uma rede com dezenas de milhares de parâmetros, o treinamento é dividido nas duas fases sequenciais já apresentadas: o pré-treinamento, no qual a rede aprende o comportamento constitutivo geral a partir do grande volume de dados sintéticos, e o ajuste fino, no qual esse conhecimento é refinado com os dados experimentais, mais escassos, porém mais precisos.

3.5.1 Pré-treinamento

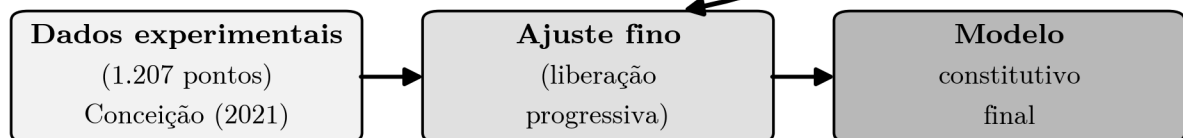
Na primeira fase, a rede é exposta exclusivamente aos dados sintéticos, com o objetivo de capturar os padrões fundamentais do comportamento elastoplástico do compósito, entre eles a não-linearidade da curva tensão-deformação e a transição entre os regimes contrátil e dilatante. Como esses dados são abundantes e fisicamente consistentes, eles permitem que a rede, partindo de pesos iniciais aleatórios, construa uma representação geral do comportamento do material antes de qualquer contato com os escassos dados experimentais. A Figura 9 resume essa estratégia em duas fases.

Figura 9 – Estratégia híbrida de treinamento em duas fases: pré-treinamento com dados sintéticos e ajuste fino com dados experimentais

Fase 1: Pré-treinamento



Fase 2: Ajuste fino



Fonte: Elaborado pelo autor.

O treinamento ajusta os pesos da rede de forma iterativa, pelo método do gradiente descendente: a cada passo, calcula-se o gradiente da função de perda em relação aos pesos, que aponta a direção de crescimento mais rápido do erro, e os pesos são deslocados no sentido oposto, reduzindo a perda (GOODFELLOW; BENGIO; COURVILLE, 2016). A cada passo, apenas um subconjunto dos dados, denominado lote (*batch*), é apresentado à rede; uma passagem completa por todos os lotes constitui uma época (*epoch*), e o treinamento percorre muitas épocas sucessivas.

O ajuste dos pesos é conduzido pelo algoritmo Adam (*Adaptive Moment Estimation*) (KINGMA; BA, 2015), uma variante do gradiente descendente estocástico que adapta automaticamente o tamanho do passo de cada peso. Para isso, o Adam mantém, ao longo do treinamento, duas médias móveis dos gradientes de cada peso: a do próprio gradiente, que atua como uma inércia, ou momento, suavizando a trajetória de ajuste

e atenuando oscilações; e a do quadrado do gradiente, que mede a magnitude típica recente das atualizações daquele peso e serve para escalonar o passo, de modo que pesos com gradientes grandes ou ruidosos recebam passos menores, e pesos com gradientes pequenos e consistentes, passos maiores. Essa adaptação individual do passo torna o método robusto e de convergência rápida, com pouca necessidade de ajuste manual, o que justifica a sua adoção, hoje padrão no treinamento de redes neurais (GOODFELLOW; BENGIO; COURVILLE, 2016).

Dois parâmetros regem essa otimização. A taxa de aprendizado (*learning rate*) controla a magnitude dos passos de ajuste: passos grandes aceleram o início do treinamento, mas podem impedir a convergência fina, ao passo que passos pequenos são mais estáveis, porém lentos. Adotou-se a taxa inicial de 10^{-3} , valor de referência recomendado para o Adam e amplamente empregado por equilibrar velocidade e estabilidade (KINGMA; BA, 2015). O tamanho do lote foi fixado em 256 exemplos: lotes maiores produzem estimativas de gradiente menos ruidosas, à custa de atualizações menos frequentes, e um lote dessa ordem é adequado ao grande volume de dados sintéticos do pré-treinamento, em contraste com o lote menor empregado no ajuste fino, em que os dados experimentais são escassos.

O número de épocas não é fixado de antemão, mas determinado por dois mecanismos que regulam a duração e a estabilidade do treinamento. O primeiro é um escalonador da taxa de aprendizado, que a reduz à metade sempre que a perda de validação deixa de diminuir por vinte épocas consecutivas; essa redução permite passos progressivamente mais finos à medida que o modelo se aproxima de um ajuste adequado, refinando-o sem oscilações. O segundo é a interrupção antecipada (*early stopping*): a cada época, mede-se a perda sobre o conjunto experimental e, quando não há melhora por quarenta épocas consecutivas, número denominado paciência, o treinamento é encerrado e o estado de melhor desempenho já observado é restaurado (GOODFELLOW; BENGIO; COURVILLE, 2016). Impôs-se ainda um teto de quinhentas épocas, como limite de segurança.

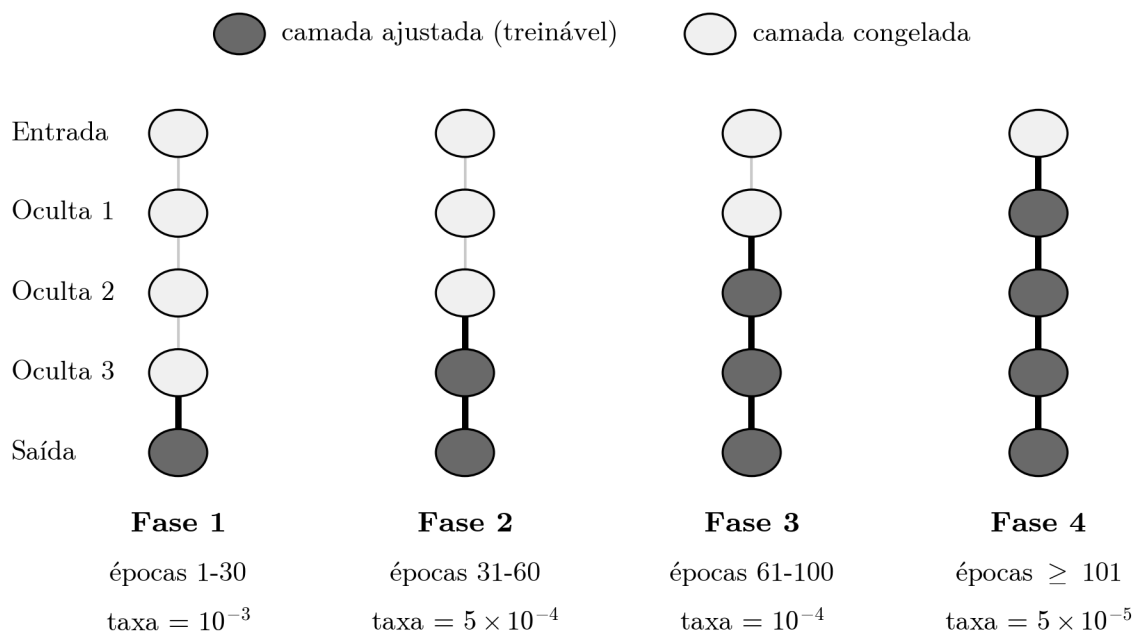
A escolha desses valores tem motivação prática. O teto de quinhentas épocas é deliberadamente generoso, para não interromper o aprendizado prematuramente; quem efetivamente determina o ponto de parada é a interrupção antecipada, que encerra o treinamento no ponto de menor perda de validação, evitando tanto o subtreinamento, caso a parada fosse precoce demais, quanto o sobreajuste e o esforço computacional inútil, caso o treinamento se prolongasse sem ganho. A paciência de quarenta épocas é uma tolerância intermediária: longa o suficiente para não encerrar o treinamento diante de uma estabilização meramente transitória da perda, e curta o suficiente para interrompê-lo pouco depois da convergência. Na prática, o encerramento ocorre bem antes do teto: a perda sobre os dados sintéticos decresce rapidamente nas primeiras épocas, e a perda de validação experimental, que rege a parada, estabiliza-se cedo, pois o refinamento adicional do ajuste à teoria pouco altera a distância da rede em relação ao comportamento experimental. O pré-treinamento é executado uma única vez, e o estado resultante serve

de ponto de partida para todas as rodadas da validação cruzada, conforme detalhado na Seção 3.3.4.

3.5.2 Ajuste fino

A segunda fase, o ajuste fino (*fine-tuning*), refina o modelo pré-treinado com os dados experimentais. É a etapa em que se concretiza o aprendizado por transferência (PAN; YANG, 2010): em vez de treinar uma nova rede do zero, parte-se dos pesos já ajustados no pré-treinamento e os adapta ao comportamento experimental. A maneira como essa adaptação é conduzida, contudo, é decisiva. Reajustar toda a rede de uma só vez, com os poucos dados experimentais, poderia apagar bruscamente o conhecimento geral adquirido na primeira fase. Adota-se, por isso, uma liberação gradual das camadas, técnica conhecida como descongelamento progressivo (*gradual unfreezing*) (HOWARD; RUDER, 2018), ilustrada na Figura 10.

Figura 10 – Liberação progressiva das camadas no ajuste fino, da camada de saída em direção à de entrada



Fonte: Elaborado pelo autor.

Descongelar uma camada significa torná-la novamente treinável, isto é, permitir que os seus pesos sejam atualizados; manter uma camada congelada significa preservar os seus pesos inalterados durante a atualização. No descongelamento progressivo, as camadas são liberadas uma a uma, da camada de saída em direção à de entrada. Essa ordem tem fundamento: em redes profundas, as camadas iniciais tendem a codificar padrões gerais, comuns a muitos problemas, enquanto as camadas finais codificam padrões mais específicos da tarefa (YOSINSKI et al., 2014). Assim, as camadas próximas à saída, responsáveis pelos ajustes mais específicos, são adaptadas primeiro, ao passo que as camadas iniciais,

que guardam a representação geral herdada do pré-treinamento, só são ajustadas ao final e com passos muito pequenos.

A liberação ocorre em quatro fases. Nas trinta primeiras épocas, ajusta-se apenas a camada de saída, com taxa de aprendizado de 10^{-3} . Da época 31 à 60, libera-se também a última camada oculta, com taxa de 5×10^{-4} ; da época 61 à 100, as duas últimas camadas ocultas, com taxa de 10^{-4} ; e, a partir da época 101, toda a rede, com taxa de 5×10^{-5} . A redução sucessiva da taxa de aprendizado, à medida que mais camadas são liberadas, garante que as camadas iniciais sejam alteradas apenas de forma sutil. Os lotes contêm 32 exemplos, compatíveis com o número reduzido de dados experimentais.

Essa progressão visa preservar o conhecimento geral adquirido no pré-treinamento e, ao mesmo tempo, permitir a adaptação aos padrões específicos do material, evitando o chamado esquecimento catastrófico (*catastrophic forgetting*), fenômeno no qual a exposição a novos dados leva uma rede a descartar abruptamente o conhecimento previamente adquirido (FRENCH, 1999). O descongelamento progressivo, combinado à redução sucessiva da taxa de aprendizado, é uma das estratégias consagradas para mitigar esse efeito no ajuste fino (HOWARD; RUDER, 2018). Adotou-se novamente a interrupção antecipada com preservação do melhor estado, monitorando-se nessa fase a perda sobre o ensaio reservado da rodada (Seção 3.3.4), com tolerância maior do que a do pré-treinamento, de cem épocas sem melhora, e teto de quatrocentas épocas.

A escolha por essa tolerância mais elevada resultou de uma comparação conduzida ao longo do desenvolvimento. Investigou-se o reforço das medidas contra o sobreajuste (*overfitting*), por meio do aumento da penalização aplicada aos pesos e da redução da tolerância da interrupção antecipada. Verificou-se que essa alteração piorou os coeficientes de determinação, pois os ensaios de ajuste mais difícil exigiam um número maior de épocas para convergir, e o sobreajuste aparente revelou-se inofensivo, uma vez que a preservação do estado de melhor desempenho já capturava o ponto ótimo do treinamento. Manteve-se, por isso, a tolerância mais elevada na versão final do modelo.

3.6 Métricas de Avaliação

O desempenho do modelo foi quantificado por meio de duas métricas complementares, amplamente empregadas na avaliação de modelos de regressão (HASTIE; TIBSHIRANI; FRIEDMAN, 2009): o coeficiente de determinação (R^2) e a raiz do erro quadrático médio (RMSE). O coeficiente de determinação,

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (11)$$

em que y_i são os valores observados, $\hat{y}_i$ os previstos, $\bar{y}$ a média dos observados e n o número de observações. O numerador é a soma dos quadrados dos resíduos, isto é, a parcela da variação dos dados que o modelo não consegue explicar, e o denominador é a variação total

dos dados em torno de sua média; o coeficiente de determinação mede, assim, a proporção da variância dos dados explicada pelo modelo (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). A raiz do erro quadrático médio,

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (12)$$

mede a magnitude média dos erros na mesma unidade da variável prevista.

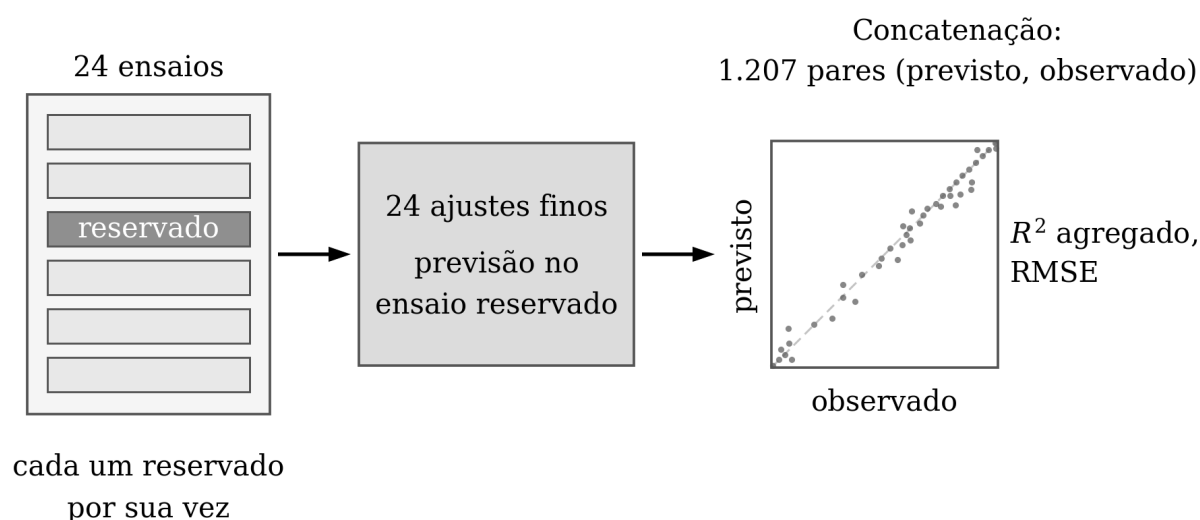
Em termos práticos, o coeficiente de determinação exprime a fração da variabilidade dos dados que o modelo consegue reproduzir: um valor igual a 1 indica previsão perfeita; um valor nulo indica desempenho equivalente ao de simplesmente prever a média dos dados; e valores negativos indicam desempenho pior do que essa referência. A raiz do erro quadrático médio, por sua vez, expressa o erro típico de previsão na unidade da própria grandeza, em quilopascal no caso das tensões, o que facilita a interpretação física da magnitude dos desvios. As duas métricas são, assim, complementares: o coeficiente de determinação é adimensional e relativo, o que permite comparar o desempenho entre variáveis de escalas distintas, ao passo que a raiz do erro quadrático médio é absoluta e dimensional, mais informativa sobre a magnitude prática do erro. Por isso ambas são reportadas em conjunto.

A escolha da raiz do erro quadrático médio, em vez do erro absoluto médio (*Mean Absolute Error*, MAE), outra métrica dimensional de uso corrente, merece justificativa. Por elevar os desvios ao quadrado antes de tomar a média, a RMSE atribui peso maior aos erros de grande magnitude, sendo mais sensível a desvios pontuais expressivos do que o MAE (CHAI; DRAXLER, 2014). Essa propriedade é conveniente no presente problema, em que a reprodução fiel de regiões críticas da curva, como o pico de resistência, importa mais do que o erro médio ao longo de todo o ensaio. Cabe registrar, em contrapartida, que essa mesma sensibilidade é objeto de crítica na literatura: Willmott e Matsuura (2005) mostram que a RMSE depende não apenas do erro médio, mas também da variabilidade dos erros e do número de observações, o que a torna, em sua interpretação, uma medida ambígua do erro típico, e defendem o MAE como métrica mais natural. Optou-se por manter a RMSE por ser amplamente reportada na literatura, o que facilita a comparação com outros trabalhos, e por sua complementaridade com o coeficiente de determinação.

Um aspecto metodológico relevante diz respeito à interpretação do coeficiente de determinação no contexto da validação cruzada adotada. O R^2 é fortemente sensível à variância do alvo: para uma mesma magnitude de erro absoluto, ele assume valores muito menores quando a variável prevista varia em uma faixa estreita. Nos ensaios, a tensão desviadora varia em uma faixa da ordem de 100 kPa sob confinamento de 50 kPa, mas da ordem de várias centenas de kPa sob confinamento de 300 kPa. Em consequência, um mesmo erro absoluto modesto produz coeficientes de determinação qualitativamente distintos entre os ensaios de baixo e de alto confinamento (ALTMAN; KRZYWINSKI,

2015; CHAI; DRAXLER, 2014). Por essa razão, adotou-se como métrica principal o coeficiente de determinação agregado, calculado sobre o conjunto concatenado dos 1.207 pontos de validação, que é imune à sensibilidade da variância por ensaio, complementado pela mediana e pelo intervalo interquartil dos coeficientes por ensaio, robustos à presença de valores extremos. O cálculo do coeficiente agregado a partir da validação cruzada está esquematizado na Figura 11: as previsões dos 24 ensaios de teste são reunidas em um único conjunto, sobre o qual se calculam o coeficiente de determinação e a raiz do erro quadrático médio globais. A média e o desvio padrão são reportados apenas para completude.

Figura 11 – Agregação das previsões da validação cruzada *leave-one-out* para o cálculo das métricas globais



Fonte: Elaborado pelo autor.

3.7 Ferramentas Computacionais

O modelo foi implementado na linguagem Python, com a biblioteca PyTorch (PASZKE et al., 2019) para a construção e o treinamento da rede neural. O pré-processamento e a análise dos dados utilizaram as bibliotecas NumPy, Pandas e scikit-learn, esta última responsável pela padronização das variáveis e pelo cálculo das métricas. A geração das figuras empregou a biblioteca matplotlib, com estilos de linha diferenciados para garantir a legibilidade em impressão em preto e branco.

O desenvolvimento e o treinamento foram realizados em um computador pessoal adquirido com recursos do projeto financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Processo nº 134533/2025-9), equipado com processador AMD Ryzen 7 9700X (8 núcleos, frequência base de 3,80 GHz), 64 GB de memória RAM

e placa de vídeo dedicada NVIDIA GeForce RTX 5070 Ti. O treinamento foi executado inteiramente sobre a unidade central de processamento (CPU), sem uso da aceleração gráfica disponível no equipamento. Ainda assim, o tempo total de uma execução completa da metodologia, incluindo o pré-treinamento e as 24 rodadas da validação cruzada *leave-one-out*, foi da ordem de dois minutos (8,4 s de pré-treinamento e 1,9 min para o conjunto das rodadas de ajuste fino e avaliação), o que evidencia a leveza computacional da arquitetura adotada, com seus 42.883 parâmetros treináveis.

Em favor da transparência e da reprodutibilidade, o código-fonte desenvolvido nesta pesquisa, incluindo os *scripts* de pré-treinamento, de ajuste fino, de avaliação e de geração das figuras, bem como os resultados numéricos da validação cruzada, encontram-se publicamente disponíveis no repositório <<https://github.com/joaopauloar2/dissertacao-solo-fibras-rna-ufu>>. Os conjuntos de dados sintético e experimental, assim como a planilha do modelo analítico que os originou, por constituírem material derivado de fontes de terceiros, não são distribuídos publicamente, podendo ser fornecidos pelo autor mediante solicitação acompanhada de justificativa.

Resultados e Discussão

Este capítulo apresenta e discute o desempenho do modelo constitutivo neural na previsão do comportamento tensão-deformação da areia uniforme reforçada com fibras. Os resultados foram obtidos pela validação cruzada exaustiva descrita na Seção 3.3.4, na qual cada um dos 24 ensaios foi previsto por um modelo que não o utilizou no treinamento. A discussão organiza-se do desempenho global para o desempenho por ensaio, examina os efeitos do reforço e do confinamento, ilustra a qualidade das curvas previstas, analisa o comportamento do treinamento, confronta o modelo com o modelo analítico que o instruiu, quantifica a contribuição do pré-treinamento sintético e, por fim, discute a interpretação adequada das métricas.

4.1 Desempenho Global do Modelo

A métrica principal de desempenho é o coeficiente de determinação agregado, calculado sobre o conjunto concatenado dos 1.207 pontos de validação, conforme justificado na Seção 3.6. Considerando a totalidade dos pontos previstos pela validação cruzada, o modelo atingiu coeficiente de determinação de 0,977 para a tensão desviadora (q), 0,991 para a tensão média efetiva (p') e 0,914 para a deformação volumétrica (ε_v). As raízes do erro quadrático médio correspondentes foram de 48,7 kPa para q , 17,7 kPa para p' e 0,39% para ε_v . Esses valores indicam que o modelo explica a quase totalidade da variância das tensões e a maior parte da variância da deformação volumétrica, considerando todo o espectro de configurações de reforço e de níveis de confinamento.

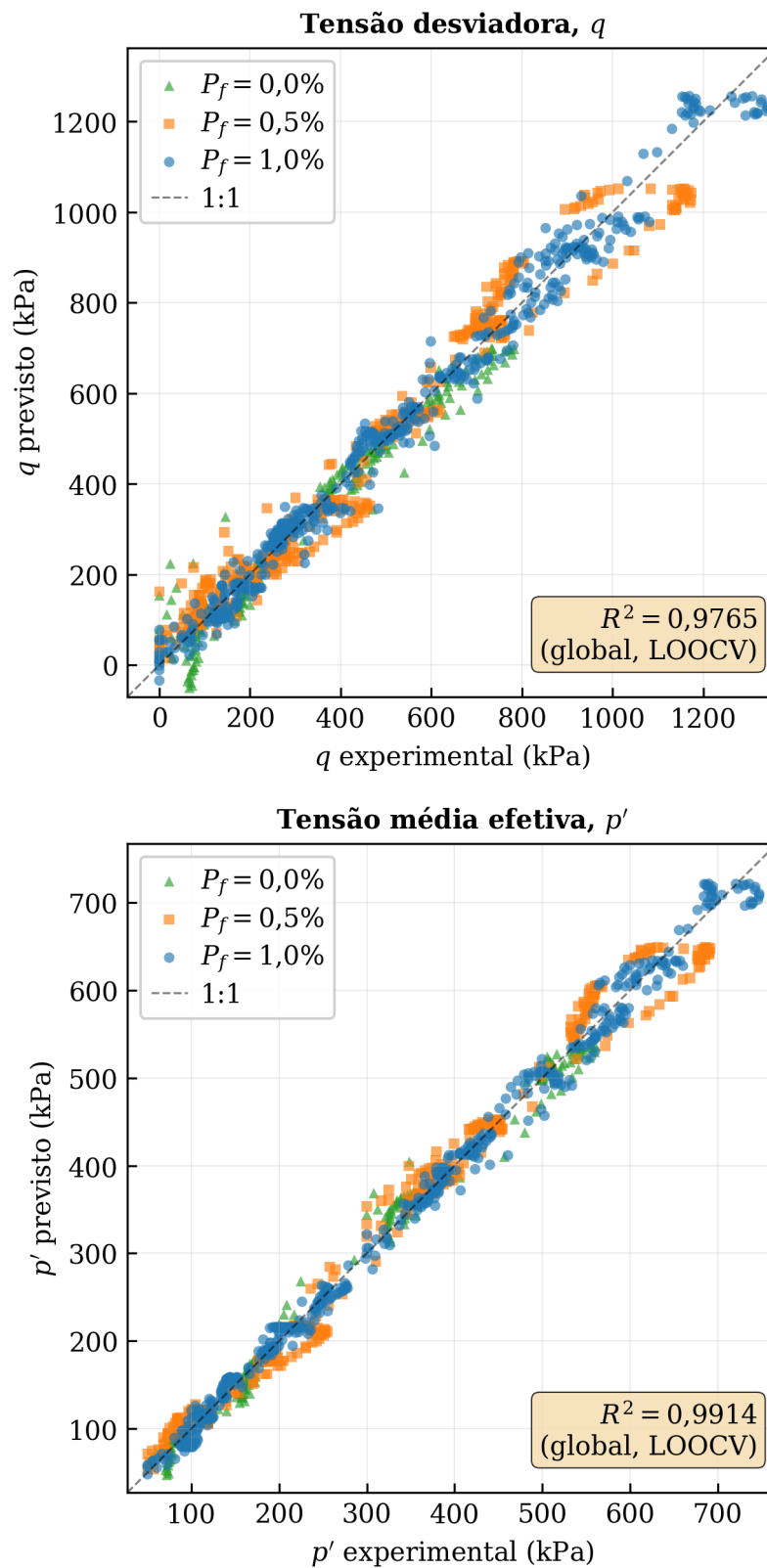
Vale observar que, dentre as três variáveis, a deformação volumétrica apresenta o menor coeficiente de determinação agregado. Esse resultado não deve ser interpretado como uma previsão volumétrica deficiente, mas como reflexo da forma de cálculo da métrica. O coeficiente agregado de q e de p' é favorecido pela ampla faixa de magnitudes que essas tensões assumem entre os níveis de confinamento, de 50 kPa a 300 kPa, o que amplia a variância de referência contra a qual o erro é comparado; a deformação

volumétrica, ao contrário, varia em uma faixa semelhante em todos os ensaios, de modo que seu coeficiente agregado não se beneficia desse efeito de escala, análogo ao discutido na Seção 4.8. A análise por ensaio, apresentada na seção seguinte, confirma essa leitura, pois é justamente a deformação volumétrica a variável de previsão mais estável entre as condições investigadas. Ainda assim, por traduzir a dilatação, um dos fenômenos fisicamente mais sutis no comportamento de areias, a resposta volumétrica permanece a saída em que um eventual refinamento do modelo seria mais proveitoso.

A Figura 12 apresenta a dispersão entre os valores previstos e os observados para a tensão desviadora e para a tensão média efetiva, com os pontos distinguidos por classe de teor de fibra. O alinhamento dos pontos em torno da reta de identidade, em toda a faixa de magnitudes, confirma a ausência de viés sistemático e a elevada capacidade preditiva do modelo nas três classes de reforço.

A dispersão restringe-se a essas duas variáveis porque ambas compartilham a mesma dimensão física, a de tensão, e assumem magnitudes comparáveis, permitindo representá-las em um único gráfico de valores previstos contra observados, com eixos de escala equivalente e uma mesma reta de identidade como referência. A deformação volumétrica, por ser adimensional, de magnitude numérica muito menor e de convenção de sinal distinta, produziria um diagrama de dispersão de leitura pouco informativa, com os pontos concentrados junto à origem e em escala incompatível com a das tensões. Por essa razão, sua qualidade preditiva é avaliada por outras vias ao longo deste capítulo: pelo coeficiente de determinação agregado já apresentado, pela distribuição por ensaio na Figura 13 e, sobretudo, pela comparação direta entre as curvas previstas e observadas na Figura 14, na qual a variação volumétrica é mostrada explicitamente em função da deformação axial.

Figura 12 – Dispersão entre valores previstos e observados na validação cruzada, para a tensão desviadora (q) e a tensão média efetiva (p')



Fonte: Elaborado pelo autor.

4.2 Distribuição do Desempenho por Ensaio

Embora a métrica agregada sintetize o desempenho do modelo, a análise da distribuição dos coeficientes de determinação por ensaio revela como a qualidade da previsão varia entre as diferentes condições. A Tabela 7 apresenta os resultados das 24 rodadas de validação, e a Figura 13 sintetiza essa distribuição por classe de teor de fibra.

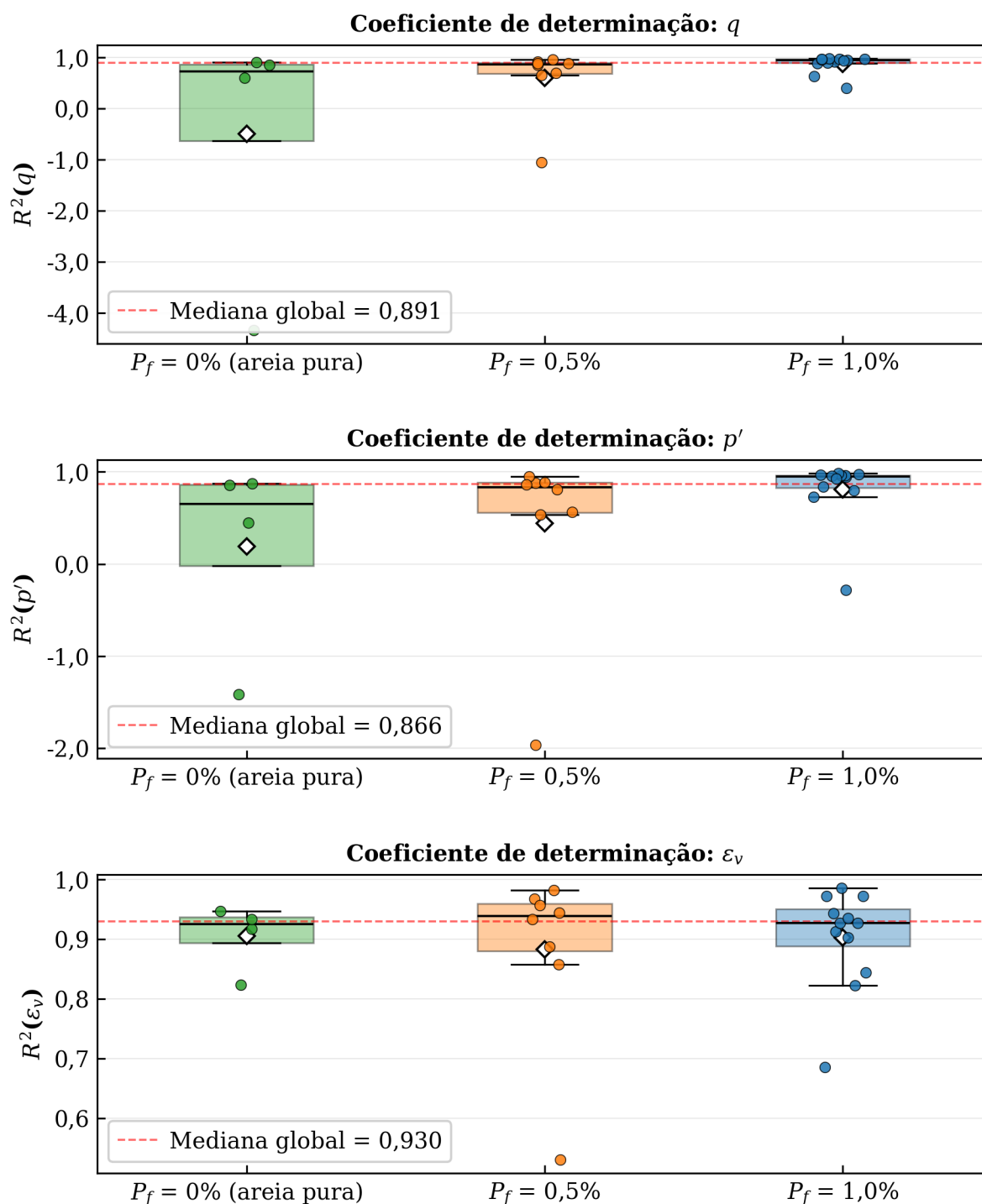
Tabela 7 – Resultados da validação cruzada *leave-one-out* por ensaio

σ_3 (kPa)	P_f (%)	L (mm)	R^2			RMSE		
			q	p'	ε_v	q (kPa)	p' (kPa)	ε_v
50	0,0	-	-4,332	-1,416	0,917	55,9	12,6	0,0052
50	0,5	25	-1,055	-1,964	0,968	39,2	15,7	0,0022
50	0,5	51	0,651	0,534	0,982	27,0	10,4	0,0019
50	1,0	12,5	0,404	-0,283	0,912	27,0	13,2	0,0043
50	1,0	25	0,635	0,726	0,927	27,3	7,9	0,0036
50	1,0	51	0,887	0,836	0,972	26,7	10,8	0,0021
100	0,0	-	0,597	0,445	0,947	33,1	13,0	0,0033
100	0,5	25	0,888	0,880	0,531	23,2	8,0	0,0063
100	0,5	51	0,693	0,564	0,858	63,4	25,2	0,0061
100	1,0	12,5	0,918	0,941	0,686	20,1	5,7	0,0063
100	1,0	25	0,895	0,798	0,823	28,1	13,0	0,0069
100	1,0	51	0,947	0,958	0,972	30,5	9,0	0,0017
200	0,0	-	0,909	0,854	0,934	36,8	15,5	0,0027
200	0,5	25	0,917	0,886	0,933	37,7	14,7	0,0036
200	0,5	51	0,962	0,948	0,956	36,6	14,2	0,0027
200	1,0	12,5	0,978	0,958	0,985	19,8	9,0	0,0016
200	1,0	25	0,962	0,982	0,927	33,7	7,7	0,0033
200	1,0	51	0,974	0,964	0,903	39,7	15,6	0,0027
300	0,0	-	0,855	0,874	0,824	70,4	21,9	0,0046
300	0,5	25	0,854	0,810	0,888	81,1	30,8	0,0037
300	0,5	51	0,884	0,858	0,945	101,6	37,4	0,0030
300	1,0	12,5	0,968	0,956	0,944	42,4	16,5	0,0032
300	1,0	25	0,941	0,926	0,935	64,8	24,2	0,0035
300	1,0	51	0,969	0,971	0,844	68,4	22,1	0,0036

Fonte: Elaborado pelo autor.

A mediana dos coeficientes de determinação por ensaio foi de 0,891 para q , 0,866 para p' e 0,930 para ε_v , com intervalos interquartis de 0,267, 0,265 e 0,069, respectivamente. A acentuada diferença entre a média e a mediana dos coeficientes de q e p' , evidenciada na Figura 13 pela posição dos losangos (média) bem abaixo das medianas, decorre de um pequeno número de ensaios com coeficientes fortemente negativos, todos associados ao confinamento de 50 kPa. Esse comportamento, longe de indicar falha do modelo, reflete uma propriedade da própria métrica, discutida na Seção 4.8. A deformação volumétrica, cuja faixa de variação é mais homogênea entre os ensaios, apresenta um intervalo interquartil estreito, indicando previsões consistentes com os resultados experimentais.

Figura 13 – Distribuição dos coeficientes de determinação por ensaio, segundo a classe de teor de fibra



Fonte: Elaborado pelo autor.

4.3 Efeito do Reforço e do Confinamento

O desempenho do modelo para a tensão desviadora melhora de forma consistente com o aumento do teor de fibra. A mediana do coeficiente de determinação de q foi de 0,726

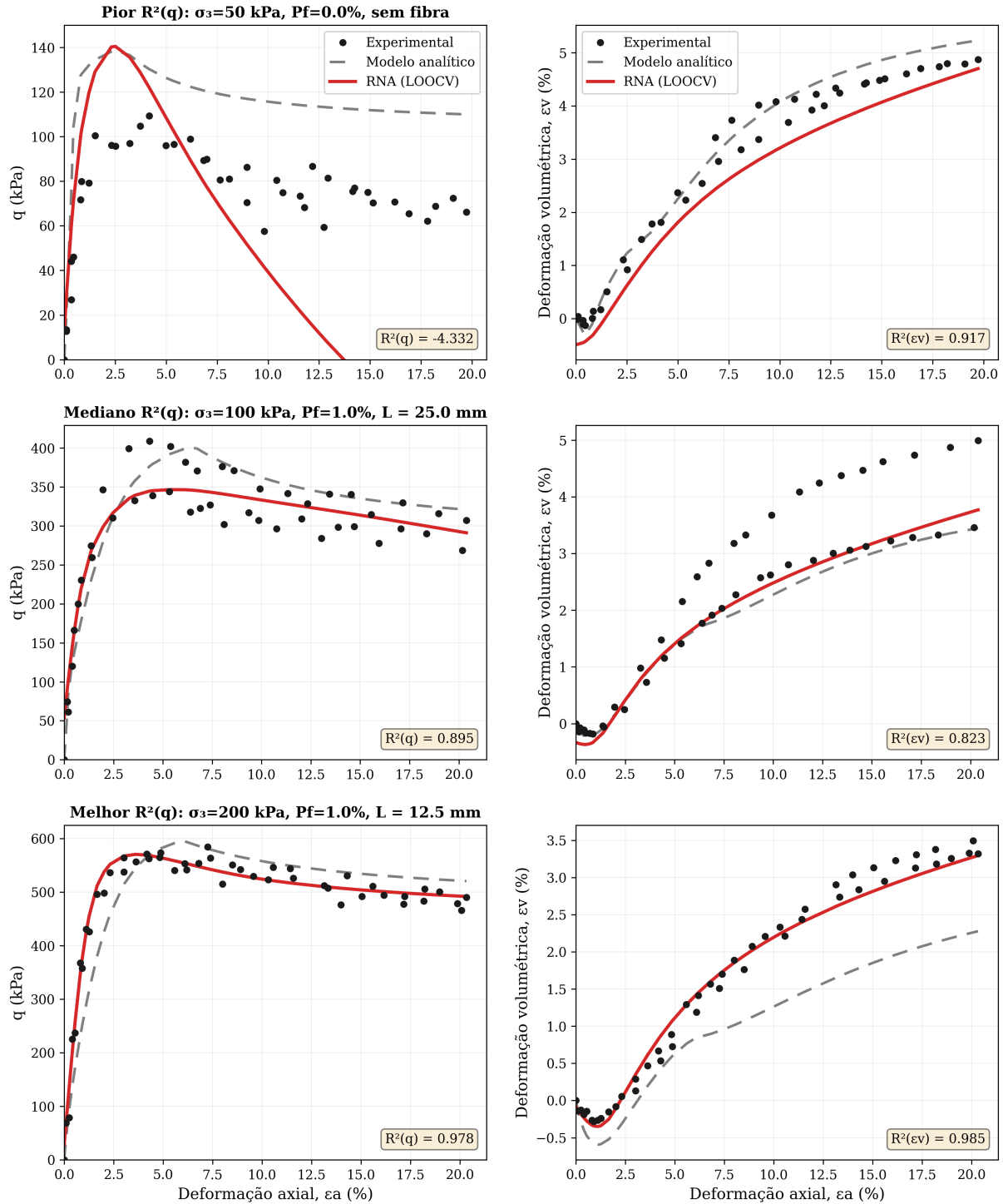
para a areia sem reforço, 0,869 para o teor de 0,5% e 0,944 para o teor de 1,0%. Essa tendência tem explicação física direta. Conforme observado experimentalmente por Conceição, Carvalho e Machado (2022), sob a tensão de confinamento de 50 kPa praticamente não há incremento de resistência com a inclusão das fibras, pois o confinamento é insuficiente para promover a ancoragem do reforço entre os grãos. Nessas condições, a resposta da areia reforçada aproxima-se da resposta da areia pura, cuja curva tensão-deformação varia em uma faixa estreita e é mais sensível ao ruído experimental.

De forma complementar, observa-se na Tabela 7 que o desempenho melhora com o aumento do confinamento até as tensões intermediárias e elevadas. Sob 200 kPa e 300 kPa, os coeficientes de determinação de q situam-se majoritariamente acima de 0,85, mesmo para a areia sem reforço, refletindo a maior amplitude da resposta de tensões e a consequente robustez da métrica nessas condições.

4.4 Curvas Tensão-Deformação Representativas

A Figura 14 apresenta a comparação entre as curvas previstas e as observadas para três ensaios representativos, selecionados pelos valores de coeficiente de determinação de q : o de pior ajuste (50 kPa, areia sem fibra), um de ajuste mediano (100 kPa, $P_f = 1,0\%$, $L = 25$ mm) e o de melhor ajuste (200 kPa, $P_f = 1,0\%$, $L = 12,5$ mm). A linha cheia representa a previsão da rede, a linha tracejada representa a previsão do modelo analítico de referência para o mesmo cenário, discutido na Seção 4.6, e os pontos representam os dados experimentais. O contraste entre as duas linhas torna visível o refinamento operado pelo ajuste fino, particularmente nos painéis de variação volumétrica, em que a trajetória analítica se afasta mais dos dados e a rede a corrige na direção do comportamento observado. Nos painéis de variação volumétrica, o sinal de ε_v é exibido invertido em relação à convenção adotada na base de dados (Seção 3.1.3), de modo que valores positivos indicam dilatação, forma usual de apresentação do comportamento dilatante de areias.

Figura 14 – Curvas tensão-deformação e de variação volumétrica previstas e observadas, para os ensaios de pior, mediano e melhor ajuste de q , com a previsão do modelo analítico de referência (Seção 4.6)



Fonte: Elaborado pelo autor.

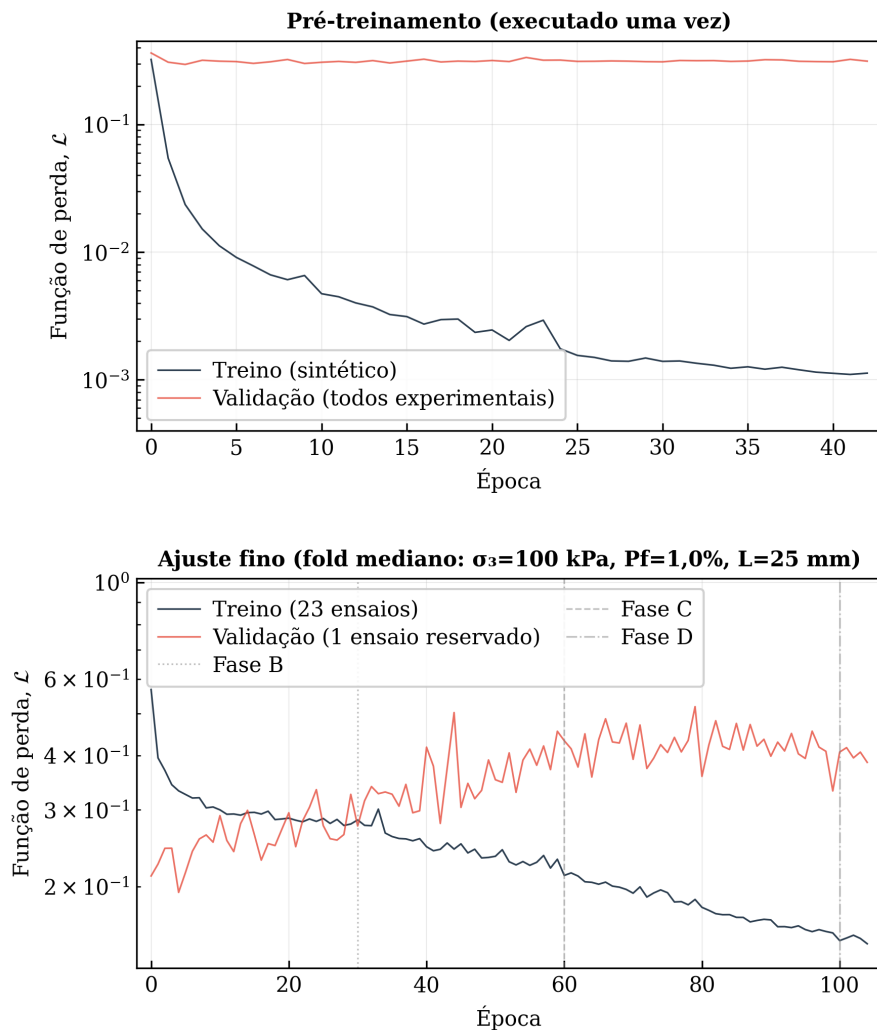
Mesmo no ensaio de pior ajuste quantitativo, a previsão captura corretamente os aspectos qualitativos essenciais do comportamento: o pico de resistência, o amolecimento pós-pico e a dilatação progressiva da deformação volumétrica. Nesse ensaio, os dados experimentais apresentam dispersão pronunciada e a tensão desviadora varia em uma faixa

reduzida, condições que penalizam fortemente o coeficiente de determinação sem que a forma da curva esteja incorreta. Nos ensaios de ajuste mediano e melhor, a aderência entre a previsão e os dados é elevada tanto para a tensão desviadora quanto para a deformação volumétrica, incluindo a transição de compressão inicial para dilatância. Vale notar que, mesmo quando o coeficiente de determinação de q é baixo, o coeficiente correspondente de ε_v permanece elevado, o que reforça a interpretação de que a deficiência de ajuste se concentra na escala da tensão desviadora, e não na física do comportamento volumétrico. As curvas previstas e observadas para a totalidade dos vinte e quatro ensaios da validação cruzada podem ser visualizadas no Apêndice A.

4.5 Comportamento do Treinamento

A Figura 15 apresenta a evolução da função de perda durante o pré-treinamento, executado uma única vez, e durante o ajuste fino de um dos ensaios, em escala logarítmica.

Figura 15 – Evolução da função de perda no pré-treinamento e no ajuste fino



Fonte: Elaborado pelo autor.

No pré-treinamento, a perda sobre os dados sintéticos decresce cerca de duas ordens de grandeza, indicando que a rede aprende com fidelidade o comportamento descrito pelo modelo analítico. A perda de validação sobre os dados experimentais permanece elevada e aproximadamente constante nessa fase, o que é esperado e não constitui deficiência: ela mede precisamente a distância entre a teoria, que a rede aprendeu, e o comportamento experimental real, distância essa que o ajuste fino tem por função reduzir. É justamente essa estabilização precoce da perda de validação que aciona a interrupção antecipada e encerra o pré-treinamento muito antes do teto de épocas. No ajuste fino, a perda de treinamento decresce de forma sustentada, enquanto a perda de validação, após uma redução inicial, volta a crescer. Esse crescimento caracteriza um sobreajuste aparente nas épocas finais. Como a estratégia de treinamento preserva o estado de melhor desempenho na validação, o modelo efetivamente adotado em cada rodada corresponde ao ponto de mínima perda de validação, de modo que o sobreajuste observado nas épocas posteriores não afeta o desempenho reportado. As linhas verticais assinalam a liberação progressiva das camadas ao longo do ajuste fino.

4.6 Comparação com o Modelo Analítico de Referência

O modelo analítico de Machado et al. (2024), empregado na geração dos dados sintéticos, constitui a referência natural contra a qual o desempenho da rede deve ser aferido. Sem essa comparação, os coeficientes de determinação apresentados nas seções anteriores careceriam de termo de contraste, e não se poderia julgar se a rede representa, de fato, um avanço em relação ao conhecimento constitutivo já disponível para o material.

A aferição foi conduzida da seguinte forma. Para cada um dos 24 ensaios, tomou-se a trajetória gerada pelo modelo analítico para o mesmo cenário, definido pela tensão de confinamento, pelo teor e pelo comprimento de fibra, e interpolaram-se as previsões nos valores de deformação axial em que há medida experimental. Aplicaram-se então as mesmas métricas e a mesma agregação utilizadas para a rede, sobre os mesmos 1.207 pares previsto-observado. A Tabela 8 reúne os resultados.

Tabela 8 – Desempenho do modelo analítico de referência e da rede sobre os 1.207 pontos experimentais

Métrica	Modelo analítico	Rede (LOOCV)
R^2 agregado, q	0,944	0,977
R^2 agregado, p'	0,983	0,991
R^2 agregado, ε_v	0,592	0,914
RMSE, q (kPa)	75,2	48,7
RMSE, p' (kPa)	25,1	17,7
RMSE, ε_v (%)	0,86	0,39
Mediana de $R^2(q)$ por ensaio	0,742	0,891
Mediana de $R^2(\varepsilon_v)$ por ensaio	0,581	0,930
Ensaio com $R^2(q) < 0,7$	6 de 24	4 de 24

Fonte: Elaborado pelo autor.

O contraste mais expressivo ocorre na deformação volumétrica: o modelo analítico explica 59% da variância experimental, enquanto a rede explica 91%. É justamente ε_v que carrega a assinatura mecânica do reforço fibroso, a transição de compressão para dilatância, e é nela que a forma funcional imposta pelo modelo analítico se mostra mais restritiva. Nas tensões, o ganho é mais discreto, porém consistente: a raiz do erro quadrático médio da tensão desviadora reduz-se de 75,2 kPa para 48,7 kPa, o que corresponde a uma diminuição de 35%, e a mediana do coeficiente de determinação por ensaio eleva-se de 0,742 para 0,891.

Convém observar que a comparação não é simétrica, e que a assimetria opera em favor do modelo analítico. Seus parâmetros foram calibrados por Machado et al. (2024) a partir desses mesmos ensaios, de modo que o modelo os reproduz em condição de dado presente na própria amostra de calibração. A rede, por sua vez, prevê cada ensaio por validação cruzada, sem que os pontos desse ensaio tenham participado do seu treinamento. A superação observada é, portanto, obtida sob condições que desfavorecem a rede, e deve ser interpretada como uma estimativa inferior da vantagem que ela oferece.

O resultado esclarece o papel de cada componente da abordagem. Todo modelo constitutivo analítico impõe uma forma funcional fixa, e o erro devido a essa forma não é redutível por calibração, por mais cuidadosa que ela seja. A rede, na qualidade de aproximador universal, captura o que a forma funcional deixa de fora. O modelo analítico não é, portanto, concorrente da rede: é a fonte da estrutura física que 1.207 pontos experimentais não seriam suficientes para determinar. Os dados experimentais, por sua vez, fornecem a correção que o modelo analítico não alcança. O produto final supera cada uma das fontes tomada isoladamente.

4.7 Contribuição do Pré-Treinamento Sintético

Estabelecido que a rede supera o modelo analítico que a instruiu, resta quantificar o quanto dessa capacidade se deve ao pré-treinamento sintético. Para isso, conduziu-se um

experimento de remoção controlada, procedimento conhecido na literatura de aprendizado de máquina como estudo de ablação (*ablation study*) (MEYES et al., 2019), no qual se suprime um componente do método, mantendo-se todo o restante inalterado, e se mede a degradação resultante do desempenho.

Compararam-se duas configurações de treinamento, ambas submetidas ao mesmo protocolo de validação cruzada da Seção 3.3.4. Na primeira, reproduziu-se integralmente a estratégia híbrida descrita no Capítulo 3, com o pré-treinamento sintético. Na segunda, suprimiu-se o pré-treinamento: a rede parte de pesos aleatórios e aprende apenas com os 23 ensaios experimentais de cada rodada, o que simula a situação em que não se dispõe de um modelo analítico gerador. Para que a comparação não penalizasse indevidamente a configuração sem pré-treinamento, adotaram-se, para ela, condições deliberadamente favoráveis: dispensou-se o descongelamento progressivo, que não teria sentido na ausência de conhecimento prévio a preservar; ampliou-se o teto de épocas de 400 para 500; e a normalização foi ajustada nos 23 ensaios de treino de cada rodada, e não nos dados sintéticos, que não existiriam no cenário simulado. Preservaram-se, nas duas configurações, a mesma arquitetura, a mesma função de perda ponderada, os mesmos 24 folds e o mesmo critério de parada.

Como o desempenho de uma rede neural depende em alguma medida do sorteio aleatório que fixa os seus pesos iniciais, o experimento foi repetido cinco vezes, cada repetição com uma inicialização distinta, obtida por uma semente diferente do gerador de números aleatórios. Reportam-se, assim, a média e o desvio padrão sobre as cinco repetições, e não o resultado de uma única execução, o que evita atribuir a uma escolha particular de pesos iniciais um efeito que poderia ser fortuito. A Tabela 9 reúne esses valores.

Tabela 9 – Estudo de ablação: desempenho com e sem o pré-treinamento sintético, em média sobre cinco repetições independentes do experimento

Métrica (média $\pm$ desvio padrão)	Com pré-treinamento	Sem pré-treinamento
R^2 agregado, q	$0,978 \pm 0,002$	$0,972 \pm 0,001$
R^2 agregado, p'	$0,992 \pm 0,001$	$0,989 \pm 0,001$
R^2 agregado, ε_v	$0,915 \pm 0,001$	$0,900 \pm 0,003$
RMSE, q (kPa)	$46,9 \pm 1,8$	$53,4 \pm 1,1$
RMSE, p' (kPa)	$17,0 \pm 0,8$	$20,3 \pm 0,8$
RMSE, ε_v (%)	$0,393 \pm 0,003$	$0,426 \pm 0,007$
Mediana de $R^2(q)$ por ensaio	$0,903 \pm 0,010$	$0,852 \pm 0,019$
Mediana de $R^2(p')$ por ensaio	$0,875 \pm 0,012$	$0,793 \pm 0,042$
Mediana de $R^2(\varepsilon_v)$ por ensaio	$0,934 \pm 0,007$	$0,917 \pm 0,006$
Ensaios com $R^2(q) < 0,7$	4 de 24	7 de 24

Fonte: Elaborado pelo autor.

Em todas as cinco repetições, e em todas as métricas da Tabela 9, a configuração com pré-treinamento superou a configuração sem pré-treinamento, sem uma única inversão. A regularidade desse resultado, aliada aos desvios padrão reduzidos, indica que a

vantagem observada não decorre de uma inicialização particularmente favorável, mas do pré-treinamento em si.

A interpretação desse resultado requer, contudo, distinguir dois modos de medir o desempenho, que aqui conduzem a conclusões diferentes. O primeiro é o coeficiente de determinação agregado, calculado sobre a totalidade dos 1.207 pontos reunidos. Por essa medida, a diferença entre as duas configurações é pequena: sem o pré-treinamento, a rede alcança 0,972 na tensão desviadora, contra 0,978 da estratégia híbrida. O resultado é esperado. Com 23 dos 24 ensaios cobrindo densamente a grade de configurações, prever o ensaio restante é, em boa parte, um problema de interpolação entre ensaios vizinhos, para o qual os dados experimentais já bastam; além disso, o coeficiente agregado é dominado pela grande variação da resposta entre os níveis de confinamento e tende, por isso, a permanecer elevado, conforme discutido na Seção 4.8.

O segundo modo é examinar cada ensaio isoladamente, e é nele que a contribuição do pré-treinamento se revela. Avaliada ensaio a ensaio, a qualidade das previsões torna-se nitidamente mais uniforme quando o pré-treinamento está presente. A mediana do coeficiente de determinação por ensaio reduz-se de 0,903 para 0,852 na tensão desviadora e de 0,875 para 0,793 na tensão média efetiva quando ele é suprimido; o número de ensaios mal descritos, com coeficiente inferior a 0,7 na tensão desviadora, sobe de quatro para sete; e a raiz do erro quadrático médio, calculada sobre todos os pontos, cresce 13,9% na tensão desviadora, 19,8% na tensão média efetiva e 8,4% na deformação volumétrica. Em outras palavras, sem o pré-treinamento a rede mantém, em média, desempenho elevado, mas apresenta degradação acentuada em ensaios específicos, de forma pouco previsível. O pré-treinamento não eleva tanto o desempenho médio quanto assegura a sua uniformidade entre os ensaios. É precisamente essa confiabilidade individual, e não a média sobre o conjunto, que se exige de um modelo constitutivo, empregado na prática para prever uma condição de cada vez.

A distribuição do ganho entre os ensaios reforça essa leitura. Discriminado por nível de confinamento, o pré-treinamento eleva a mediana de $R^2(q)$ de 0,798 para 0,888 a 100 kPa e de 0,913 para 0,959 a 200 kPa, mas praticamente não a altera a 300 kPa, em que passa de 0,933 para 0,934. Nesse último nível, a rede sem pré-treinamento já opera próximo ao teto de desempenho, e não há margem para melhoria. Sob 50 kPa, nenhuma das duas estratégias descreve satisfatoriamente o comportamento, o que é coerente com a ancoragem insuficiente das fibras e com a menor identificabilidade física discutidas na Seção 4.8. O benefício do pré-treinamento concentra-se, portanto, na faixa em que há simultaneamente margem para melhoria e informação útil a transferir, o que corrobora a associação, registrada nas conclusões, entre a eficácia da abordagem e a coerência do gerador sintético com a base experimental.

4.8 Discussão da Métrica sob Baixo Confinamento

A ocorrência de coeficientes de determinação negativos em alguns ensaios de 50 kPa merece interpretação cuidadosa. O coeficiente de determinação compara o erro do modelo à variância dos dados em torno de sua média. Quando a variável prevista varia em uma faixa estreita, como a tensão desviadora sob baixo confinamento, da ordem de 100 kPa, a variância de referência é pequena, e um erro absoluto modesto, da ordem de algumas dezenas de kPa, é suficiente para produzir um coeficiente negativo, ainda que a curva prevista seja qualitativamente correta (CHAI; DRAXLER, 2014; ALTMAN; KRZYWINSKI, 2015). Esse efeito é agravado pela maior dispersão experimental relativa observada nesses ensaios e pela já mencionada ausência de efeito do reforço sob 50 kPa.

Por essa razão, o coeficiente de determinação por ensaio não constitui, isoladamente, uma medida adequada de desempenho quando as escalas das respostas são heterogêneas, como ocorre ao se reunir ensaios de baixo e de alto confinamento. A adoção do coeficiente agregado como métrica principal, complementado pela mediana e pelo intervalo interquartil, evita que essa sensibilidade da métrica seja confundida com deficiência do modelo. As curvas representativas da Figura 14 confirmam que, mesmo nos ensaios de coeficiente negativo, o modelo reproduz adequadamente a física do comportamento mecânico.

Cabe, no entanto, uma distinção adicional. Afirmar que o coeficiente negativo é uma propriedade da métrica não equivale a afirmar que o comportamento sob baixo confinamento esteja tão bem determinado quanto sob confinamentos elevados. A faixa estreita de tensões, a mobilização reduzida do reforço e a maior dispersão experimental relativa fazem dessa região uma zona de menor identificabilidade física, na qual os dados restringem de forma mais frouxa a resposta a ser aprendida e as previsões carregam, por isso, maior incerteza, independentemente do valor assumido pela métrica. Em termos mais gerais, um coeficiente de determinação agregado elevado atesta a adequação estatística do ajuste, mas não basta, isoladamente, para garantir a fidelidade física das previsões. É por essa razão que a avaliação não se apoia apenas em índices numéricos: a análise qualitativa das curvas previstas, na Figura 14, fornece a evidência complementar de que a forma do comportamento mecânico, e não somente a sua escala, é reproduzida pelo modelo.

Conclusões

Esta dissertação desenvolveu um modelo constitutivo baseado em Redes Neurais Artificiais para a previsão do comportamento tensão-deformação de uma areia uniforme reforçada com fibras poliméricas. O modelo foi construído como um mapeamento direto, no qual a configuração do ensaio, dada pela tensão de confinamento, pelo teor e pelo comprimento de fibra, juntamente com o nível de deformação axial, é mapeada nas variáveis de estado de interesse: a tensão desviadora, a tensão média efetiva e a deformação volumétrica. O treinamento adotou uma estratégia híbrida em duas fases, com pré-treinamento sobre um grande volume de dados sintéticos gerados pelo modelo analítico de Machado et al. (2024) e ajuste fino, por aprendizado por transferência, sobre os dados experimentais de Conceição, Carvalho e Machado (2022).

A partir dos resultados obtidos, destacam-se as conclusões a seguir.

A estratégia híbrida mostrou-se eficaz, e o modelo resultante supera o modelo analítico que o instruiu. Sobre os mesmos 1.207 pontos experimentais, o modelo de Machado et al. (2024) alcança coeficiente de determinação de 0,944 para a tensão desviadora e de 0,592 para a deformação volumétrica, contra 0,977 e 0,914 obtidos pela rede em validação cruzada, conforme a Seção 4.6. O contraste é tanto mais significativo quanto se observe que o modelo analítico prevê esses ensaios em condição de dado visto, pois foi calibrado com eles, ao passo que a rede jamais os utilizou no treinamento. O ganho concentra-se na deformação volumétrica, precisamente a variável que carrega a assinatura mecânica do reforço fibroso, e evidencia que o erro decorrente da forma funcional imposta por um modelo analítico não é redutível por calibração.

A contribuição do pré-treinamento está na consistência, e não na média. O estudo de ablação da Seção 4.7 mostrou que uma rede treinada apenas com os dados experimentais também atinge coeficiente agregado elevado, da ordem de 0,972 para a tensão desviadora. O que se perde ao suprimir o pré-treinamento é a confiabilidade de cada previsão individual: a mediana do coeficiente de determinação por ensaio cai de 0,903 para 0,852 na tensão desviadora e de 0,875 para 0,793 na tensão média efetiva, o número de ensaios com ajuste deficiente sobe de quatro para sete, e a raiz do erro quadrático médio global

crece entre 8% e 20% conforme a variável. É essa confiabilidade caso a caso, e não o desempenho médio sobre o conjunto, que se exige de um modelo constitutivo. A coerência entre a fonte dos dados experimentais e a fonte dos dados sintéticos, ambos referentes ao mesmo material, mostrou-se determinante: o ganho do pré-treinamento concentra-se nos níveis de confinamento em que o modelo analítico é fiel ao comportamento observado, e desaparece onde não o é.

O modelo apresentou alta capacidade preditiva. Avaliado por validação cruzada exaustiva do tipo *leave-one-out*, na qual cada ensaio é previsto por um modelo que não o utilizou no treinamento, o coeficiente de determinação agregado sobre os 1.207 pontos de validação foi de 0,977 para a tensão desviadora, 0,991 para a tensão média efetiva e 0,914 para a deformação volumétrica. As curvas previstas reproduzem corretamente os aspectos essenciais do comportamento mecânico, incluindo o pico de resistência, o amolecimento pós-pico e a transição de compressão para dilatação.

O desempenho do modelo acompanha os mecanismos físicos do reforço. A qualidade das previsões melhora de forma sistemática com o aumento do teor de fibra e do nível de confinamento. Esse comportamento é coerente com a observação experimental de que, sob baixo confinamento, a ancoragem das fibras é insuficiente e o efeito do reforço é reduzido.

A interpretação das métricas exige cautela. Os coeficientes de determinação negativos observados em alguns ensaios de baixo confinamento decorrem da reduzida variância da tensão desviadora nessas condições, uma propriedade da própria métrica, e não de deficiência do modelo, ainda que o baixo confinamento permaneça a condição de maior incerteza preditiva. Esse achado reforça a importância de adotar uma métrica agregada e medidas robustas, como a mediana e o intervalo interquartil, na avaliação de modelos sobre conjuntos de respostas com escalas heterogêneas.

5.1 Limitações

O trabalho apresenta limitações que devem ser consideradas na interpretação dos resultados. O modelo foi desenvolvido e validado para um único material, uma areia uniforme com uma fibra de copolímero específica, de modo que sua generalização para outros solos e fibras não foi avaliada. A base de treinamento restringiu-se às trajetórias de compressão triaxial drenada convencional, não contemplando trajetórias de tensão não convencionais. No plano da avaliação, os critérios de parada do treinamento foram informados pelos dados experimentais, inclusive, no ajuste fino, pelo próprio ensaio reservado de cada rodada; embora nenhum gradiente seja calculado sobre esses dados, a seleção do ponto de parada com base neles tende a introduzir leve otimismo nas métricas por ensaio (CAWLEY; TALBOT, 2010). Além disso, o modelo prevê a resposta tensão-deformação, mas não foi integrado a um ambiente de análise por elementos finitos, etapa que exigiria a obtenção do operador tangente consistente a partir da rede. Por fim, a estratégia adotada pressupõe

a existência de um modelo analítico capaz de gerar dados de pré-treinamento fisicamente consistentes para o material de interesse. A eficácia da abordagem mostrou-se associada à coerência entre o gerador sintético e a base experimental, ambos relativos ao mesmo material; o estudo de ablação da Seção 4.7 quantificou o custo de sua supressão para a presente base, mas a transposição da estratégia para situações em que não se disponha de um modelo analítico adequado, ou em que este esteja mal calibrado, permanece não avaliada e não pode ser presumida, sobretudo em bases experimentais mais esparsas do que a aqui empregada.

5.2 Sugestões para Trabalhos Futuros

A partir das limitações identificadas, sugerem-se as seguintes direções para trabalhos futuros:

- ❑ incorporar as trajetórias de tensão não convencionais disponíveis no programa experimental, com a correspondente ampliação da geração de dados sintéticos, de modo a generalizar o modelo para diferentes caminhos de carregamento;
- ❑ avaliar a transferência do modelo para outros materiais, incluindo a areia aluvionar de Pinto (2021), ensaiada com a mesma fibra e com resultados publicados em Pinto e Machado (2022), como segundo material de validação;
- ❑ integrar o modelo a um ambiente de análise por elementos finitos, mediante a obtenção do operador tangente consistente a partir da rede, viabilizando a sua aplicação na solução de problemas de valor de contorno.

Referências

- AKIBA, T. et al. Optuna: A next-generation hyperparameter optimization framework. In: **Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**. Anchorage, AK: [s.n.], 2019. p. 2623–2631. Disponível em: <<https://doi.org/10.1145/3292500.3330701>>.
- ALTMAN, N.; KRZYWINSKI, M. Points of significance: Simple linear regression. **Nature Methods**, Nature Publishing Group, v. 12, n. 11, p. 999–1000, 2015. Disponível em: <<https://doi.org/10.1038/nmeth.3627>>.
- BISHOP, C. M. **Pattern Recognition and Machine Learning**. New York: Springer, 2006. (Information Science and Statistics).
- CALVELLO, M.; FINNO, R. J. Selecting parameters to optimize in model calibration by inverse analysis. **Computers and Geotechnics**, Elsevier, v. 31, n. 5, p. 410–424, 2004. Disponível em: <<https://doi.org/10.1016/j.compgeo.2004.03.004>>.
- CAWLEY, G. C.; TALBOT, N. L. C. On over-fitting in model selection and subsequent selection bias in performance evaluation. **Journal of Machine Learning Research**, v. 11, p. 2079–2107, 2010.
- CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. **Geoscientific Model Development**, Copernicus GmbH, v. 7, n. 3, p. 1247–1250, 2014. Disponível em: <<https://doi.org/10.5194/gmd-7-1247-2014>>.
- CHOU, J.-S.; YANG, K.-H.; LIN, J.-Y. Peak shear strength of discrete fiber-reinforced soils computed by machine learning and metaensemble methods. **Journal of Computing in Civil Engineering**, v. 30, n. 6, p. 04016036, 2016. Disponível em: <[https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000595](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000595)>.
- CLEVERT, D.-A.; UNTERTHINER, T.; HOCHREITER, S. Fast and accurate deep network learning by exponential linear units (ELUs). In: **4th International Conference on Learning Representations (ICLR)**. San Juan, Porto Rico: [s.n.], 2016. Disponível em: <<https://doi.org/10.48550/arXiv.1511.07289>>.
- CONCEIÇÃO, M. P. d. S. **Análise do comportamento tensão-deformação de um solo reforçado com fibra**. 93 p. Dissertação (Dissertação (Mestrado em Engenharia Civil) – Escola Politécnica) — Universidade Federal da Bahia, Salvador, 2021.

- CONCEIÇÃO, M. P. d. S.; CARVALHO, M. d. F.; MACHADO, S. L. Análise do comportamento tensão-deformação de uma areia uniforme reforçada com fibra polimérica. In: **XX Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica**. Campinas, SP: ABMS, 2022. p. 7552–7559. Disponível em: <<https://doi.org/10.4322/cobramseg.2022.0950>>.
- CONCEIÇÃO, M. P. d. S. et al. Fiber reinforcement effectiveness in two different sand specimens. **Soils and Rocks**, v. 46, n. 2, p. e2023012422, 2023. Disponível em: <<https://doi.org/10.28927/SR.2023.012422>>.
- CONSOLI, N. C. et al. Fiber reinforcement effects on sand considering a wide cementation range. **Geotextiles and Geomembranes**, Elsevier, v. 27, n. 3, p. 196–203, 2009. Disponível em: <<https://doi.org/10.1016/j.geotexmem.2008.11.005>>.
- DIAMBRA, A.; IBRAIM, E. Fibre-reinforced sand: interaction at the fibre and grain scale. **Géotechnique**, ICE Publishing, v. 65, n. 4, p. 296–308, 2015. Disponível em: <<https://doi.org/10.1680/geot.14.P.206>>.
- DIAMBRA, A. et al. Fibre reinforced sands: from experiments to modelling and beyond. **International Journal for Numerical and Analytical Methods in Geomechanics**, Wiley, v. 37, p. 2427–2455, 2013. Disponível em: <<https://doi.org/10.1002/nag.2142>>.
- DIAMBRA, A. et al. Fibre reinforced sands: Experiments and modelling. **Geotextiles and Geomembranes**, Elsevier, v. 28, n. 3, p. 238–250, 2010. Disponível em: <<https://doi.org/10.1016/j.geotexmem.2009.09.010>>.
- FRENCH, R. M. Catastrophic forgetting in connectionist networks. **Trends in Cognitive Sciences**, Elsevier, v. 3, n. 4, p. 128–135, 1999. Disponível em: <[https://doi.org/10.1016/S1364-6613\(99\)01294-2](https://doi.org/10.1016/S1364-6613(99)01294-2)>.
- GHABOUSSI, J.; JR, J. H. G.; WU, X. Knowledge-based modeling of material behavior with neural networks. **Journal of Engineering Mechanics**, ASCE, v. 117, n. 1, p. 132–153, 1991. Disponível em: <<https://doi.org/frq7pb>>.
- GLOROT, X.; BORDES, A.; BENGIO, Y. Deep sparse rectifier neural networks. In: **Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS)**. Fort Lauderdale, FL: [s.n.], 2011. (Proceedings of Machine Learning Research, v. 15), p. 315–323.
- GOH, A. T. C.; GOH, S. H. Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data. **Computers and Geotechnics**, Elsevier, v. 34, n. 5, p. 410–421, 2007. Disponível em: <<https://doi.org/10.1016/j.compgeo.2007.06.001>>.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>.
- GRAY, D. H.; OHASHI, H. Mechanics of fiber reinforcement in sand. **Journal of Geotechnical Engineering**, ASCE, v. 109, n. 3, p. 335–353, 1983. Disponível em: <[https://doi.org/10.1061/\(ASCE\)0733-9410\(1983\)109:3\(335\)](https://doi.org/10.1061/(ASCE)0733-9410(1983)109:3(335))>.

- HASHASH, Y. M. A.; JUNG, S.; GHABOUSSI, J. Numerical implementation of a neural network based material model in finite element analysis. **International Journal for Numerical Methods in Engineering**, Wiley, v. 59, n. 7, p. 989–1005, 2004. Disponível em: <<https://doi.org/10.1002/nme.905>>.
- HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. 2. ed. New York: Springer, 2009. (Springer Series in Statistics). Disponível em: <<https://doi.org/cd7nhz>>.
- HAYKIN, S. **Neural networks and learning machines**. 3. ed. Upper Saddle River, NJ: Pearson, 2009.
- HEINECK, K. S.; COOP, M. R.; CONSOLI, N. C. Effect of microreinforcement of soils from very small to large shear strains. **Journal of Geotechnical and Geoenvironmental Engineering**, ASCE, v. 131, n. 8, p. 1024–1033, 2005. Disponível em: <[https://doi.org/10.1061/\(ASCE\)1090-0241\(2005\)131:8\(1024\)](https://doi.org/10.1061/(ASCE)1090-0241(2005)131:8(1024))>.
- HEJAZI, S. M. et al. A simple review of soil reinforcement by using natural and synthetic fibers. **Construction and Building Materials**, Elsevier, v. 30, p. 100–116, 2012. Disponível em: <<https://doi.org/10.1016/j.conbuildmat.2011.11.045>>.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, Elsevier, v. 2, n. 5, p. 359–366, 1989. Disponível em: <[https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)>.
- HOWARD, J.; RUDER, S. Universal language model fine-tuning for text classification. In: **Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**. Melbourne, Australia: Association for Computational Linguistics, 2018. p. 328–339. Disponível em: <<https://doi.org/10.18653/v1/P18-1031>>.
- JAVADI, A. A.; REZANIA, M. Applications of artificial intelligence and data mining techniques in soil modeling. **Geomechanics and Engineering**, Techno Press, v. 1, n. 1, p. 53–74, 2009.
- JEFFERIES, M.; BEEN, K. **Soil liquefaction: a critical state approach**. 2. ed. Boca Raton, FL: CRC Press, 2015. Disponível em: <<https://doi.org/10.1201/b19114>>.
- JEFFERIES, M. G. Nor-Sand: a simple critical state model for sand. **Géotechnique**, ICE Publishing, v. 43, n. 1, p. 91–103, 1993. Disponível em: <<https://doi.org/btw5xs>>.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: **3rd International Conference on Learning Representations (ICLR)**. San Diego, CA: [s.n.], 2015. Disponível em: <<https://doi.org/10.48550/arXiv.1412.6980>>.
- KOHAVI, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)**. Montreal, Canada: Morgan Kaufmann, 1995. v. 2, p. 1137–1143.
- LECUN, Y. A. et al. Efficient BackProp. In: MONTAVON, G.; ORR, G. B.; MÜLLER, K.-R. (Ed.). **Neural Networks: Tricks of the Trade**. 2. ed. Berlin, Heidelberg:

Springer, 2012, (Lecture Notes in Computer Science, v. 7700). p. 9–48. Disponível em: <https://doi.org/10.1007/978-3-642-35289-8_3>.

LESHNO, M. et al. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. **Neural Networks**, Elsevier, v. 6, n. 6, p. 861–867, 1993. Disponível em: <[https://doi.org/10.1016/S0893-6080\(05\)80131-5](https://doi.org/10.1016/S0893-6080(05)80131-5)>.

LI, L.; ZORNBERG, J. G. Mobilization of reinforcement forces in fiber-reinforced soil. **Journal of Geotechnical and Geoenvironmental Engineering**, ASCE, v. 139, n. 1, p. 107–115, 2013. Disponível em: <<https://doi.org/gksmsf>>.

MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. Rectifier nonlinearities improve neural network acoustic models. In: **Proceedings of the 30th International Conference on Machine Learning (ICML), Workshop on Deep Learning for Audio, Speech and Language Processing**. Atlanta, GA: [s.n.], 2013. v. 28.

MACHADO, S. L.; CARVALHO, M. F.; VILAR, O. M. Constitutive model for municipal solid waste. **Journal of Geotechnical and Geoenvironmental Engineering**, ASCE, v. 128, n. 11, p. 940–951, 2002. Disponível em: <[https://doi.org/10.1061/\(ASCE\)1090-0241\(2002\)128:11\(940\)](https://doi.org/10.1061/(ASCE)1090-0241(2002)128:11(940))>.

MACHADO, S. L. et al. Modeling the mechanical behaviour of fibre reinforced sands. **Geomechanics and Geoengineering**, Taylor & Francis, v. 19, n. 5, p. 800–822, 2024. Disponível em: <<https://doi.org/10.1080/17486025.2024.2319619>>.

MASI, F. et al. Thermodynamics-based artificial neural networks for constitutive modeling. **Journal of the Mechanics and Physics of Solids**, Elsevier, v. 147, p. 104277, 2021. Disponível em: <<https://doi.org/10.1016/j.jmps.2020.104277>>.

MEYES, R. et al. **Ablation studies in artificial neural networks**. 2019. Disponível em: <<https://arxiv.org/abs/1901.08644>>. Disponível em: <<https://doi.org/10.48550/arXiv.1901.08644>>.

MICHALOWSKI, R. L. Limit analysis with anisotropic fibre-reinforced soil. **Géotechnique**, ICE Publishing, v. 58, n. 6, p. 489–501, 2008. Disponível em: <<https://doi.org/10.1680/geot.2007.00055>>.

MICHALOWSKI, R. L.; ČERMÁK, J. Triaxial compression of sand reinforced with fibers. **Journal of Geotechnical and Geoenvironmental Engineering**, ASCE, v. 129, n. 2, p. 125–136, 2003. Disponível em: <<https://doi.org/dprwp4>>.

MICHALOWSKI, R. L.; ZHAO, A. Failure of fiber-reinforced granular soils. **Journal of Geotechnical Engineering**, ASCE, v. 122, n. 3, p. 226–234, 1996. Disponível em: <[https://doi.org/10.1061/\(ASCE\)0733-9410\(1996\)122:3\(226\)](https://doi.org/10.1061/(ASCE)0733-9410(1996)122:3(226))>.

NAIR, V.; HINTON, G. E. Rectified linear units improve restricted Boltzmann machines. In: **Proceedings of the 27th International Conference on Machine Learning (ICML)**. Haifa, Israel: Omnipress, 2010. p. 807–814.

OWUSU-ANSAH, D. et al. Prediction of elastic modulus for fibre-reinforced soil-cement mixtures: a machine learning approach. **Applied Sciences**, v. 12, n. 17, p. 8540, 2022. Disponível em: <<https://doi.org/10.3390/app12178540>>.

- PAN, S. J.; YANG, Q. A survey on transfer learning. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 22, n. 10, p. 1345–1359, 2010. Disponível em: <<https://doi.org/10.1109/TKDE.2009.191>>.
- PASZKE, A. et al. PyTorch: An imperative style, high-performance deep learning library. In: **Advances in Neural Information Processing Systems**. [s.n.], 2019. p. 8026–8037. Disponível em: <<https://doi.org/10.48550/arXiv.1912.01703>>.
- PENUMADU, D.; ZHAO, R. Triaxial compression behavior of sand and gravel using artificial neural networks (ANN). **Computers and Geotechnics**, Elsevier, v. 24, n. 3, p. 207–230, 1999. Disponível em: <[https://doi.org/10.1016/S0266-352X\(99\)00002-6](https://doi.org/10.1016/S0266-352X(99)00002-6)>.
- PINTO, C. M. T. **Estudo do comportamento mecânico de uma areia aluvionar reforçada com fibras**. 83 p. Dissertação (Dissertação (Mestrado em Engenharia Civil) – Escola Politécnica) — Universidade Federal da Bahia, Salvador, 2021. Publicado como artigo em Pinto e Machado (2022).
- PINTO, C. M. T.; MACHADO, S. L. Comportamento mecânico de uma areia aluvionar reforçada com fibras poliméricas. In: **XX Congresso Brasileiro de Mecânica dos Solos e Engenharia Geotécnica**. Campinas, SP: ABMS, 2022. p. 7632–7639. Disponível em: <<https://doi.org/10.4322/cobramseg.2022.0960>>.
- POTTS, D. M.; ZDRAVKOVIC, L. **Finite element analysis in geotechnical engineering: theory**. London: Thomas Telford, 1999. Disponível em: <<https://doi.org/10.1680/feaiget.27534>>.
- ROSCOE, K. H.; BURLAND, J. B. On the generalized stress-strain behaviour of wet clay. In: HEYMAN, J.; LECKIE, F. A. (Ed.). **Engineering plasticity**. Cambridge: Cambridge University Press, 1968. p. 535–609.
- SANTAMARINA, J. C.; KLEIN, K. A.; FAM, M. A. **Soils and waves: particulate materials behavior, characterization and process monitoring**. Chichester: John Wiley & Sons, 2001.
- SCHANZ, T.; VERMEER, P. A.; BONNIER, P. G. The hardening soil model: formulation and verification. In: BRINKGREVE, R. B. J. (Ed.). **Beyond 2000 in computational geotechnics**. Rotterdam: Balkema, 1999. p. 281–296. Disponível em: <<https://doi.org/10.1201/9781315138206-27>>.
- SHAHIN, M. A.; JAKSA, M. B.; MAIER, H. R. Artificial neural network applications in geotechnical engineering. **Australian Geomechanics**, v. 36, n. 1, p. 49–62, 2001.
- STOFFEL, M.; BAMER, F.; MARKERT, B. Artificial neural networks and intelligent finite elements in non-linear structural mechanics. **Thin-Walled Structures**, Elsevier, v. 131, p. 102–106, 2018. Disponível em: <<https://doi.org/10.1016/j.tws.2018.06.035>>.
- WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research**, Inter-Research, v. 30, n. 1, p. 79–82, 2005. Disponível em: <<https://doi.org/10.3354/cr030079>>.
- WOOD, D. M. **Geotechnical modelling**. London: Spon Press, 2004. Disponível em: <<https://doi.org/10.1201/9781315273556>>.

YOSINSKI, J. et al. How transferable are features in deep neural networks? In: **Advances in Neural Information Processing Systems 27 (NIPS)**. Montreal, Canada: [s.n.], 2014. p. 3320–3328. Disponível em: <<https://doi.org/10.48550/arXiv.1411.1792>>.

YU, H. S. CASM: a unified state parameter model for clay and sand. **International Journal for Numerical and Analytical Methods in Geomechanics**, Wiley, v. 22, n. 8, p. 621–653, 1998. Disponível em: <<https://doi.org/d9thtg>>.

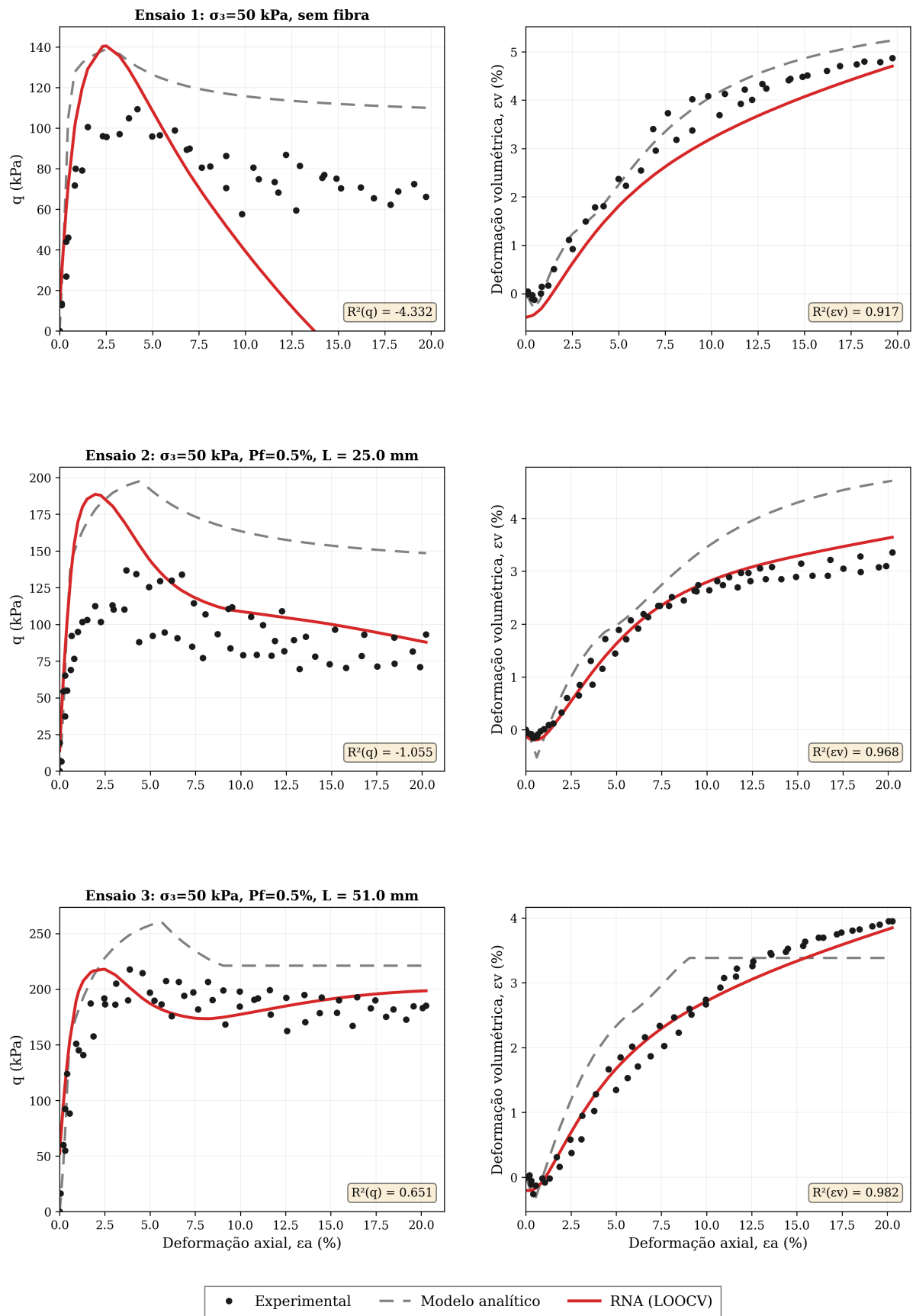
ZHANG, P. et al. Intelligent modelling of clay compressibility using hybrid meta-heuristic and machine learning algorithms. **Geoscience Frontiers**, Elsevier, v. 12, n. 1, p. 441–452, 2021. Disponível em: <<https://doi.org/10.1016/j.gsf.2020.02.014>>.

Apêndices

Curvas Tensão-Deformação de Todos os Ensaios

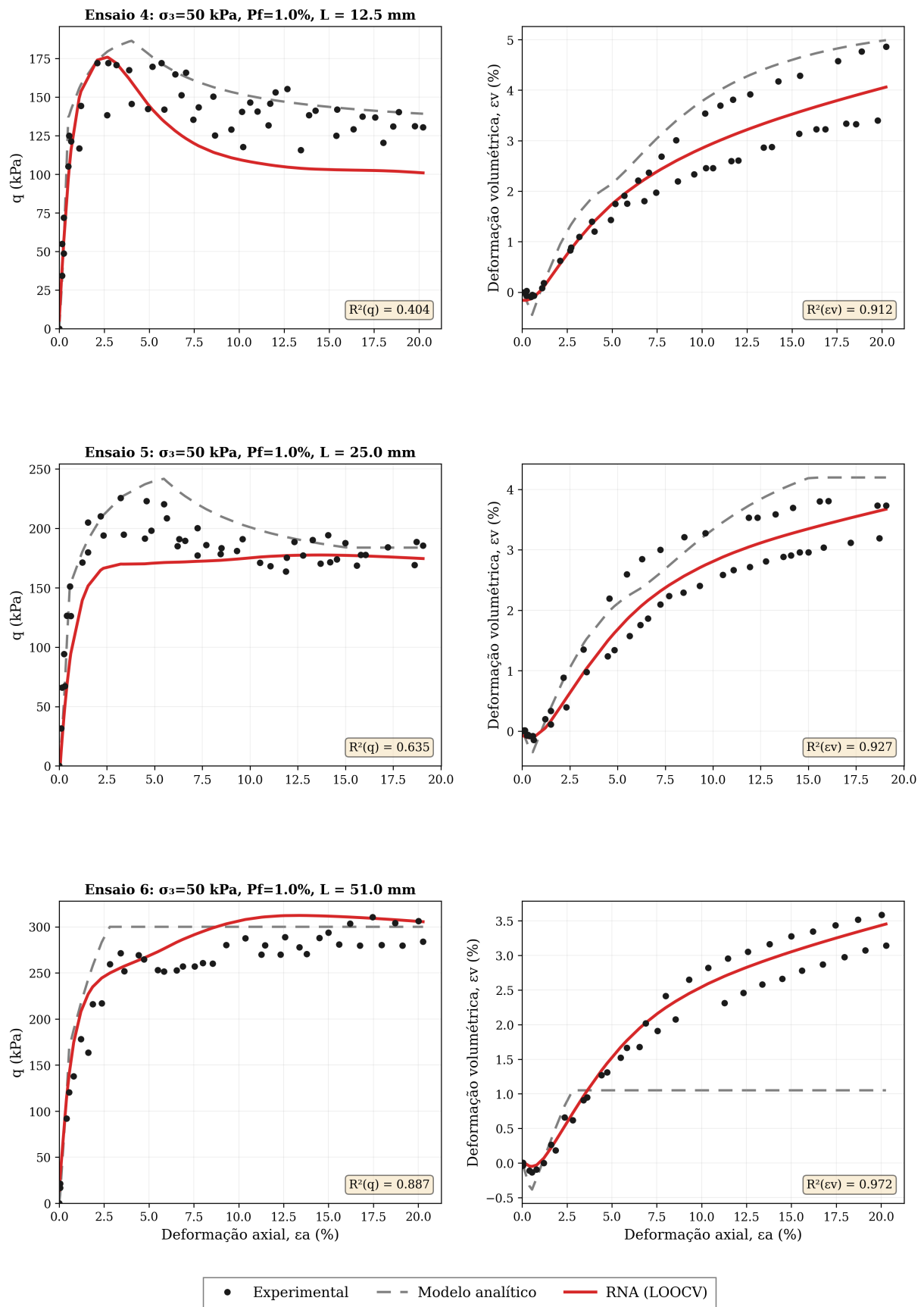
Reúnem-se neste apêndice as curvas previstas e observadas para os vinte e quatro ensaios da validação cruzada *leave-one-out*, na mesma ordem em que figuram na Tabela 7, três ensaios por figura. Em cada painel, os pontos representam os dados experimentais, a linha cheia representa a previsão da rede para o ensaio reservado da rodada correspondente e a linha tracejada representa a previsão do modelo analítico de referência para o mesmo cenário, discutido na Seção 4.6. O sinal de ε_v é exibido invertido em relação à convenção adotada na base de dados, de modo que valores positivos indicam dilatação, como na Seção 4.4.

Figura 16 – Ensaios 1 a 3



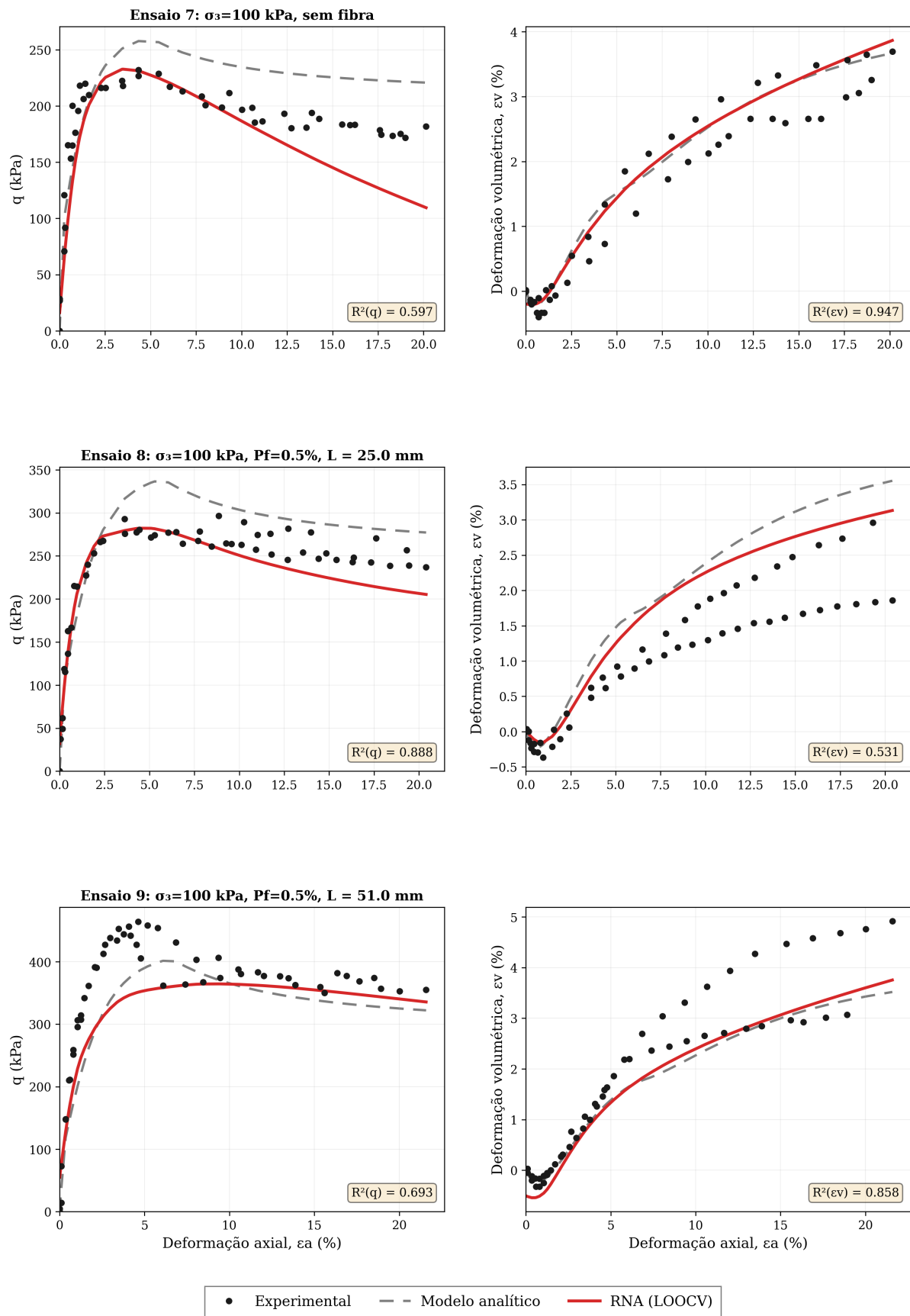
Fonte: Elaborado pelo autor.

Figura 17 – Ensaios 4 a 6



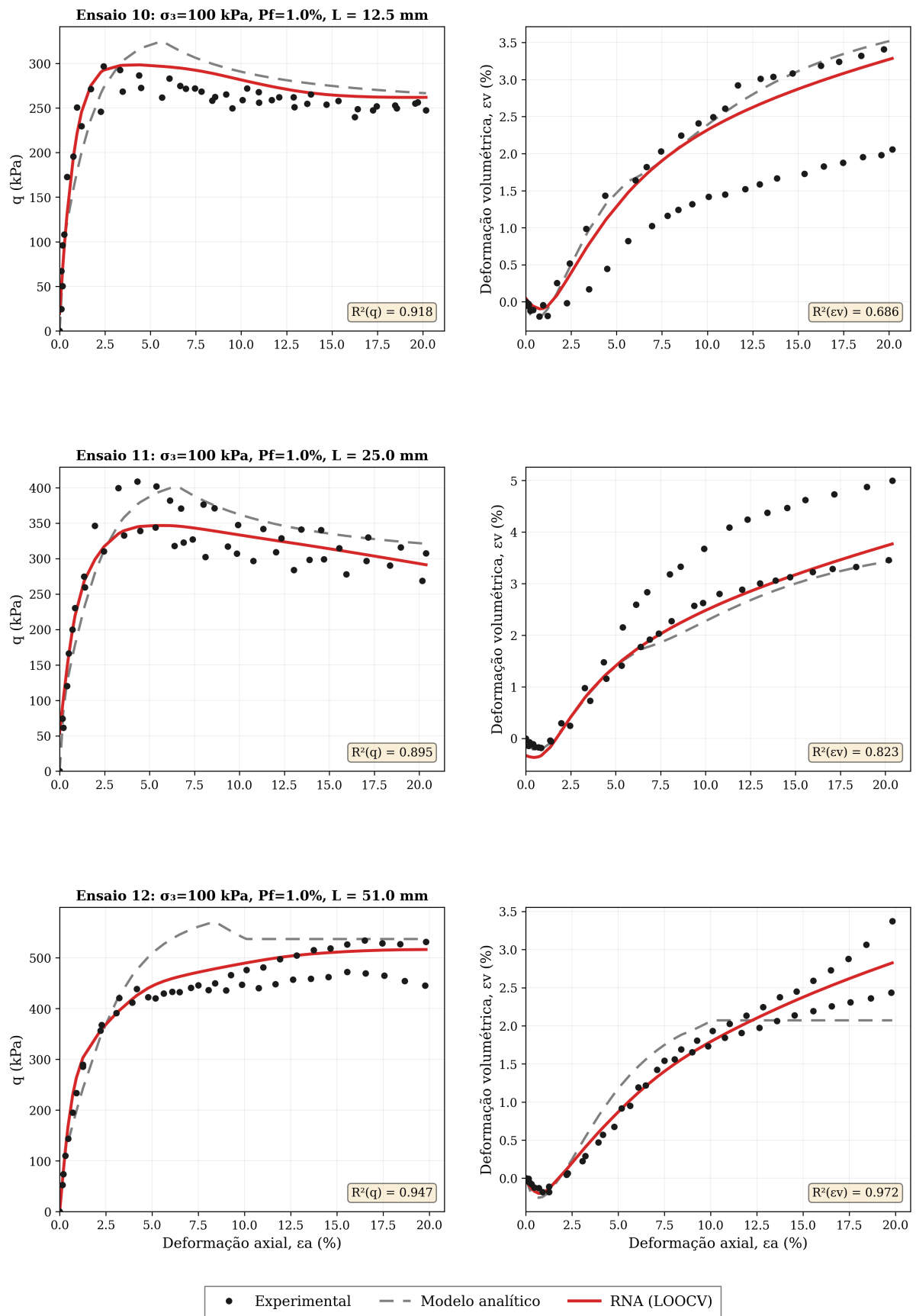
Fonte: Elaborado pelo autor.

Figura 18 – Ensaios 7 a 9



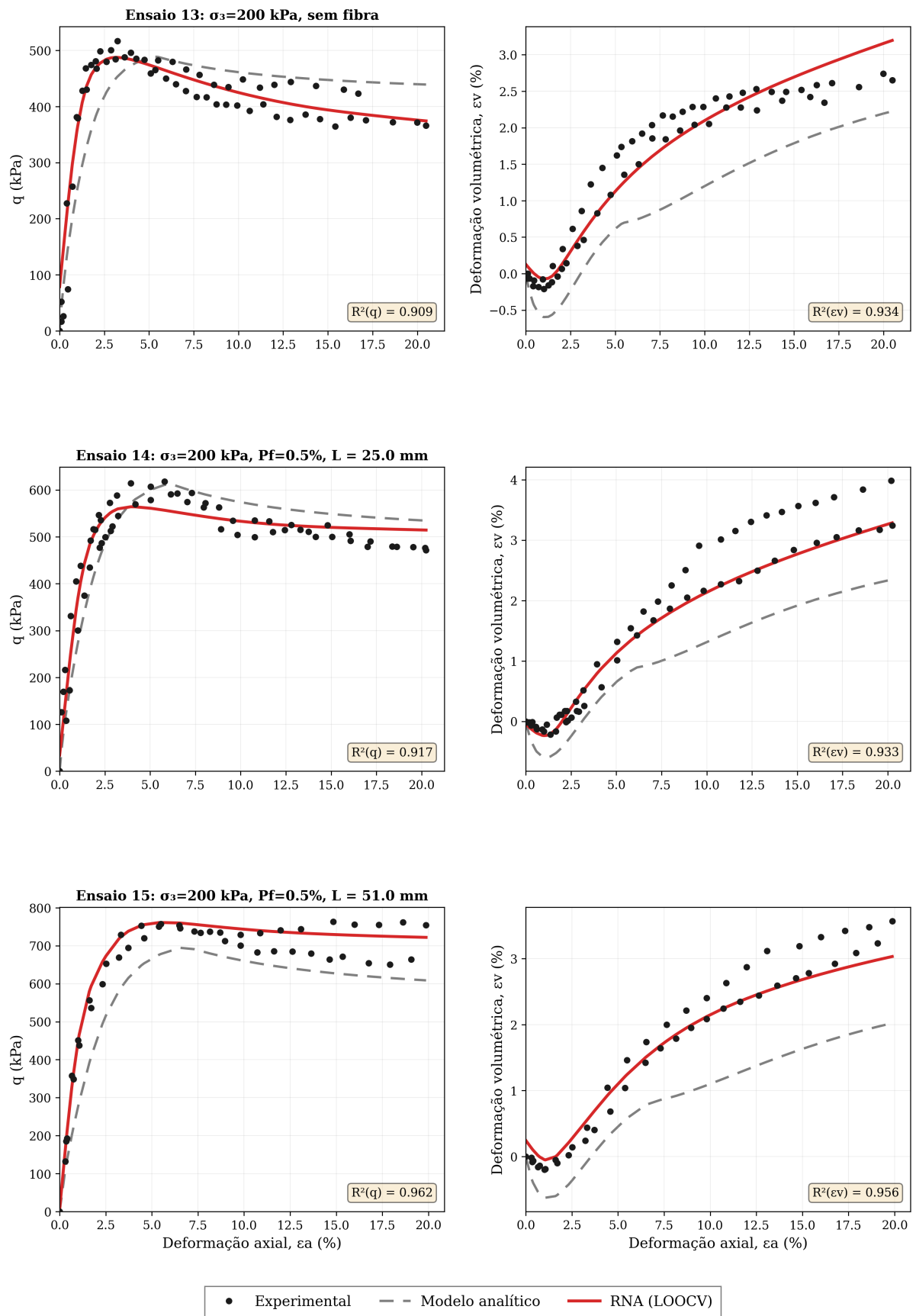
Fonte: Elaborado pelo autor.

Figura 19 – Ensaios 10 a 12



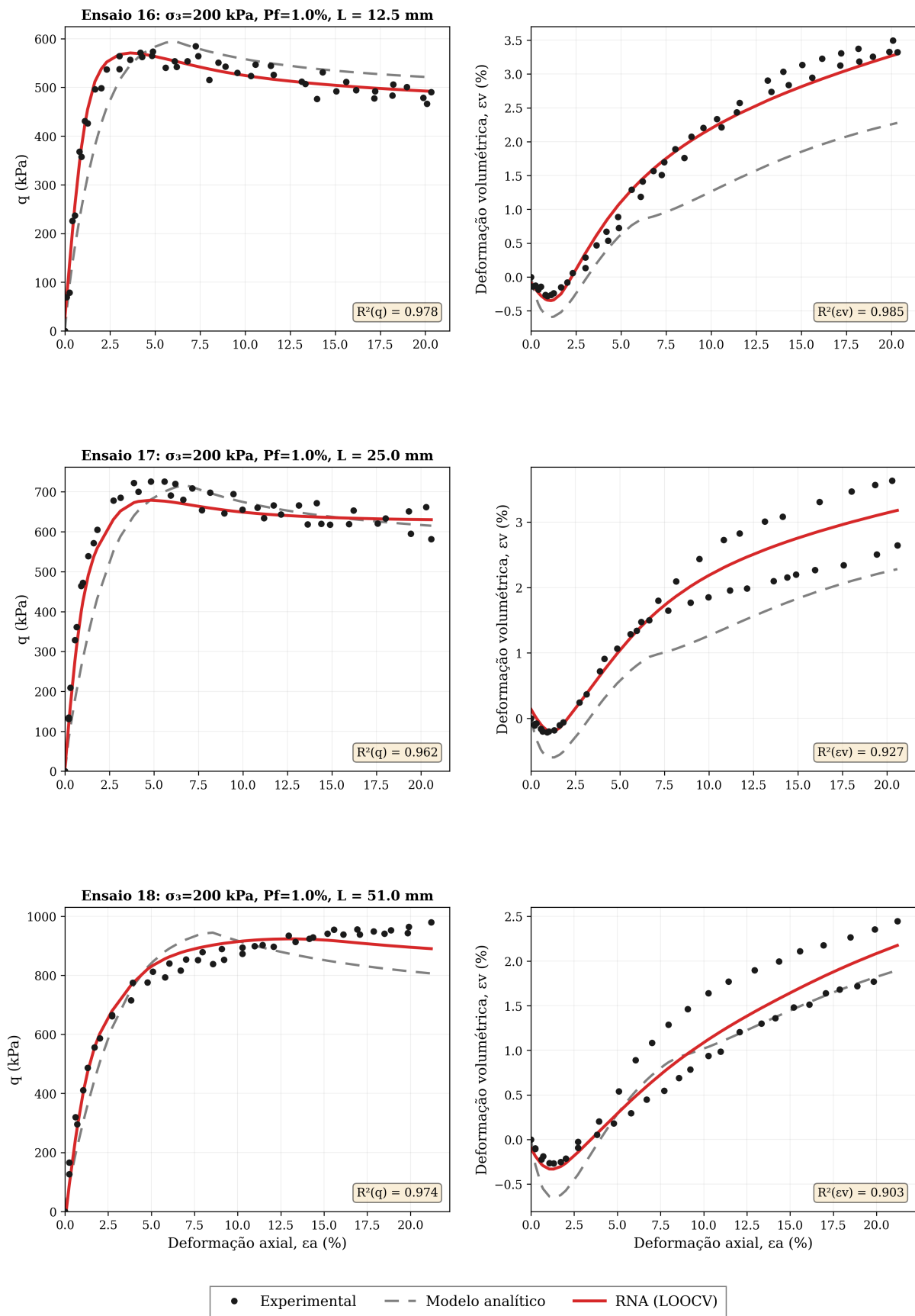
Fonte: Elaborado pelo autor.

Figura 20 – Ensaios 13 a 15



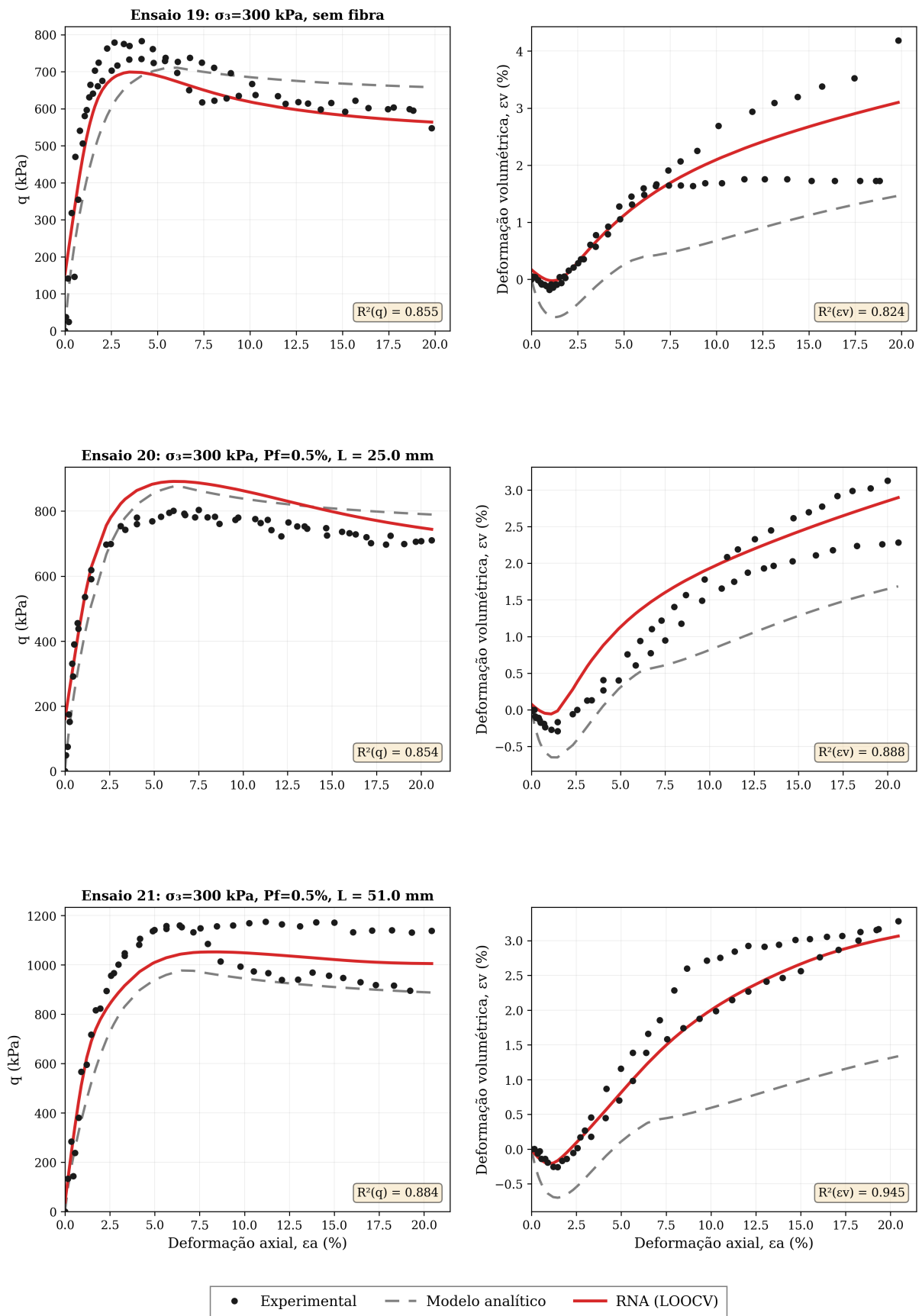
Fonte: Elaborado pelo autor.

Figura 21 – Ensaios 16 a 18



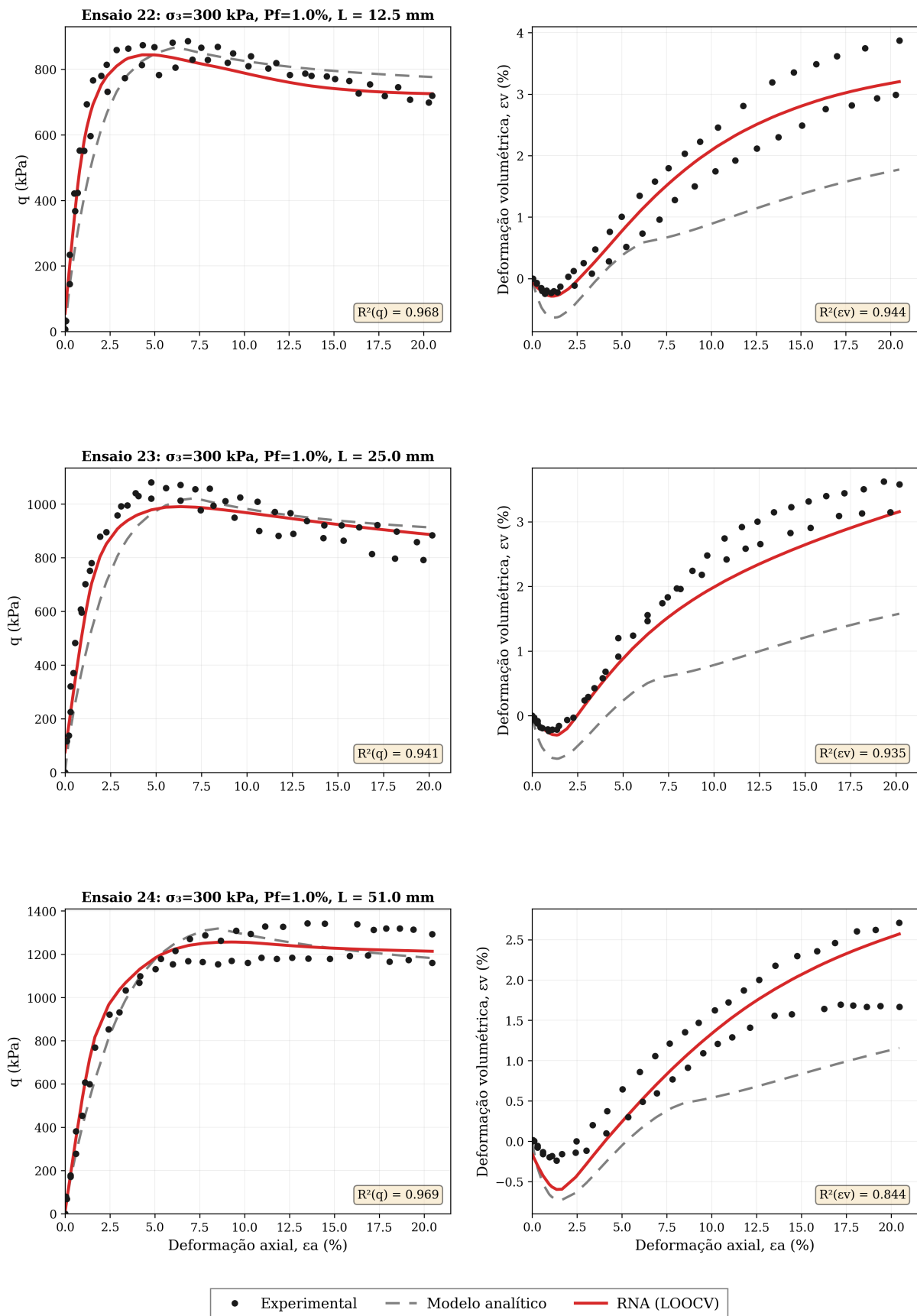
Fonte: Elaborado pelo autor.

Figura 22 – Ensaios 19 a 21



Fonte: Elaborado pelo autor.

Figura 23 – Ensaios 22 a 24



Fonte: Elaborado pelo autor.