

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

SANDILEIA RECALCATTI

**ESTIMATIVA DA QUANTIDADE DE RESÍDUOS DA CONSTRUÇÃO CIVIL
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

CURITIBA

2023

SANDILEIA RECALCATTI

**ESTIMATIVA DA QUANTIDADE DE RESÍDUOS DA CONSTRUÇÃO CIVIL
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

Estimating the amount of construction waste using artificial neural networks

Dissertação apresentada como requisito para obtenção do título de Mestre em Engenharia Civil da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. André Nagalli.

CURITIBA

2023



Esta licença permite remixe, adaptação e criação a partir do trabalho, para fins não comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Curitiba



SANDILEIA RECALCATTI

ESTIMATIVA DA QUANTIDADE DE RESÍDUOS DA CONSTRUÇÃO CIVIL UTILIZANDO REDES NEURAIS ARTIFICIAIS

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestra Em Engenharia Civil da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Construção Civil.

Data de aprovação: 05 de Abril de 2023

Dr. Andre Nagalli, Doutorado - Universidade Tecnológica Federal do Paraná

Dra. Andrea Parisi Kern, Doutorado - Universidade do Vale do Rio dos Sinos (Unisinos)

Dra. Karina Querne De Carvalho Passig, Doutorado - Universidade Tecnológica Federal do Paraná

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 05/04/2023.

Dedico este trabalho à minha família, em especial aos meus pais, Honorino e Natalina Recalcatti, pelo carinho, incentivo e apoio.

AGRADECIMENTOS

Agradeço à Deus, por ter me permitido chegar até aqui.

Ao meu orientador Prof. Dr. André Nagalli, por todo conhecimento técnico compartilhado, pela sua atenção, paciência e excelente orientação ao longo do desenvolvimento dessa pesquisa.

Aos membros da banca de avaliação, Prof.^a Dra. Karina Q. de Carvalho e Prof.^a Andrea P. Kern pelas contribuições e incentivos a essa pesquisa.

Agradeço aos demais professores do Programa de Pós-graduação em Engenharia Civil (PPGEC) da Universidade Tecnológica Federal do Paraná (UTFPR) pelo ensino de qualidade e pela excelente oportunidade de ter sido Representante Discente do programa nos dois anos de curso.

Sou grata pelo apoio prestado por minha família, que mesmo distante sempre se fez presente e acreditou junto comigo nesse sonho.

Registro também, o meu agradecimento a todas as pessoas e empresas que de alguma forma contribuíram para o desenvolvimento dessa pesquisa.

Por fim, agradeço à UTFPR e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Capes (código de financiamento 001), pelas bolsas de incentivo à pesquisa.

RESUMO

A geração de resíduos é um dos impactos negativos causados pela construção civil, e enfrentado pelo poder público e por empresas construtoras. Estimar a quantidade de resíduos da construção civil (RCC) é fundamental no processo de gerenciamento, porém, essa tarefa não é simples, pois há diversos materiais e processos construtivos que podem ser utilizados no mesmo empreendimento. Os construtores estão em busca de soluções rápidas na construção civil, porém muitos dos métodos de quantificação disponíveis não atendem às suas necessidades de maneira adequada. As redes neurais artificiais (RNA's) podem suprir essa necessidade, devido à sua capacidade de aprendizado e solução de sistemas não lineares. Nesse sentido, o objetivo geral desta pesquisa é investigar a utilização de RNA's para estimar a geração de RCC em obras de construção civil. A pesquisa adotou como método a simulação, utilizando as amostras A e B com 5.000 e 10.000 dados fictícios, respectivamente, ambas com áreas de construção total entre 75 m² e 125.050 m², e a amostra R com dados de 360 obras, com áreas entre 906 m² a 138.824 m². Os dados fictícios foram criados com base em taxas de geração disponíveis na literatura, e os dados reais foram obtidos através de contato com empresas de construção civil localizadas em Curitiba/PR, e abrangem obras construídas entre 2006 e 2021. Para realizar as simulações, foi utilizado o software MATLAB® versão R2022a. Diferentes configurações de redes neurais foram treinadas com as amostras A e B, e foi possível verificar que o melhor resultado preditivo foi no treinamento da amostra B, com a rede neural *feed-forward* com duas variáveis de entrada (classificação do resíduo e área total construída), dez neurônios na camada oculta, uma variável de saída (quantidade de resíduos) e três ciclos de treinamento com o algoritmo *Bayesian Regularization*, apresentando valores de R² igual a 1,0, MSE igual a 42,87 kg e MAPE igual a 0,00013%. Na validação do modelo com a amostra R, o resultado de R² igual a 0,83 indicou um bom desempenho da rede neural em explicar a variação dos dados de saída com base nos dados de entrada. Além disso, modelo proposto apresentou um MSE de 4.337,69 m³ e o resultado de MAPE apontou que o modelo de rede neural apresentado é capaz de estimar com precisão mais de 60% dos casos, bem como foi verificada uma ótima estimativa apresentada pelo modelo de RNA quando comparado a outros modelos da literatura. Para essa pesquisa, o modelo de rede neural que apresentou os melhores resultados de previsão, foi a rede neural *feed-forward* com dez neurônios na camada oculta e três ciclos de treinamento com o algoritmo *Bayesian Regularization*. Essa pesquisa traz uma contribuição importante para o setor da construção civil, colaborando com a quantificação de RCC de maneira ágil, além de contribuir com o gerenciamento dos resíduos dentro e fora do canteiro de obras, bem como contribui na conscientização dos profissionais da área e é útil para basear ações que minimizem a geração de resíduos.

Palavras-chave: gestão de resíduos; aprendizado de máquina; inteligência artificial; quantificação de resíduos da construção civil.

ABSTRACT

Waste generation is one of the negative impacts caused by civil construction, and faced by the public authorities and construction companies. Estimating the amount of civil construction waste (CCW) is fundamental in the management process, however, this task is not simple, because there are several materials and construction processes that can be used in the same enterprise. Builders are in search of quick solutions in civil construction, but many of the available quantification methods do not meet their needs adequately. Artificial neural networks (ANNs) can meet this need, due to their ability to learn and solve non-linear systems. In this sense, the general objective of this research is to investigate the use of ANN's to estimate the generation of CCW in civil construction. The research adopted the simulation method, using samples A and B with 5,000 and 10,000 dummy data, respectively, both with total construction areas between 75 m² and 125,050 m², and sample R with data from 360 construction sites, with areas between 906 m² and 138,824 m². The dummy data was created based on generation rates available in the literature, and the real data was obtained through contact with construction companies located in Curitiba/PR, and covers buildings built between 2006 and 2021. To perform the simulations, the software MATLAB® version R2022a was used. Different configurations of neural networks were trained with samples A and B, and it was possible to verify that the best predictive result was in the training of sample B, with the feed-forward neural network with two input variables (waste classification and total built-up area), ten neurons in the hidden layer, one output variable (amount of waste) and three training cycles with the Bayesian Regularization algorithm, presenting R² values equal to 1.0, MSE equal to 42.87 kg and MAPE equal to 0.00013%. In the validation of the model with the R sample, the result of R² equal to 0.83 indicated a good performance of the neural network in explaining the variation of the output data based on the input data. In addition, the proposed model presented an MSE of 4,337.69 m³ and the MAPE result pointed out that the presented neural network model is able to accurately estimate more than 60% of the cases, as well as an optimal estimation presented by the ANN model when compared to other models in the literature. For this research, the neural network model that presented the best prediction results was the feed-forward neural network with ten neurons in the hidden layer and three training cycles with the Bayesian Regularization algorithm. This research brings an important contribution to the civil construction sector, collaborating with the quantification of CCW in an agile manner, besides contributing to the management of waste inside and outside the construction site, as well as contributing to the awareness of professionals in the area and is useful to base actions that minimize waste generation.

Keywords: waste management; machine learning; artificial intelligence; quantification of construction waste.

LISTA DE FIGURAS

Figura 1 - Características das publicações analisadas	23
Figura 2 - Número de publicações por ano	24
Figura 3 - Número de documentos por autor.....	24
Figura 4 - Número de documentos por instituição.....	25
Figura 5 - Número de documentos por país	25
Figura 6 - Distribuição dos documentos por área temática	26
Figura 7 - Acoplamento bibliográfico por autores	26
Figura 8 - Mapa de co-citações por referências	27
Figura 9 - Termos presentes no título e <i>abstract</i> das publicações.....	28
Figura 10 - Etapas da estratégia de pesquisa.....	46
Figura 11- Configurações e sequências das simulações	51
Figura 12 - Gráficos de Regressão Linear para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra A	60
Figura 13 - Gráfico do Erro Quadrático Médio para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra A	61
Figura 14 - Histograma de Erro para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra A	62
Figura 15 - Gráfico de Regressão Linear para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra B.....	69
Figura 16 - Gráfico do Erro Quadrático Médio para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra B	70
Figura 17 - Histograma de Erro para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra B	71
Figura 18 - Gráfico de Regressão Linear para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra R.....	74
Figura 19 - Gráfico do Erro Quadrático Médio para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo <i>Bayesian Regularization</i> - Amostra R	76
Figura 20 – Histograma de Erro da rede neural com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento <i>Bayesian Regularization</i> – Amostra R.....	78

LISTA DE QUADROS

Quadro 1 - Área da pesquisa, assunto e tema.....	22
Quadro 2 - Composição dos termos e <i>string</i> de busca	22
Quadro 3 - Tipo de sistema produtivo	36
Quadro 4 - Tipos de Redes Neurais.....	39
Quadro 5 - Critérios e classificação da pesquisa.....	45
Quadro 6 - Planejamento das simulações.....	50

LISTA DE GRÁFICOS

Gráfico 1 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Levenberg-Marquardt</i> - Amostra A	54
Gráfico 2 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Levenberg-Marquardt</i> - Amostra A.....	55
Gráfico 3 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Scaled Conjugate Gradient</i> - Amostra A	56
Gráfico 4 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Scaled Conjugate Gradient</i> - Amostra A	57
Gráfico 5 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Bayesian-Regularization</i> - Amostra A	58
Gráfico 6 - Quantidades de resíduos estimadas pela rede neural <i>feed-forward</i> com dez neurônios na camada oculta, um ciclo de treinamento e algoritmo de treinamento <i>Bayesian Regularization</i> - Amostra A	59
Gráfico 7 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Levenberg-Marquardt</i> - Amostra B	64
Gráfico 8 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Levenberg-Marquardt</i> - Amostra B.....	65
Gráfico 9 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Scaled Conjugate Gradient</i> - Amostra B	66
Gráfico 10 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando <i>Bayesian Regularization</i> - Amostra B	67
Gráfico 11 - Quantidades de resíduos estimadas pela rede neural <i>feed-forward</i> com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento <i>Bayesian Regularization</i> - Amostra B	68
Gráfico 12 - Gráfico <i>boxplot</i> da Amostra R.....	72
Gráfico 13 - MAPE da Amostra R separado por faixas de dados.....	79
Gráfico 14 - Percentual de erro absoluto da Amostra R	80
Gráfico 15 - Quantidades de resíduos reais e quantidades de resíduos estimadas pela rede neural <i>feed-forward</i> com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento <i>Bayesian Regularization</i>	81

LISTA DE ABREVIATURAS E SIGLAS

ABRELPE	Associação Brasileira de Emp. de Limp. Pública e Resíduos Especiais
ANFIS	<i>Adaptative Inference Neuro Fuzzy System</i>
BIM	<i>Building Information Modeling</i>
CONAMA	Conselho Nacional do Meio Ambiente
LER	Lista Europeia de Resíduos
MAPE	<i>Mean Absolut Percentual Error</i>
ML	<i>Machine Learning</i>
MSE	<i>Mean Squared Error</i>
ODS	Objetivos de Desenvolvimento Sustentável
PNRS	Política Nacional de Resíduos Sólidos
PPGEC	Programa de Pós Graduação em Engenharia Civil
RCC	Resíduos da Construção Civil
RMSE	<i>Root Mean Squared Error</i>
RNA	Rede Neural Artificial
RSU	Resíduos Sólidos Urbanos
SINIR	Sistema Nacional de Informações sobre a Gestão dos Res. Sólidos
SVM	<i>Support Vector Machine</i>
UTFPR	Universidade Tecnológica Federal do Paraná
WGA	<i>Waste Generation per groos floor Area</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Contextualização	15
1.2	Problema de Pesquisa	16
1.3	Objetivos	17
1.3.1	Objetivo Geral.....	17
1.3.2	Objetivos Específicos	17
1.4	Justificativa	17
1.5	Questões Orientadoras	19
1.6	Delimitação da Pesquisa	19
1.7	Estrutura da Dissertação	19
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	Análise Bibliométrica	21
2.1.1	Busca na Base Bibliográfica.....	21
2.1.2	Análise dos Resultados da Pesquisa Bibliográfica	23
2.2	Referencial Teórico	29
2.2.1	Resíduos da Construção Civil	29
2.2.2	Métodos de Quantificação de RCC	31
2.2.3	<i>Machine Learning</i> e Redes Neurais Artificiais	38
2.3	Revisão de Literatura	42
2.3.1	Estudos Utilizando Redes Neurais Artificiais	42
3	ESTRATÉGIA DE PESQUISA	45
3.1	Classificação da Pesquisa	45
3.2	Etapas da Pesquisa	46
3.2.1	Planejamento da Simulação	47
3.2.2	Execução da Simulação	50
3.2.3	Análise e Validação do Modelo Computacional	52
4	RESULTADOS E DISCUSSÕES	54
4.1	Simulação de Redes Neurais com a Amostra A	54
4.2	Simulação de Redes Neurais com a Amostra B	63
4.3	Validação do Modelo de Rede Neural com a Amostra R	72
5	CONCLUSÃO	86
	REFERÊNCIAS	89
	APÊNDICE A - Dados da Amostra R	97

APÊNDICE B - Passo a passo para realizar simulações com Redes Neurais Artificiais no MATLAB®.....	107
--	------------

1 INTRODUÇÃO

A indústria da construção civil causa impactos ambientais em toda a sua cadeia produtiva, da extração ao consumo. Esses impactos negativos são verificados em inúmeras etapas da obra que geram resíduos sólidos, desde a execução de novas construções, até reparos, reformas e demolições (GÁLVEZ-MARTOS et al., 2018; OPOKU, 2019).

A geração de resíduos não é apenas um problema ambiental, mas também econômico, uma vez que o desperdício aumenta o custo do construtor e do poder público quando responsável pela coleta. Um dos principais desafios do processo de gerenciamento dos RCC, por parte dos geradores, está relacionado a minimizar a geração de resíduos. Por outro lado, para a gestão municipal, a maior dificuldade é enfrentar seu descarte incorreto.

Zhang et al. (2019) apontaram que a China gerou 4,1 bilhões de toneladas métricas de RCC em 2016, sendo que deste montante aproximadamente 85% foram resíduos de demolição. Na União Europeia, as atividades de construção e demolição geram cerca de 850 milhões de toneladas de RCC por ano (VILLORIA SÁEZ; OSMANI, 2019). Nos Estados Unidos, foram geradas 569 milhões de toneladas de RCC em 2017, mais do que o dobro da quantidade de Resíduos Sólidos Urbanos - RSU gerados pelo Brasil, de acordo com a Agência de Proteção Ambiental (EPA, 2020).

No Brasil, dados do Sistema Nacional de Informações sobre a Gestão dos Resíduos Sólidos - SINIR, indicam que são geradas mais de 100 milhões de toneladas por ano de RCC, sendo que o país possui taxa de reciclagem menor do que 10% (SINIR, 2020). O Panorama dos Resíduos Sólidos no Brasil 2018/2019, elaborado pela Associação Brasileira de Empresas de Limpeza Pública e Resíduos Especiais - ABRELPE, aponta que os municípios coletaram cerca de 45 milhões de toneladas de RCC, descartados, em sua maioria, em vias e logradouros públicos em 2018, representando 36% do total de Resíduos Sólidos Urbanos - RSU produzidos no país (ABRELPE, 2019).

Quantificar resíduos não é algo simples, pois há incontáveis materiais e processos construtivos que podem ser utilizados no mesmo empreendimento. Dessa forma, predizer características, tipos e quantidades de resíduos se torna difícil, já que a geração dos RCC engloba variáveis como, processo construtivo, tamanho e

experiência da equipe, dentre outros aspectos (NAGALLI, 2021b). Nesse âmbito, estimar a quantidade de resíduos que será gerado em uma obra é de suma importância para seu gerenciamento, pois é possível determinar as quantidades que podem ser reutilizadas, recicladas ou em últimos casos, dispostas em aterros sanitários (GUERRA et al., 2019).

Considerando que o Brasil não dispõe de informações suficientes acerca dos resíduos e características das construções para possibilitar a estruturação de um modelo de predição de resíduos da construção civil, é conveniente contar com métodos de predição que correlacionem as características dos empreendimentos com a geração de resíduos. Uma das maneiras de estudar tais correlações é a partir de Redes Neurais Artificiais (RNA's), apesar de existirem poucos estudos desse tipo na literatura (NAGALLI, 2021a).

Redes neurais artificiais são sistemas de processamento de informações celulares projetadas e desenvolvidas com base na noção percebida do cérebro humano e seu sistema neural. Uma das características mais significativas das RNA's em modelos preditivos é sua capacidade de aprendizado. A RNA pode construir um sistema não linear complexo, a partir de um conjunto de dados de entrada/saída. Em virtude disso, a estrutura não linear da geração de resíduos faz com que a utilização de redes neurais artificiais seja ideal para cálculos de predição (FIRAT; TURAN; YURDUSEV, 2010).

Embora os métodos que utilizam regressões lineares múltiplas tenham descrito com sucesso a geração de resíduos para as amostras utilizadas nos estudos, os erros associados ao extrapolá-los são significativos. As redes neurais foram capazes de prever os RCC dos trabalhos estudados com qualidade superior ($R^2 > 0,94$) quando comparados aos modelos de regressão convencionais disponíveis na literatura. Portanto, é preferível trabalhar com modelos de aprendizado de máquina que permitam manipular grandes quantidades de dados (NAGALLI, 2021a).

Na literatura, não é comum encontrar pesquisas que tratam sobre estimativas de quantidades de resíduos de construção utilizando redes neurais artificiais. Assim sendo, essa pesquisa busca investigar a utilização de RNA's na estimativa de resíduos da construção civil, proporcionando uma ferramenta de tomada de decisões aos gestores que possibilite estimar quantidades geradas de RCC.

1.1 Contextualização

A previsão de taxas confiáveis de geração de resíduos é um elemento fundamental no planejamento e implementação de estratégias sustentáveis de gestão de resíduos sólidos (COSKUNER et al., 2021).

Existem vários métodos de quantificação de resíduos de construção e demolição presentes na literatura, como por exemplo, visita ao local, cálculos de taxa de geração, acumulação de sistema de classificação e análise de ciclo de vida das construções. No entanto, esses métodos possuem limitações, incluindo implementação demorada, falta de precisão e de dados confiáveis, generalizações excessivas e ausência de meios de verificação que provem a adequação do método adotado (GUERRA et al., 2019).

Li et al. (2013) apresentaram um modelo de quantificação de resíduos por área bruta (*Waste Generation per gross floor Area - WGA*), com base no princípio do balanço de massa para a construção civil na China. Os autores relacionaram a quantidade de material necessária para construção de um edifício com a quantidade de material comprado, e puderam concluir que o método é adequado e prático.

Nos últimos anos, modelos de inteligência artificial e aprendizado de máquina, como Máquina de Vetor de Suporte (*Support Vector Machine - SVM*), Sistema de Inferência *Neuro-fuzzy* Adaptativo (*Adaptative neuro-fuzzy inference system - ANFIS*) e Rede Neural Artificial (*Artificial Neural Network - ANN*) têm ganhado destaque devido à sua elevada flexibilidade e habilidades de predição comprovadas. Modelos inteligentes são capazes de prever a geração de resíduos a curto, médio e longo prazo (ABBASI; EL HANANDEH, 2016).

Abbasi e El Hanandeh (2016) identificaram cinco categorias para previsão de RSU, sendo elas, métodos estatísticos descritivos, análise de regressão, modelo de fluxo de material, análise de série temporal e modelos de inteligência artificial. A comparação entre as técnicas de regressão linear múltipla e redes neurais artificiais mostrou que os resultados obtidos com as RNA's são mais precisos e estatisticamente significativos.

No âmbito computacional, a modelagem de informações de construção (*Building Information Modeling - BIM*) também vem sendo utilizada em conjunto com as RNA's. Akinade et al. (2018) utilizaram dados de 117 edifícios e destacaram o

potencial do uso do BIM em conduzir um processo eficaz para previsão de resíduos em projeto, possibilitando melhor tomada de decisões à equipe.

Nesse sentido, o método de estimativa da geração de resíduos utilizando aprendizado de máquina se mostra como mais preciso e de rápida implementação, quando comparado a métodos que dependem de visitas *in loco* e levantamentos quantitativos de materiais, por exemplo.

1.2 Problema de Pesquisa

O crescimento das atividades de construção nos últimos anos resultou na geração de grandes quantidades de resíduos. Estimar a quantidade de RCC gerado é um passo fundamental no processo de gerenciamento. Através da quantificação é que se forma toda a logística da obra, uma vez que se pode definir, por exemplo, a área para armazenamento dos resíduos, frequência de coleta e melhor forma de transporte interno e externo (NAGALLI, 2022).

Atualmente, a busca por soluções rápidas na construção civil vem crescendo gradativamente e, os métodos de quantificação disponíveis não são suficientemente adequados para os construtores utilizarem sem exigir muito tempo e esforço. As RNA's podem suprir essa necessidade, devido à sua capacidade de aprendizado, podendo construir um sistema não linear complexo por meio de um conjunto de exemplos de entrada e saída, sendo empregada com sucesso na modelagem de sistemas não lineares (FIRAT; TURAN; YURDUSEV, 2010).

Redes Neurais Artificiais são utilizadas para estimar ou aproximar funções que podem depender de grande número de entradas que geralmente são desconhecidas (LU et al., 2016). O modelo de RNA's é uma ferramenta poderosa que fornece soluções viáveis para problemas complexos. Além disso, a estrutura não linear da geração de RCC torna a RNA ideal para a previsão da geração de resíduos (NAGALLI, 2021a).

Considerando a relevância do tema, o problema desta pesquisa é verificar se Redes Neurais Artificiais podem ser consideradas instrumentos para prever a geração de Resíduos da Construção Civil.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral deste trabalho foi investigar a utilização de Redes Neurais Artificiais na estimativa da quantidade de Resíduos da Construção Civil.

1.3.2 Objetivos Específicos

São objetivos específicos estabelecidos:

- Definir bibliograficamente quais variáveis a serem consideradas no modelo;
- Simular diferentes configurações de Redes Neurais Artificiais para estimativa dos RCC;
- Definir configuração de rede neural mais adequada e avaliar seu desempenho;
- Utilizar uma base de dados reais para validar o modelo desenvolvido e;
- Analisar a sensibilidade do modelo computacional quanto à natureza e à quantidade de dados.

1.4 Justificativa

O consumo de recursos naturais em atividades de construção civil e os impactos da disposição de RCC em aterros, são exemplos de como a indústria da construção contribui com a degradação do ambiente. Parte desses impactos podem ser mitigados adotando-se o princípio dos 3R's, reduzir, reutilizar e reciclar. Ter uma estimativa, a nível de projeto, da quantidade de resíduos gerados, é de suma importância para seu gerenciamento, pois permite identificar o que pode ser reutilizado e reciclado na própria obra. Por esse motivo, é necessário um método eficiente, que utilize dados confiáveis de projeto para obter uma estimativa de RCC mais precisa e direta (GUERRA et al., 2019).

Lu et al. (2021) afirmam que o que não pode ser medido, não pode ser melhorado e que informações sobre a geração de resíduos são um pré-requisito para muitas estratégias de gestão e atividades de engajamento público, incluindo planejamento da área do aterro, determinação de taxas de poluidores ou subsídios para recicladores e empresas que implementam políticas de gestão de resíduos.

Sistemas avançados de previsão têm mostrado superioridade aos modelos convencionais em problemas de engenharia, bem como em pesquisa de gestão de resíduos. Uma vez que, modelos teóricos relacionam resíduos e características de construção, mas as práticas locais podem resultar em comportamentos diferentes (ALI ABDOLI et al., 2012; TEIXEIRA et al., 2020).

A visualização, ainda na fase de projeto, das perdas e desperdícios gerados na construção civil, permite aos projetistas averiguarem medidas corretivas e identificarem processos críticos de geração de resíduos, planejando assim, estratégias de controle para maior eficiência da construção, o que reduz custos relacionados à mobilização nos canteiros de obras, desperdício de materiais e transporte e descarte dos resíduos. Ainda, o modelo pode servir como subsídio para construtoras, prefeituras e órgãos fiscalizadores realizarem um gerenciamento efetivo dos resíduos, que conseqüentemente irá promover redução na geração e descarte ambientalmente incorreto, reduzindo os impactos ambientais causados pelos RCC.

Essa pesquisa contribui com as ações dos Objetivos de Desenvolvimento Sustentável (ODS) para os quais as Nações Unidas estão trabalhando a fim de se possa atingir a Agenda 2030 no Brasil (IPEA, 2018). Principalmente, o ODS 6: Disponibilidade e gestão sustentável da água e saneamento para todas e todos, ODS 11: Tornar as cidades e os assentamentos humanos inclusivos, seguros, resilientes e sustentáveis bem como, o ODS 12: Assegurar padrões de produção e de consumo sustentáveis. Esses três objetivos possuem metas para reduzir os impactos ambientais negativos sobre a saúde humana e meio ambiente, causados pela geração e disposição irregular de resíduos sólidos, além de implementar a gestão de resíduos municipais e planos de gerenciamento de resíduos sólidos.

Além da sua contribuição auxiliando o trabalho de diversos profissionais, como projetistas e formuladores de políticas públicas, o tema desta pesquisa está diretamente ligado ao grupo de pesquisa Gestão de Resíduos da Construção Civil, do Programa de Pós-Graduação em Engenharia Civil - PPGEC da Universidade Tecnológica Federal do Paraná - UTFPR, que já desenvolveu diversos trabalhos nessa área. Podem ser citadas as pesquisas desenvolvidas por Biju (2015), que utilizou o Sistema de Informação Geográfica (SIG) na indicação de possíveis áreas aptas à disposição de resíduos da construção; Schamne (2016) que avaliou os preceitos da aplicação da logística reversa no setor da construção; Silva (2020) que

investigou o potencial de recuperação de materiais e componentes de edificações; e Geraldo Filho (2021) que avaliou o gerenciamento de resíduos sólidos em uma obra portuária no município de Paranaguá/PR.

1.5 Questões Orientadoras

A presente pesquisa busca responder às seguintes questões orientadoras:

- Quais características das obras de construção civil podem ser relacionadas com os dados de geração de resíduos obtidos?
- É possível incorporar um modelo de predição na base de dados existente?
- Qual configuração de RNA apresenta resultados eficazes na predição da geração de resíduos de construção e demolição?

1.6 Delimitação da Pesquisa

Para a elaboração dessa pesquisa, foram utilizados dados fictícios de obras de construção e demolição, e dados reais de obras de construção civil.

Os dados das amostras fictícias trataram-se de área bruta construída e/ou demolida (m^2) e quantidades de resíduos gerados (kg). As áreas variaram de $75 m^2$ a $125.050 m^2$ e as quantidades de resíduos variaram de 11.250 a 100.040.000 kg.

Na amostra real, os dados consistiram em área total construída (m^2) e quantidades de resíduos de construção gerados (m^3). As áreas variaram de $906 m^2$ a $138.824 m^2$ e as quantidades de resíduos variaram de $572 m^3$ a $98.280 m^3$.

O estudo foi desenvolvido na cidade de Curitiba/PR durante o ano de 2022, e os dados das obras reais compreendem edificações construídas entre 2006 e 2021. No entanto, devido à parte dos dados que não são de obras reais, não houve delimitação espacial para a pesquisa, que pode ser aplicada em demais localidades.

1.7 Estrutura da Dissertação

Esta pesquisa foi estruturada em cinco capítulos, sendo Introdução, Fundamentação Teórica, Estratégia de Pesquisa, Resultados e Discussões e Conclusões.

Na Introdução apresenta-se o assunto a ser abordado, a contextualização acerca do tema, problema de pesquisa, objetivos, justificativa, questões orientadoras e delimitação da pesquisa, bem como a estrutura do trabalho.

No segundo capítulo é apresentada a Fundamentação Teórica, com análise bibliométrica do tema, referencial teórico contendo os principais conceitos abordados e revisão de literatura, que evidencia as principais pesquisas já realizadas e autores relacionados com o tema.

Em seguida, o capítulo três discorre sobre a Estratégia de Pesquisa, onde é apresentada a classificação da pesquisa e as etapas adotadas para sua realização, bem como os critérios e procedimentos empregados para coleta e análise dos dados.

O quarto capítulo trata dos Resultados e Discussões, onde apresentam-se as análises realizadas, assim como as discussões sobre os resultados obtidos.

Por fim, o quinto capítulo aponta a Conclusão sobre os dados analisados, tais como limitações encontradas e recomendações para trabalhos futuros. Além disso, são apresentadas as referências e apêndices utilizados neste trabalho.

2 FUNDAMENTAÇÃO TEÓRICA

Com o intuito de fundamentar a pesquisa, neste capítulo é abordada a revisão bibliográfica dividida em três subcapítulos. Primeiramente será apresentada uma análise geral do tema baseada em bibliometria, com a finalidade de identificar as principais publicações, bem como os principais autores pertinentes à área. Na sequência será apresentado o referencial teórico, com os principais temas e conceitos relacionados à pesquisa, tais como: resíduos sólidos, resíduos de construção e demolição, métodos de quantificação de RCC e redes neurais artificiais. Por fim, serão apresentados os principais estudos análogos aos objetivos desta dissertação.

2.1 Análise Bibliométrica

A bibliometria é utilizada como um método de análise quantitativa para pesquisa científica. Os dados estatísticos provenientes dos estudos bibliométricos determinam a contribuição do conhecimento científico resultante das publicações em determinadas áreas. Esses dados podem ser utilizados na representação das tendências de pesquisa, assim como na identificação de novos temas de pesquisa (SU; LEE, 2010).

A produção acadêmica é influenciada por inúmeros fenômenos sociais e naturais, e a análise bibliométrica auxilia no entendimento de tais fenômenos, visto que é uma ferramenta amplamente difundida e fundamental na avaliação da produção científica (LIU et al., 2014).

2.1.1 Busca na Base Bibliográfica

Para realização da pesquisa foi utilizada a base de dados *Scopus*, que é o maior banco de dados de resumos e citações da literatura. Oferecendo uma visão geral e abrangente da produção mundial de pesquisas em inúmeras áreas, como ciências, tecnologia, medicina, ciências sociais, e artes e humanidades. O *Scopus* apresenta ferramentas inteligentes para rastrear, analisar e visualizar pesquisas (ELSEVIER, 2021).

O primeiro passo para realização da pesquisa foi a definição da área de pesquisa, assunto e tema a ser abordado. No Quadro 1 são apresentadas informações da definição da área, assunto e tema abordado da pesquisa.

Quadro 1 - Área da pesquisa, assunto e tema

Área de pesquisa	Gestão e Sustentabilidade na Construção Civil
Assunto	Predição da geração de RCC em obras da Construção Civil
Tema	Estimativa da quantidade de RCC gerados em canteiros utilizando RNA's

Fonte: A autoria própria (2021)

Na sequência, foram definidas as palavras-chave e delimitações de busca, formando-se uma *string* (sequência de caracteres utilizados para representar palavras, frases ou textos) com termos relevantes à pesquisa. A *string* de busca utilizou apenas palavras-chave em inglês, associadas a operadores *booleanos* (*and*, *or* e *not*). Além disso, foram elencados apenas artigos, capítulos de livros e livros, publicados entre 2011 e 2021. No Quadro 2 é apresentada a composição dos termos para a pesquisa.

Quadro 2 - Composição dos termos e *string* de busca

Palavras-chave	Resíduos de Construção e Demolição, Rede Neural, Modelos, Aprendizado de Máquina
Termos em inglês	Construction and Demolition Wastes, Neural Network, Models, Machine Learning
Delimitações da busca (<i>limit to</i>)	<i>Subject area: Environmental Science, Engineering, Computer Science, Mathematics.</i> <i>Document type: Article, Book Chapter</i>
<i>String</i> de busca	<i>(TITLE-ABS-KEY("construction and demolition waste") AND TITLE-ABS-KEY("neural network") OR TITLE-ABS-KEY("models") OR TITLE-ABS-KEY("machine learning")) AND (LIMIT-TO (SUBJAREA,"ENVI") OR LIMIT-TO (SUBJAREA,"ENGI") OR LIMIT-TO (SUBJAREA,"COMP") OR LIMIT-TO (SUBJAREA,"MATH")) AND (LIMIT-TO (DOCTYPE,"ar") OR LIMIT-TO (DOCTYPE,"ch") OR LIMIT-TO (DOCTYPE,"bk")) AND (LIMIT-TO (PUBYEAR,2021) OR LIMIT-TO (PUBYEAR,2020) OR LIMIT-TO (PUBYEAR,2019) OR LIMIT-TO (PUBYEAR,2018) OR LIMIT-TO (PUBYEAR,2017) OR LIMIT-TO (PUBYEAR,2016) OR LIMIT-TO (PUBYEAR,2015) OR LIMIT-TO (PUBYEAR,2014) OR LIMIT-TO (PUBYEAR,2013) OR LIMIT-TO (PUBYEAR,2012) OR LIMIT-TO (PUBYEAR,2011))</i>

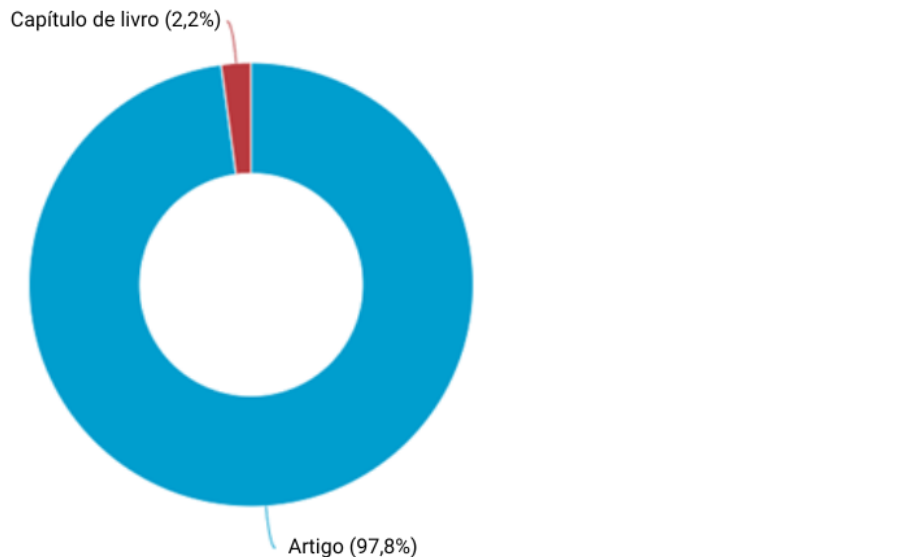
Fonte: A autoria própria (2021)

A partir da definição da *string* de busca, realizou-se a pesquisa na base de dados, e em seguida foi feita a análise das publicações relevantes para a pesquisa utilizando a própria ferramenta de análise dos resultados do *Scopus*.

2.1.2 Análise dos Resultados da Pesquisa Bibliográfica

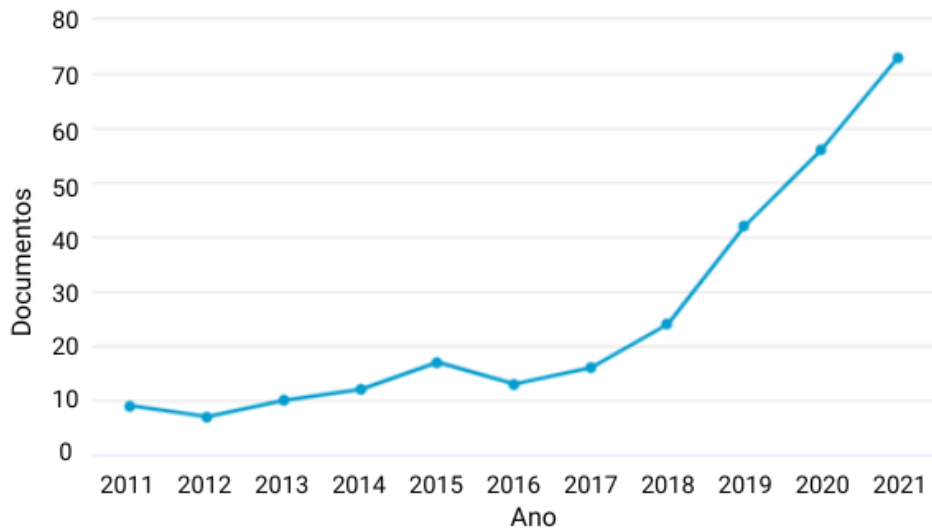
A *string* de busca apresentada foi realizada no banco de dados *Scopus* em Julho de 2021, resultando em 280 documentos. Desse total, foram 174 artigos e 6 capítulos de livros, conforme apresentado na Figura 1.

Figura 1 - Características das publicações analisadas



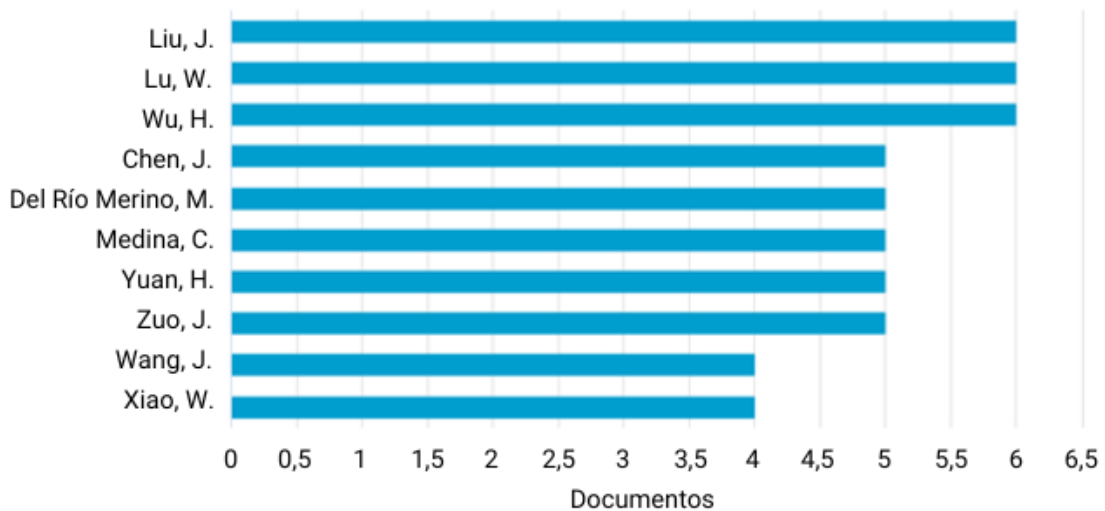
Fonte: Scopus (2021)

Quanto ao número de publicações por ano, verificou-se que o tema teve aumento no número de pesquisas relacionadas, de 9 publicações em 2011 para 74 publicações em 2021, conforme apresentado na Figura 2.

Figura 2 - Número de publicações por ano

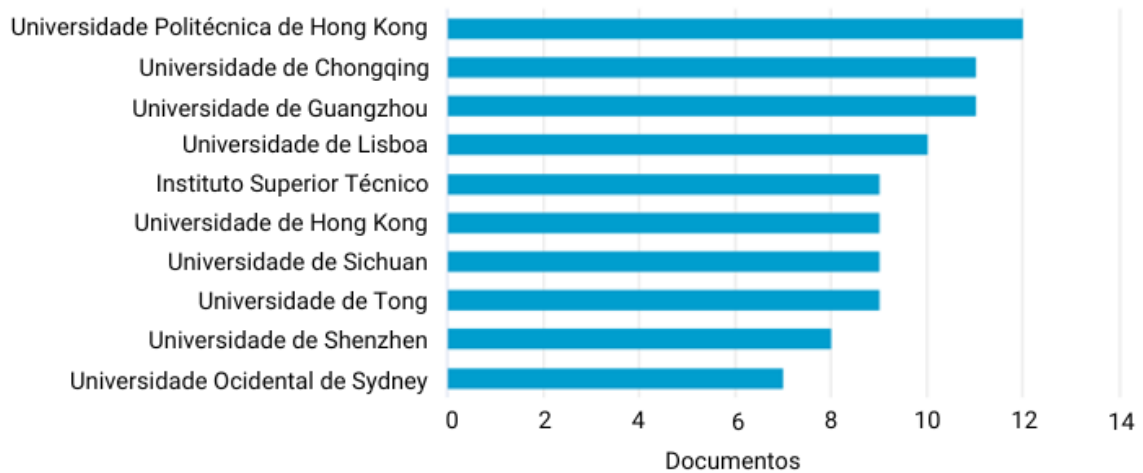
Fonte: Scopus (2021)

Jingkuang Liu e Weisheng Lu da China e Huanyu Wu da Austrália são os autores com maior número de publicações acerca do tema em questão, com total de 6 publicações, conforme apresentado na Figura 3.

Figura 3 - Número de documentos por autor

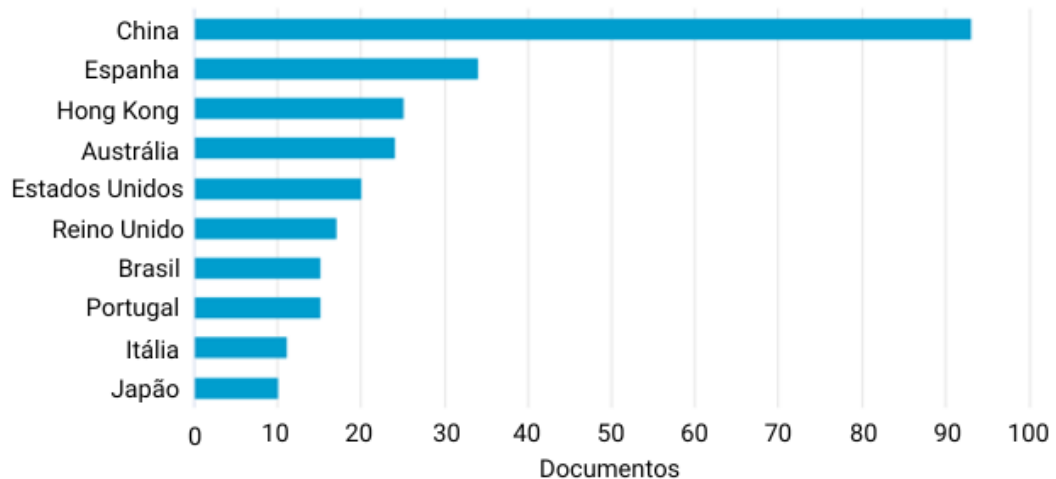
Fonte: Scopus (2021)

Com relação aos centros de pesquisa mais relevantes, a Universidade Politécnica de Hong Kong possui o maior número de publicações, com total de 12 documentos publicados, e a Universidade Ocidental de Sydney possui o menor número de publicações, com 7 documentos publicados sobre o tema (Figura 4).

Figura 4 - Número de documentos por instituição

Fonte: Scopus (2021)

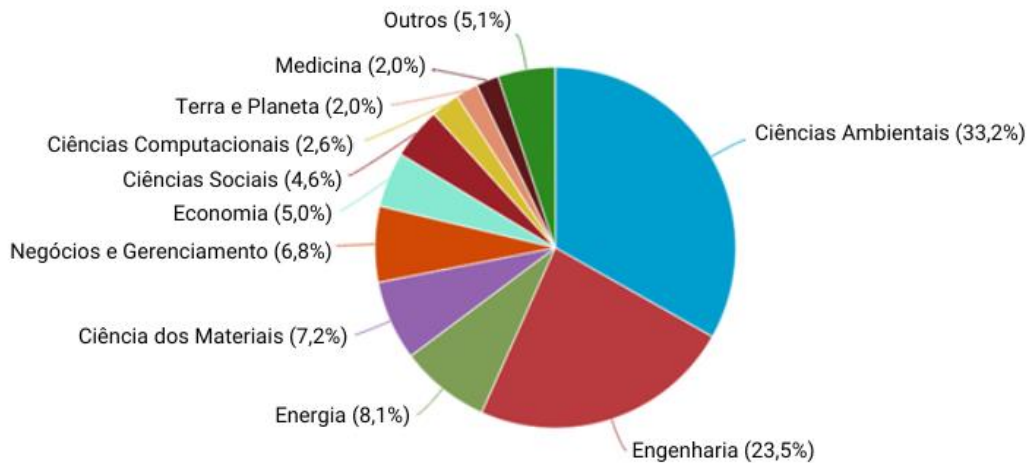
A respeito dos países com o maior número de publicações sobre o tema, a China lidera com 94 documentos, seguida pela Espanha com 34 documentos e por Hong Kong com 25 documentos. O Brasil ocupa a sétima posição, com 15 documentos publicados na área, conforme apresentado na Figura 5.

Figura 5 - Número de documentos por país

Fonte: Scopus (2021)

Acerca do número de publicações por área temática, a área de Ciências Ambientais possui mais documentos publicados, com 182 publicações, seguida pela área de Engenharia com 128 documentos e pela área de Energia com 44 documentos, conforme indicado na Figura 6.

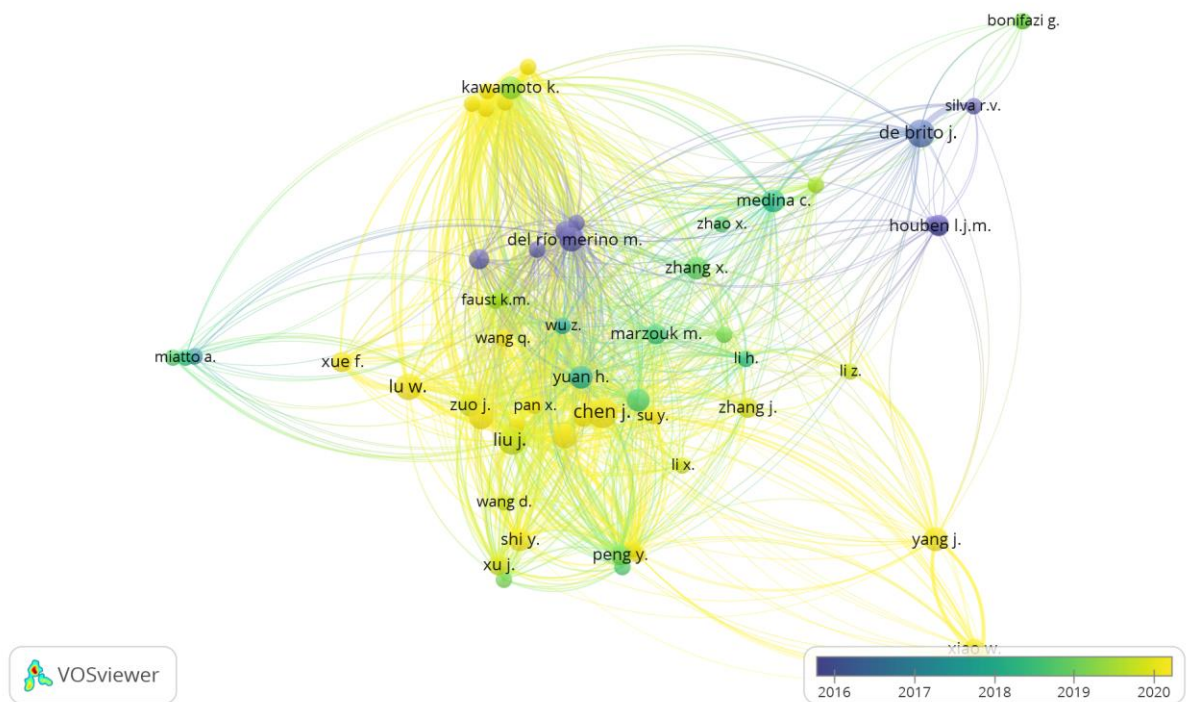
Figura 6 - Distribuição dos documentos por área temática



Fonte: Scopus (2021)

Utilizando o *software VOSviewer* para analisar os dados extraídos do *Scopus*, é possível obter o mapa de acoplamento bibliográfico por autores, apresentado na Figura 7. Esta configuração de mapa é importante, pois apresenta o *link* entre dois artigos que citam o mesmo, ou os mesmos documentos. Em outras palavras, noo mapa é apontado o conjunto de documentos que utilizam a mesma base bibliográfica.

Figura 7 - Acoplamento bibliográfico por autores

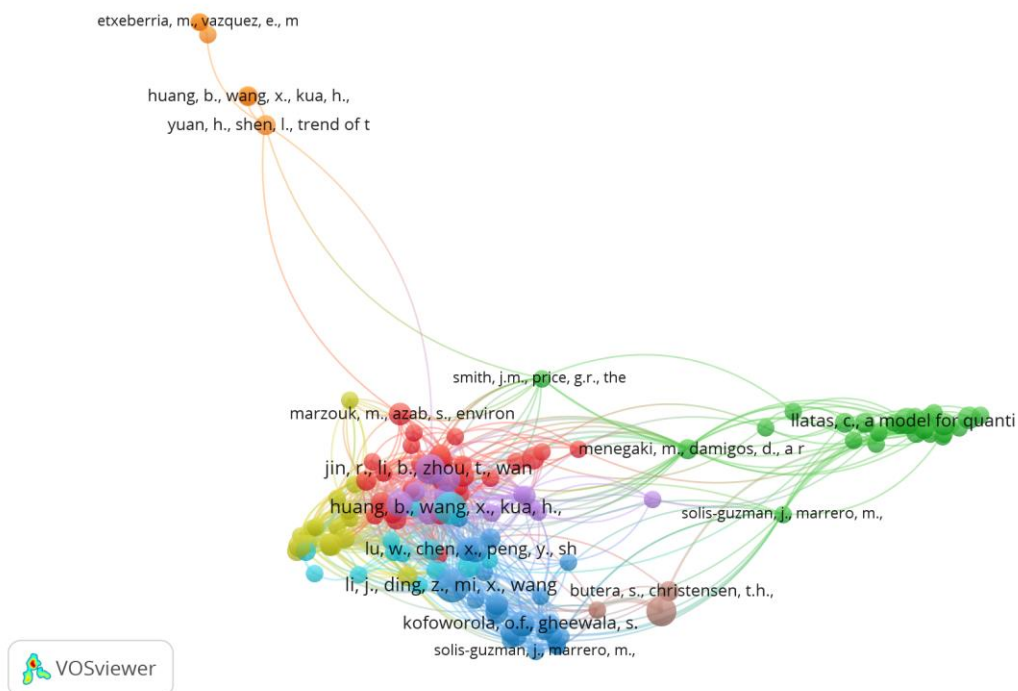


Fonte: VOSviewer (2021)

Na Figura 7 é apresentada uma escala temporal no formato *Overlay Visualization*, na qual os autores que estão em amarelo utilizam uma base de dados mais recente, e os que estão em tons de verde e roxo, utilizam uma base mais antiga. Além disso, pode-se notar que os autores que estão mais próximos indicam a utilização da mesma base bibliográfica, e os que encontram-se mais distantes, possuem uma abordagem um pouco diferente.

Na Figura 8 é apresentado um mapa de co-citações por referências, demonstrando o *link* entre dois artigos quando ambos são citados pelo mesmo documento. Essa análise é relevante para verificar quais artigos são mais frequentemente citados pelos autores, bem como para visualizar os métodos e conceitos nos quais os autores normalmente fundamentam-se.

Figura 8 - Mapa de co-citações por referências

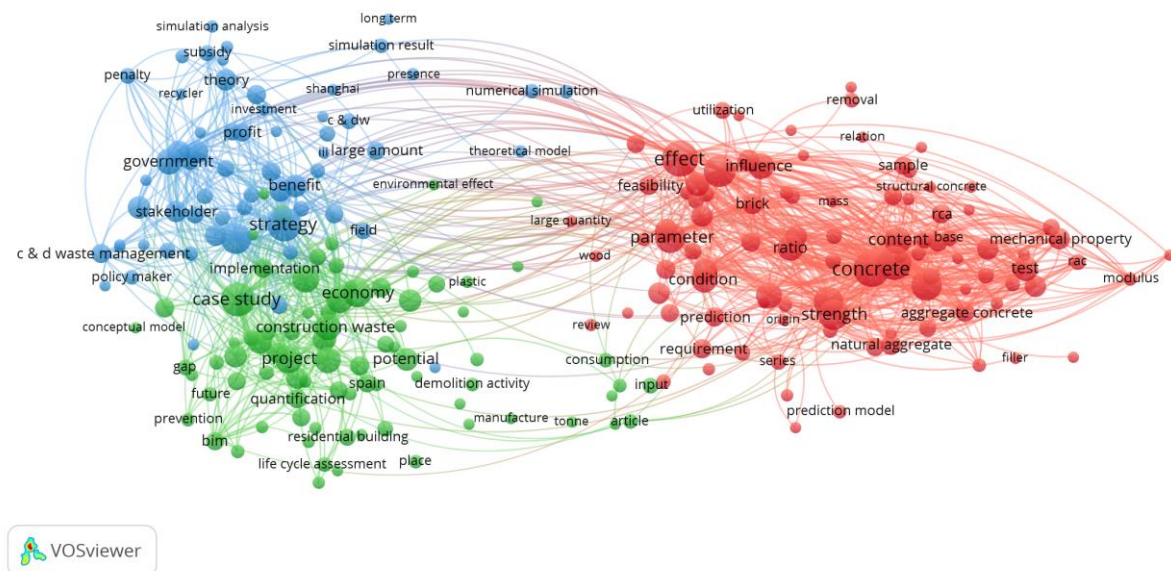


Fonte: VOSviewer (2021)

Os autores mais próximos são os que utilizam a mesma base conceitual, enquanto os mais distantes utilizam conceitos ou métodos diferentes. Além do mais, o tamanho do marcador que traz o nome do autor, indica a força do artigo ou a frequência com que este é citado. Quanto maior o indicador, mais vezes o artigo foi citado.

Na Figura 9 é apresentado o mapa baseado em texto relativo a análise dos termos mais significativos presentes no título e resumo (*abstract*) das publicações. Nessa análise, foi estipulado que o mapa apresentasse termos citados no mínimo 5 vezes em cada documento. É comum a análise dos termos utilizados como palavras-chave. No entanto, acredita-se que esse mapa a ser exibido torna-se mais representativo, uma vez que as palavras-chave normalmente ocorrem de maneiras repetitivas nos trabalhos do mesmo assunto.

Figura 9 - Termos presentes no título e *abstract* das publicações



Fonte: VOSviewer (2021)

Na Figura 9 são apresentados três *cluster* com cores diferentes, sendo vermelha, verde e azul. O *cluster 1*, representado na cor vermelha, é composto por termos relacionados mais às características do concreto e seus agregados, mostrando também os termos previsão e modelo de previsão, ambos relacionados com esta pesquisa. O *cluster 2*, indicado na cor verde, trata-se de uma abordagem relacionada à gestão dos RCC, estudos de caso, quantificações, construções residenciais, projetos, economia, incluindo também BIM e RNA. O *cluster 3*, apresentado na cor azul, corresponde ao conjunto de termos referentes à teorias e estratégias de gerenciamento de RCC, envolvendo partes governamentais, subsídios, benefícios e decisões, englobando análises e resultados de simulação, bem como simulações numéricas.

2.2 Referencial Teórico

O referencial teórico apresentado neste capítulo visa demonstrar a base teórica e conceitual sobre o tema desta pesquisa, apontando os conceitos básicos e consolidados da literatura, fundamentais para o desenvolvimento deste trabalho.

2.2.1 Resíduos da Construção Civil

A gestão de resíduos sólidos tornou-se um dos desafios globais mais críticos da atualidade, proporcionando efeitos negativos na economia e na sustentabilidade, principalmente em países menos desenvolvidos (DAOUD et al., 2021; SANTOS; TUBINO, 2021).

No Brasil, a Lei Federal nº 12.305 de 2 de Agosto de 2010 que institui a Política Nacional dos Resíduos Sólidos (PNRS), denomina resíduo sólido todo material, substância, objeto ou bem descartado resultante de atividades humanas em sociedade, cuja destinação se dá em estado sólido ou semissólido, bem como gases contidos em recipientes e líquidos cujas particularidades tornem inviável seu lançamento na rede pública de esgoto ou em corpos d'água, ou exijam para isso soluções técnica e economicamente inviáveis (BRASIL, 2010).

Na PNRS é proposta uma ordem de prioridade para a gestão e gerenciamento de resíduos, que consiste na não geração, redução, reutilização, reciclagem, tratamento e disposição final ambientalmente adequada. Além disso, a política adota alguns instrumentos considerando os princípios da responsabilidade compartilhada do gerador e do ciclo de vida dos produtos, entre os quais se enfatizam: planos e inventários de resíduos sólidos, coleta seletiva, sistema de logística reversa, incentivo ao desenvolvimento de cooperativas e associações de catadores, bem como extinção de lixões (BRASIL, 2010).

No âmbito dos resíduos da construção, na Resolução nº 307 de 5 de Julho de 2002, do Conselho Nacional do Meio Ambiente (CONAMA), são estabelecidos diretrizes, critérios e procedimentos para gestão de resíduos de construção civil. Além disto, os resíduos são definidos como sendo aqueles provenientes de construções, reformas, reparos e demolições de obras de construção civil, e os resultantes de escavação e preparação de terrenos, bem como tijolos, blocos cerâmicos e concreto em geral, solos, metais, madeiras, argamassa, gesso, telhas,

vidros, plásticos, fiação, dentre outros, também conhecidos como entulho de obras, calça ou metralha (BRASIL, 2002).

Na Resolução nº 307 (CONAMA, 2002) alterada pela Resolução nº 431/2011 (BRASIL, 2011), pela Resolução nº 448/2012 (BRASIL, 2012) e pela Resolução nº 469/2015 (BRASIL, 2015), é estabelecida a classificação dos resíduos de construção civil como (BRASIL, 2002):

- Classe A – resíduos reutilizáveis ou recicláveis como agregados, tais como: solos, resíduos de pavimentação, componentes cerâmicos, argamassa, concreto, peças pré-moldadas, blocos e tubos.
- Classe B – resíduos recicláveis para outras destinações, tais como: plásticos, papéis, papelão, metais, vidro, madeira, embalagens vazias de tintas imobiliárias e gesso.
- Classe C – resíduos para os quais não foram desenvolvidas tecnologias ou aplicações economicamente viáveis que permitam sua reciclagem ou recuperação.
- Classe D – resíduos perigosos, tais como: tintas, solventes, óleos, resíduos oriundos de clínicas radiológicas, materiais com amianto ou aqueles contaminados e prejudiciais à saúde.

Na Lei Federal nº 12.305 é proposto que cada município deve estabelecer políticas de gestão de RCC, tal como garantir sua efetiva aplicação. No entanto, é na Resolução nº 307 que é regulamentada a gestão obrigatória dos resíduos de construção em cada município pela implementação de um plano estratégico e concessão de responsabilidades claras, tanto para os geradores quanto para o órgão governamental responsável (BRASIL, 2002; SOUZA et al., 2022).

O problema da gestão de resíduos não afeta apenas o Brasil, mas também países onde a pequena extensão territorial e produção de recursos naturais agravam o reaproveitamento e reciclagem. No ano de 2016, cerca de 113 milhões de toneladas de RCC foram gerados na Europa. Nos Estados Unidos (EUA), foram produzidas 600 milhões de toneladas de resíduos de construção em 2018, sendo que desse total, mais de 90% tratava-se de resíduos provenientes de reformas, e menos de 10% foram resíduos de construção (SANTOS; TUBINO, 2021; TABOADA et al., 2020).

O Panorama dos Resíduos Sólidos no Brasil apresenta aumento da quantidade de RCC coletado, passando de 33 milhões de toneladas em 2010 para

48 milhões de toneladas em 2021. A quantidade coletada por habitante foi de aproximadamente 227 kg por ano e, grande parte dessa quantidade foi coletada em vias e logradouros públicos. Ainda, cerca de 6% desse total coletado é reciclado, e grande parte é destinado a aterros inertes e terrenos baldios (ABRELPE, 2022).

Considerando a necessidade pela busca de soluções técnica e economicamente viáveis para promover a gestão integrada eficiente dos RCC, os modelos preditivos de geração de resíduos podem ser utilizados como ferramentas para melhorar essa gestão. Os modelos preditivos contribuem com a não geração, primeiro item da hierarquia de gestão proposta pela PNRS, pois, uma vez que se conhece os tipos e quantidades de resíduos gerados, é possível adotar medidas que evitem a geração de determinado tipo de resíduo. Além disso, é possível reutilizar alguns resíduos, por exemplo, os de classe A, em outros serviços dentro do canteiro.

2.2.2 Métodos de Quantificação de RCC

Existem diversos métodos para estimar a quantidade de resíduos gerados em um empreendimento, sendo que, normalmente, eles são quantificados em peso ou volume. Apesar da complexidade do processo de geração, que é influenciado por inúmeros fatores diretos e indiretos, grande parte dos métodos de quantificação correlacionam esses fatores com as características das obras.

Wu et al. (2014) identificaram seis categorias de metodologias de quantificação de RCC, sendo elas, método de visita *in loco*, método da taxa de geração de resíduos, método de análise de vida útil, método de acumulação de sistema de classificação, método de modelagem de variáveis e outros métodos específicos como, por exemplo, o que considera porcentagem fixa de materiais comprados.

O método de visita *in loco* caracteriza-se por abordagens diretas ou indiretas de aferição do volume gerado de resíduos nas obras. O método de cálculo da taxa de geração de resíduos é o mais utilizado na literatura e, se baseia em utilizar a taxa de geração pela área bruta construída ou demolida (WU et al., 2014).

O método de análise de vida útil tem como princípio o balanço de massa do material. Presume-se que os edifícios construídos serão eventualmente demolidos, e conseqüentemente, a quantidade de resíduos de demolição deverá ser igual à massa da estrutura construída, e pode ser prevista determinando um tempo de vida razoável para edifícios (WU et al., 2014).

A metodologia de acumulação do sistema de classificação fornece uma plataforma para quantificar diferentes materiais. O sistema pode ser estabelecido de acordo com sistemas de classificação já existentes, como por exemplo a Lista Europeia de Resíduos (LER). A quantidade total de resíduos é calculada acumulando todos os itens a partir dos métodos de visita *in loco* e de cálculo de taxa de geração, uma vez que as quantidades dos itens foram determinadas (WU et al., 2014).

Ainda segundo os mesmos autores, o método de modelagem de variáveis permite prever quantidades de resíduos de construção e demolição utilizando *softwares* de modelagem. Essa metodologia permite entender a inter-relação entre as variáveis e fornece dados mais sistemáticos para a tomada de decisões.

Dentre os métodos identificados por Wu et al. (2014) existe ainda a quantificação de resíduos a partir de uma porcentagem fixa de materiais comprados. Por exemplo, se determinado que a taxa de geração de resíduos é de 10%, então a quantidade total de resíduos será 10% da quantidade de material comprado. Conforme os autores, deve-se escolher a metodologia adequada a cada cenário, de acordo com objetivos e condições reais de quantificação.

Llatas (2011) propôs um modelo para estimar a geração de RCC na fase de projeto, tendo como base a Lista Europeia de Resíduos (LER). O modelo permite identificar a origem dos resíduos e a partir disso adotar medidas para eliminar resíduos perigosos e reduzir a geração dos RCC.

O método se dá a partir de duas variáveis, a primeira trata-se da quantidade de edificações ou canteiros de obra e seus respectivos materiais de construção e componentes, e a segunda diz respeito ao processo de transformação desses materiais no decorrer da obra.

A primeira variável é obtida por meio de planilhas orçamentárias e medições em projeto. Já a segunda variável está relacionada com alguns fatores de transformação, que dependem da fonte e tipos de resíduos gerados, bem como a maneira que os materiais são fornecidos, embalados e coletado, padrão de qualidade da execução da obra, dentre outros. (LLATAS, 2011). A partir da Equação 1 é possível estimar a quantidade de RCC gerado.

$$CW_B = \sum_j CW_{SBEj} = \sum_{ji} CW_{BEi} = \sum_{ji} CW_{Pi} + \sum_{ji} CW_{Ri} + \sum_{ji} CW_{Si} \quad (1)$$

Em que:

CW_B : volume de RCC estimado para a construção;

CW_{SBEj} : volume de RCC estimado para o processo construtivo “j”;

CW_{BEi} : volume de RCC estimado para o elemento construtivo “i”;

CW_{Pi} : volume de resíduo de embalagem associado ao elemento construtivo “i”;

CW_{Ri} : volume de resíduos estimado para o elemento construtivo “i”;

CW_{Si} : volume de solo estimado para o elemento construtivo “i”.

Tratando-se dos resíduos de embalagens de cada elemento construtivo, seus tipos e quantidades, codificados conforme a Lista Europeia de Resíduos, podem ser estimados utilizando a Equação 2.

$$CW_{Pi} = \sum_k (EWL)_{Pk} \cdot Q_i \cdot F_P \cdot F_C \cdot F_I \quad (2)$$

Em que:

CW_{Pi} : volume de resíduo de embalagem associado ao elemento construtivo “i”;

EWL_{Pk} : código do resíduo de embalagem número “k”, conforme a LER;

Q_i : quantidade de elemento construtivo “i”, na unidade de medida do projeto (U);

F_P : fator de resíduos de empacotamento/embalagem;

F_C : fator de conversão;

F_I : fator de empolamento.

A Equação 2 pode ser utilizada para estimar a quantidade de perdas de materiais e solo, para cada elemento da construção, apenas substituindo o fator F_P pelo fator correspondente ao que se deseja estimar. Deste modo, Llatas (2011) definiu os fatores para geração de resíduos com base em estudos de caso, conforme apresentado na Tabela 1.

Tabela 1 - Fatores para geração de resíduos

Fator	Valor
F_P	0,750000
F_R	0,040000
F_S	0,787320
F_I	1,100000

Fonte: Nagalli (2022)

O fator F_C é utilizado somente em casos em que a unidade de medida do material de construção ou embalagem for diferente da unidade de medida do resíduo gerado, como por exemplo, o serviço de alvenaria que é normalmente quantificado em m^2 . Contudo que se queira estimar os resíduos em m^3 , o fator F_P corrigiria essa diferença.

Li et al. (2013) apontaram um método para estimar a quantidade de resíduos considerando a área de piso coberta da obra, incluindo as paredes externas e excluindo os telhados. Os autores utilizam as nomenclaturas “majoritário” e “minoritário” para classificar os materiais de acordo com a quantidade adquirida. A estimativa para resíduos gerados a partir dos materiais “majoritários” é feita a partir da quantidade de material adquirida. Para os resíduos gerados por materiais “minoritários”, a estimativa é feita calculando um percentual da quantidade total de resíduos de construção.

A partir da Equação 3 é possível calcular a quantidade total de resíduo gerado pelo empreendimento.

$$WG = \sum_{i=1}^n M_i \cdot r_i + W_0 \quad (3)$$

Em que:

WG: total de resíduo gerado pelo empreendimento, em massa (kg);

M_i : quantidade adquirida de material “majoritário” “i”, em massa (kg);

r_i : taxa de desperdício do material “i”;

W_0 : resíduo restante;

n: número de tipos de materiais “majoritários”.

Na sequência, é calculada a taxa de geração de resíduos por área de piso coberta (WGA), utilizando a Equação 4.

$$WGA = \frac{WG}{GFA} \quad (4)$$

Em que:

WGA: taxa de geração de resíduos por área de piso coberta da edificação, em m²;

WG: total de resíduo gerado pelo empreendimento, em massa (kg);

GFA: área bruta do projeto de construção, em m².

Por fim, a terceira etapa do método trata do cálculo da taxa de geração de resíduos para materiais majoritários “i”, dada pela Equação 5.

$$WGA_i = \frac{(M_i \times \Gamma_i)}{GFA} \quad (5)$$

Dias (2013) e Kern et al. (2015) apresentaram um método para quantificar resíduos de obras verticais. No modelo são considerados aspectos de projeto, como área, forma/compacidade, repetição de pavimentos e quantidade interna de paredes, bem como aspectos de construção, sistema construtivo, organização do canteiro de obras e formas de reaproveitamento. A fórmula proposta pelas autoras, ajustada estatisticamente, é expressa pela Equação 6.

$$VR = -5.202,886 + \left(5.138,519 * \frac{T}{T}\right) + (1,411 * ATP) + (22,968 * IeC) \\ + (375,155 * SP) + (-783,296 * RR) + \varepsilon \quad (6)$$

Em que:

VR: volume total de resíduos (m³);

$\frac{T}{T}$: razão entre o número de pavimentos tipo e o número total de pavimentos;

ATP: área do pavimento tipo a ser construído, em m²;

IeC: índice econômico de compacidade;

SP: sistema produtivo (conforme Quadro 3);

RR: reaproveitamento de resíduos (igual a 0 em caso de inexistência, igual a 1 se elas existem);

ε : erro estatístico.

Quadro 3 - Tipo de sistema produtivo

Sistema produtivo	Descrição	Valor na equação
Artesanal	Estrutura em concreto armado moldado <i>in loco</i> , alvenaria de vedação sem recortes, alvenaria interna de tijolos, produção de argamassa e corte e dobra de aço no canteiro, e poucos equipamentos de transporte na obra.	1
Intermediário quanto à industrialização	Estrutura em concreto armado moldado <i>in loco</i> , alvenaria de vedação sem recortes, alvenaria interna de tijolos, produção de argamassa e corte e dobra de aço industrializados, utilização de pré-fabricados e alguns equipamentos de transporte na obra.	2
Construtivo com práticas industrializadas	Estrutura em concreto armado moldado <i>in loco</i> , alvenaria de vedação externa em blocos racionalizados com projeto de paginação, alvenaria de divisão interna dos apartamentos em gesso acartonado, corte e dobra do aço e argamassa industrializada, utilização de pré-fabricados, ampla aplicação de equipamentos de transporte na obra e emprego da filosofia da construção enxuta, especialmente tratando-se do abastecimento otimizado de materiais nos postos de trabalho (ferramentas de gestão como <i>kanban</i> e <i>just in time</i>).	3

Fonte: Dias (2013)

O IeC considera arestas e planos curvos no cálculo através de aumento no perímetro, e pode ser calculado utilizando a Equação 7.

$$IeC = \frac{\sqrt[2]{A_p \cdot \pi}}{Pep} \times 100 \quad (7)$$

Em que:

A_p : superfície de projeto;

Pep : perímetro econômico de projeto.

O Pep é uma relação estabelecida entre o perímetro de um círculo de área igual ao do projeto e o perímetro das paredes exteriores do projeto, podendo ser calculado pela Equação 8.

$$Pep = Ppr + 1,5pc + \frac{nA}{2} \quad (8)$$

Em que:

Pep: perímetro das paredes exteriores retas;

Ppr: perímetro das paredes exteriores curvas;

nA: número de arestas das fachadas.

Outro método disponível para estimar a quantidade de resíduos gerados em obras de edifícios verticais foi proposto por Nagalli e Carvalho (2018), a partir dos seguintes critérios: área total construída, criticidade do cronograma de execução da obra, tipo de alvenaria de fechamento (cerâmica ou concreto), dimensão das equipes de trabalho e frequência da supervisão do processo de gestão dos resíduos. Este modelo considera resíduos de terraplenagem e é calculado a partir da Equação 9.

$$Q = 0,63 * A + 2.846,8 * T - 788,6 * L - 1.861,4 * C + 5.744,6 * M - 1.866,6 * Sc - 511,8 * R - 1.844,9 * P + 247,9 * SO - 916,1 * FC + 1.066,3 * FO + 6.686,6 \quad (9)$$

Em que:

Q: quantidade de resíduos estimada para a obra;

A: área total construída, em m²;

T: treinamento da equipe de trabalho (0 se treinada, 1 se não treinada);

L: tamanho da equipe de trabalho é menor que o desejável (1 se verdadeiro, 0 se falso);

C: tamanho da equipe de trabalho é compatível com o desejável (1 se verdadeiro, 0 se falso);

M: tipo de alvenaria de fechamento (1 se concreto, 0 se cerâmica);

Sc: fiscalização do gerenciamento de resíduos é rara ou inexistente (1 se verdadeiro, 0 se falso);

R: fiscalização do gerenciamento de resíduos de uma a quatro vezes ao mês (1 se verdadeiro, 0 se falso);

P: fiscalização do gerenciamento de resíduos mais do que uma vez na semana (1 se verdadeiro, 0 se falso);

SO: cronograma executivo “apertado” (1 se verdadeiro, 0 se falso);

FC: cronograma executivo regular (1 se verdadeiro, 0 se falso);

FO: cronograma executivo flexível e com folga (1 se verdadeiro, 0 se falso);

Além dos métodos apresentados, existem ainda métodos computacionais de cálculo, como por exemplo, os baseados em Redes Neurais Artificiais (RNA's). Estudos sobre este tipo de modelo preditivo vêm sendo reportados na literatura, como os de Lu et al. (2021), Coskuner et al. (2021) e Nagalli (2021a), que serão apresentados na sequência.

Independentemente do método adotado, vale a pena ressaltar que a geração de resíduos está diretamente ligada aos aspectos tecnológicos e culturais da região, formas de embalagem e transporte dos materiais, grau de industrialização do setor, entre outros fatores. No entanto, presume-se que o futuro da predição de resíduos esteja na automatização do processo, quer seja utilizando aprendizado de máquina ou outra técnica que a substitua parcial ou complemente (NAGALLI, 2021b).

2.2.3 *Machine Learning* e Redes Neurais Artificiais

Sendo visto como um dos pilares da Inteligência Artificial, o *Machine Learning* (ML) (aprendizado de máquina, em português) refere-se ao uso de algoritmos para organizar dados, reconhecer padrões e fazer com que os computadores aprendam com esses modelos, sem necessitar de pré-programação. É o que permite tornar a inteligência artificial real. Os modelos de aprendizado tomam decisões com base em experiências acumuladas pela solução bem sucedida de problemas anteriores. Além disso, possuem características particulares e comuns que permitem sua classificação quanto à linguagem de descrição, modo, paradigma e forma de aprendizado utilizado (BRASIL, 2018; MONARD; BARANAUSKAS, 2003).

Nesse contexto, a Rede Neural Artificial (RNA) é uma poderosa técnica de *machine learning* que tem sido adotada na previsão de resíduos sólidos. RNA's são construções matemáticas simplificadas baseadas no modelo biológico do sistema nervoso, contudo, com um conjunto limitado de neurônios. Esses neurônios processam simultaneamente os dados e os propagam através de uma complexa malha de interconexão. Assim como o cérebro humano, as RNA's são capazes de interagir com o meio externo e adaptar-se a ele (FINOCCHIO, 2014; MONARD; BARANAUSKAS, 2003).

A maneira como os neurônios estão estruturados está intimamente ligada com o algoritmo de aprendizagem utilizado para treinar a rede. Assim sendo, o treinamento de uma arquitetura específica consiste na aplicação de passos ordenados para ajustar os pesos e intensidades mínimas para que os neurônios produzam uma resposta. Esse processo de ajuste, também conhecido como algoritmo de aprendizagem, visa sintonizar a rede para que suas respostas estejam próximas dos valores desejados (HAYKIN, 2009; SILVA; SPATTI; FLAUZINO, 2016).

As principais arquiteturas de RNA's, considerando a disposição dos neurônios, formas de ligação entre eles e a constituição de suas camadas, podem ser divididas em: redes *feedforward* (alimentação à frente) de camada simples, redes *feedforward* de camadas múltiplas, redes recorrentes e redes reticuladas.

O Quadro 4 apresenta uma síntese sobre alguns tipos de redes neurais.

Quadro 4 - Tipos de Redes Neurais (continua)	
Rede Neural	Descrição
<i>Rede Perceptron e Adaline</i>	Esse tipo de rede possui apenas uma camada de entrada e apenas uma camada oculta, que é a própria camada de saída. O fluxo de informações é unidirecional, da entrada para a saída e, a quantidade de saídas desse tipo de rede sempre coincidirá com o número de neurônios. Essas redes são empregadas para classificação de padrões e filtragem linear.
<i>Rede Perceptron multicamadas e Redes de base radial</i>	As redes com esta arquitetura possuem uma ou mais camadas ocultas de neurônios. A quantidade de camadas ocultas e de neurônios dependem da complexidade do problema e da quantidade e qualidade dos dados disponíveis. O número de neurônios que compõem a primeira camada escondida é normalmente diferente do número de sinais que compõem a camada de entrada da rede. Entretanto, assim como nas redes <i>feedforward</i> de camadas simples, a quantidade de sinais de saída sempre coincidirá com o número de neurônios daquela respectiva camada.
Redes recorrentes	São redes em que as saídas dos neurônios são realimentadas como sinais de entrada para outros neurônios. Esse tipo de rede pode ser utilizada em sistemas variantes em relação ao tempo, como previsão de séries temporais, otimização e identificação de sistemas e controle de processos, etc. Entre os principais tipos de redes com retroalimentação estão a rede de <i>Hopfield</i> e a <i>Perceptron</i> multicamadas com retroalimentação

	entre neurônios situados em camadas distintas.
Redes com estrutura reticulada	As principais características desse tipo de rede estão na disposição espacial dos neurônios, que está diretamente relacionada com o processo de ajuste de seus pesos e limiares. Podem ser aplicadas em problemas de agrupamento (<i>clustering</i>), reconhecimento de padrões, otimização de sistemas, entre outros. A principal rede com esta arquitetura é a rede de <i>Kohonen</i> .

Fonte: Adaptado de Silva; Spatti; Flauzino (2016)

A modelagem de RNA divide o conjunto total de amostras em dois subconjuntos, sendo eles treinamento e teste. O subconjunto de treinamento é composto por cerca de 60 a 90% das amostras totais, usadas no processo de aprendizagem da rede. Em contrapartida, o subconjunto de teste, é composto por 10 a 40% do total das amostras, utilizadas para verificar se a generalização de soluções da rede está em níveis aceitáveis, possibilitando a validação do modelo definido. Em cada fase de treinamento os dados de validação medem a capacidade de generalização do modelo, e o processo de treinamento é encerrado quando essa não melhora mais (LU et al., 2021; SILVA; SPATTI; FLAUZINO, 2016).

Para analisar a qualidade de ajuste dos modelos computacionais, é comum utilizar como métrica o coeficiente de determinação (R^2) e o coeficiente de correlação (R). O R^2 representa o percentual da variância dos dados que são explicados pelo modelo. É um número não negativo que assume valores entre 0 e 1, sendo que, quanto mais próximo de 1 maior é a correlação entre as variáveis, ou seja, mais explicativo é o modelo. Já o R , é uma medida do grau de relação entre duas variáveis, assumindo valor positivo ou negativo, entre -1 e 1 (JÚNIOR OLIVEIRA, 2021).

É importante utilizar outros parâmetros de análise como este, para ter uma visão global sobre a performance do modelo. Além do R^2 , também podem ser utilizados como métricas o erro quadrático médio e erro percentual médio absoluto (PALA, 2019).

O Erro Quadrático Médio (*Mean Squared Error – MSE*) trata-se da média da diferença quadrática entre o valor de predição do modelo e o valor real. Quanto maior o valor de MSE, significa que o modelo não performou bem com relação às estimativas. Já o RMSE (*Root Mean Square Error*) é calculado como a raiz quadrada do MSE, o que faz com que a unidade de medida seja a mesma do dado original.

Por exemplo, se o dado original for em metros, o RMSE será em metros também. O RMSE é uma medida mais interpretável e intuitiva do erro, já que está na mesma escala dos dados originais.

O Erro Percentual Médio Absoluto (*Mean Absolute Percentual Error – MAPE*) é uma métrica que mostra o percentual de erro com relação aos valores reais. Por exemplo, se o valor de MAPE for igual a 40%, significa que o modelo faz previsões que em média a diferença entre o valor previsto e o real equivale a 40% (JÚNIOR OLIVEIRA, 2021).

Os primeiros estudos envolvendo redes neurais ocorreram em 1943, ano em que McCulloch e Pitts apresentaram um modelo de neurônio como uma unidade de processamento binária e revelaram que esta unidade seria capaz de realizar inúmeras operações lógicas. A estrutura ficou conhecida como a primeira rede neural e, apesar da simplicidade, foi fundamental para a produção dos primeiros computadores digitais e elaboração dos primeiros modelos matemáticos de dispositivos artificiais que visavam analogias biológicas (FINOCCHIO, 2014).

Finocchio (2014) aponta ainda que, na década de 50, foi criado o primeiro neurocomputador, chamado de Mark I Perceptron e um novo elemento de processamento de redes neurais, chamado de *Adaline (Adaptive Linear Element)*. No entanto, ainda que o desenvolvimento do *Perceptron* e da *Adaline* tenham proporcionado avanços na área de redes neurais, o fato da maioria das pesquisas ser de natureza heurística e não apresentar resultados teóricos que justificassem o interesse científico pela área, proporcionou uma redução no número de novas pesquisas acerca do tema.

A partir de 1970, os pesquisadores descobriram novas possibilidades na área de engenharia e da lógica matemática, e na década de 80, as pesquisas em redes neurais voltaram definitivamente, permitindo sua aplicação em sistemas reais. Desde então, as RNA's tem tido diversas aplicações, dentre elas: identificação de sistemas, reconhecimento de padrões e processamento de imagens (FINOCCHIO, 2014). Na área de construção civil, inúmeros trabalhos apresentam o emprego de redes neurais para prever taxas de geração de resíduos, resistência à compressão do concreto e identificação de RCC em canteiros de obras. Na sequência, serão apresentados alguns estudo sobre esse tema.

2.3 Revisão de Literatura

2.3.1 Estudos Utilizando Redes Neurais Artificiais

No contexto de resíduos sólidos, Coskuner et al. (2021) aplicaram a rede neural artificial *perceptron* multicamadas para verificar a previsão das taxas de geração anual de resíduos domésticos, comerciais e de construção e demolição. A pesquisa utilizou dados coletados no aterro Askar, no Reino do Bahrein, entre 1997 e 2016. O erro quadrático médio (MSE) e o coeficiente de determinação (R^2) foram utilizados como indicadores de desempenho para avaliar a eficácia dos modelos desenvolvidos. Os modelos apresentaram uma forte precisão na estimativa de resíduos com valores de R^2 iguais a 0,95, 0,99, 0,91, para os resíduos domésticos, comerciais e RCC, respectivamente. Os autores concluíram que modelos utilizando RNA's são eficazes para a previsão de resíduos de diferentes fontes e podem ser considerados como uma boa abordagem econômica para o planejamento de sistemas integrados de gestão de resíduos.

Versando sobre resíduos da construção, Nagalli (2021a) avaliou o desempenho de RNA's na previsão da quantidade de resíduos gerados em 330 obras de edifícios verticais. O modelo utilizou redes neurais com dois, cinco e dez neurônios na camada oculta, bem como área de piso e duração do trabalho como dados de entrada. O coeficiente de determinação (R^2) obtido foi igual a 0,96 e a melhor configuração proposta foi capaz de prever com precisão 43% dos casos. O autor concluiu que os modelos de predição utilizando RNA's funcionam melhor do que os modelos que utilizam regressão múltipla linear, comumente encontrados na literatura.

Abordando resíduos de demolição, Cha; Hong e Kim (2023) desenvolveram um estudo sobre a melhoria de desempenho do modelo de aprendizado de máquina usando codificador automático para prever a taxa de geração de resíduos de demolição. Os autores combinaram três algoritmos, sendo eles redes neurais (*Perceptron Multicamadas*), regressão vetorial de suporte e floresta aleatória com um *autoencoder* para desenvolver e testar modelos híbridos de aprendizado de máquina. Foram coletadas informações de 782 edifícios antes de suas demolições, posteriormente, foram coletadas as quantidades de resíduos que os edifícios geraram após sua demolição. A partir das simulações utilizando essas informações, o modelo que apresentou melhor resultado foi a combinação de redes neurais com o

autoencoder, que quando comparado a um modelo não híbrido, melhorou em 49% o erro médio absoluto, em 27% a raiz quadrada do erro médio, em 49% o coeficiente de determinação e em 22% o coeficiente de correlação. Com a melhora nos resultados, os autores concluíram que o modelo híbrido é considerado uma abordagem nova e vantajosa para a gestão dos resíduos de demolição utilizando aprendizado de máquina.

No âmbito do reconhecimento de imagens, Davis et al. (2021) projetaram uma rede convolucional profunda para classificar resíduos de construção e demolição usando imagens digitais de resíduos depositados em uma lixeira no canteiro de obras. Em 442 imagens, a lixeira continha resíduos de madeira, plásticos, concreto, tijolos, papelão e poliestireno. Em 1.023 lixeiras havia mais de um tipo de resíduo e em 183 imagens a lixeira estava vazia. O conjunto de imagens foi testado em dois experimentos, o primeiro foi treinado para classificar as imagens em 8 classes, de acordo com o tipo de resíduo da lixeira. No segundo experimento, a rede foi treinada para classificar os resíduos da lixeira em simples e mistos. O primeiro experimento resultou em 94% de acerto na classificação das categorias, e o segundo, apresentou 92% de precisão na classificação.

Tratando-se de concretos com agregados reciclados, Liu et al. (2021) utilizaram o aprendizado de máquina para investigar a profundidade de carbonatação no composto. Um conjunto de 593 dados foi utilizado para treinar, validar e testar modelos de *machine learning*. Os resultados mostraram que o modelo utilizando floresta aleatória teve desempenho superior ao modelo de regressão do progresso gaussiano e modelo de RNA autônoma. Todos os modelos de aprendizado de máquina apresentaram maior precisão do que os modelos de código e modelos estatísticos existentes, com R^2 acima de 0,90.

Nesse segmento, Tam et al. (2022) utilizaram análise de regressão e a rede neural *perceptron* multicamadas para prever a força compressiva de concreto com injeção de CO_2 . As variáveis consideradas foram relação água/cimento, adição de agregado reciclado, tempo de câmara, pressão da câmara, cimento, água e areia. Os resultados apontaram um R^2 de 0,88 na análise de regressão e 0,95 para a RNA. Além disso, a rede neural apresentou excelente precisão, com erro médio de 1,24 MPa ou 3,43%. O erro máximo apresentado pela rede neural foi de 10%, que, segundo os autores, se deve ao fato das variações relacionadas à natureza do agregado reciclado.

Ainda no âmbito de concretos, Ray et al. (2022), desenvolveram um modelo computacional capaz de prever a resistência à tração e à compressão do concreto feito com pó de pedra e fibra de nylon. Para avaliar a precisão do modelo desenvolvido, utilizando redes neurais, foram avaliadas as resistências à compressão dos concretos aos 7 e 28 dias, e os resultados apresentaram um R^2 de 0,89 e 0,94, para os concretos nessas idades, respectivamente. Os autores apontaram que o modelo foi capaz de interpretar cerca de 90% dos dados estimados, sendo um resultado satisfatório, uma vez que os valores previstos ficaram próximos aos valores reais de resistência.

Na área de pavimentação, Ghorbani et al. (2021) utilizaram aprendizado de máquina para avaliar as propriedades térmicas e mecânicas de resíduos de demolição em pavimentos geotérmicos. Os testes foram feitos em diferentes concretos com agregados reciclados, resíduos de tijolos, resíduos de rochas e pavimento asfáltico recuperado, realizados em temperaturas de 5, 20, 35 e 50°C. Os concretos com adição de resíduos de tijolos e agregados reciclados apresentaram maiores condutividades térmicas do que os concretos com resíduos de rochas e pavimento asfáltico recuperado. O modelo computacional apresentou R^2 igual a 0,99, e os autores concluíram que redes neurais são ferramentas úteis e confiáveis para prever a deformação de concretos com materiais reciclados.

O aprendizado de máquina está presente em diversas empresas de construção civil, não apenas como ferramentas de previsão de resíduos, mas também em plataformas que detectam a segurança, qualidade e uso de equipamentos no canteiro. Além disso, o aprendizado de máquina também pode ser utilizado para aumentar a assertividade de orçamentos e auxiliar nas vendas imobiliárias, reconhecendo padrões e prevendo tendências.

3 ESTRATÉGIA DE PESQUISA

Este capítulo apresenta a estratégia utilizada para o desenvolvimento desse estudo, dividido em classificação da pesquisa e etapas empregadas.

3.1 Classificação da Pesquisa

Para classificação desta pesquisa, foram adotados alguns critérios descritos a seguir. Previamente, o método de pensamento científico adotado foi o indutivo. Tratando-se da sua finalidade, a pesquisa é classificada como pesquisa aplicada, uma vez que está voltada à resolução de um problema presente. De acordo com seus objetivos, a pesquisa é classificada como descritiva, pois visa descrever características e estabelecer relações entre variáveis (GIL, 2002).

A forma de intervenção relacionada ao fenômeno trata de estudo experimental, já que as variáveis foram manipuladas em ambiente controlado, visando verificar os efeitos da intervenção sob o objeto de estudo. Com relação às fontes de dados, pode ser classificada como laboratório, pois os dados foram obtidos de maneira artificial e com certo grau de controle.

Tratando-se da abordagem, a pesquisa se encaixa como quantitativa, pois são utilizadas técnicas matemáticas e estatísticas para análise dos dados. Quanto ao tempo, a pesquisa trata-se de estudo transversal, uma vez que os dados foram produzidos em um único momento.

Com relação às técnicas de pesquisa utilizadas, empregou-se a bibliometria, simulação, *machine learning* e Redes Neurais Artificiais, nas etapas de método de pesquisa e obtenção e análise de resultados.

No Quadro 5 são sintetizados os critérios utilizados nessa pesquisa.

Quadro 5 - Critérios e classificação da pesquisa

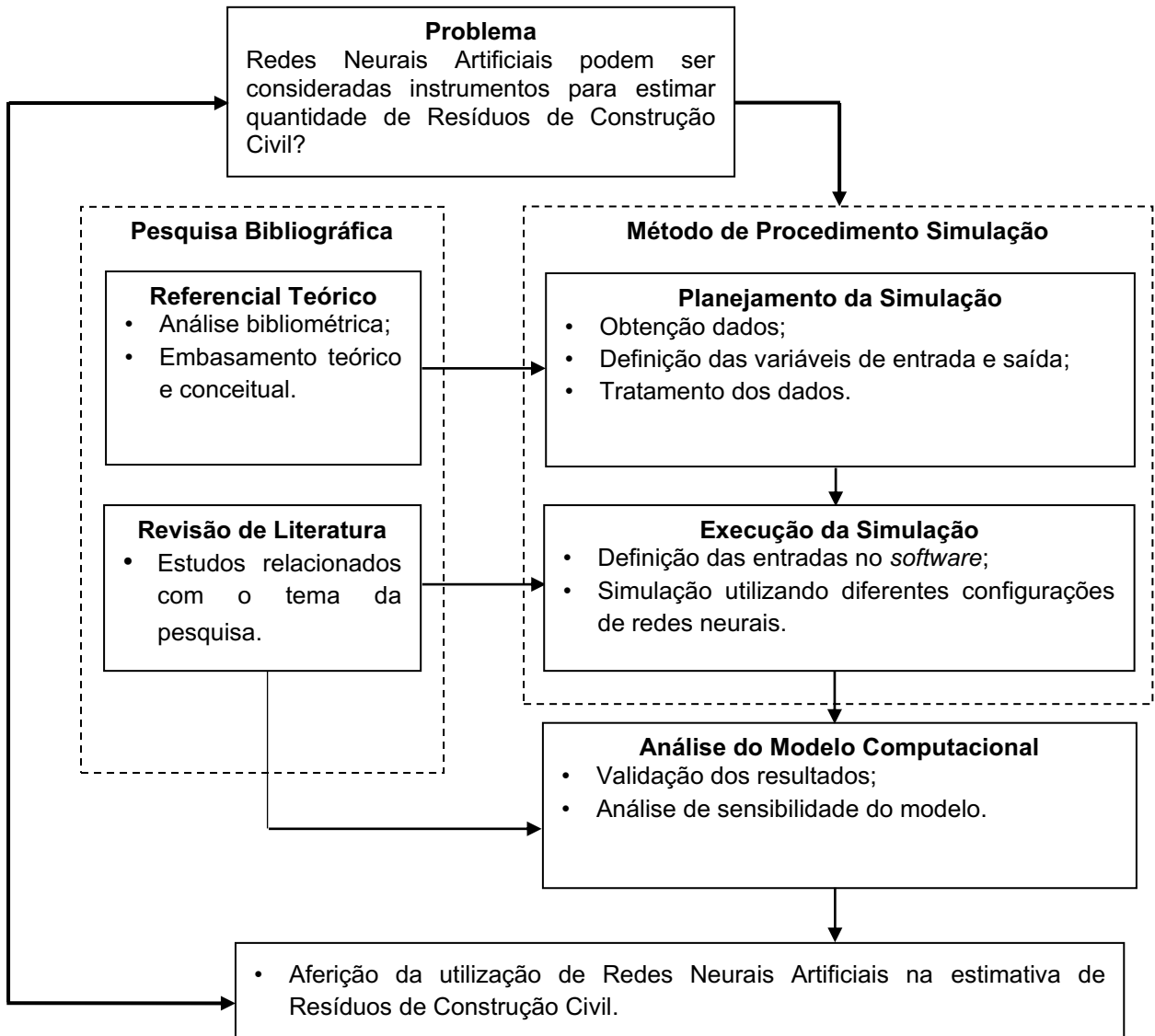
Critério	Classificação
Finalidade	Aplicada
Objetivo	Descritiva
Forma de Intervenção	Experimental
Fonte de Dados	Laboratório
Abordagem	Quantitativa
Tempo de Intervenção	Transversal
Procedimentos	Pesquisa Bibliográfica e Simulação

Fonte: Autoria própria (2022)

3.2 Etapas da Pesquisa

Na Figura 10 é apresentado um fluxograma com as principais etapas da pesquisa.

Figura 10 - Etapas da estratégia de pesquisa



Fonte: Autoria própria (2022)

A fim de facilitar a compreensão da estratégia utilizada nesta pesquisa, na Figura 10 são apresentadas as etapas do método de pesquisa e suas inter-relações. possível notar que, na Pesquisa bibliográfica, as etapas Referencial teórico e Revisão de literatura, fundamentam o método de procedimento e as análises dos dados, a fim de responder o problema de pesquisa. Em seguida, as etapas serão descritas de forma mais detalhada.

3.2.1 Planejamento da Simulação

Para aplicação da simulação, foram utilizadas duas fontes de dados, sendo a primeira de natureza fictícia, criada utilizando o editor de planilhas Excel, e a segunda, de natureza real, com informações obtidas junto a empresas de construção. As informações dos dados de natureza fictícia se referem a áreas totais construídas e quantidades de resíduos de obras de construção e de demolição, e as informações dos dados de natureza real, se referem a áreas totais construídas e quantidades de resíduos de construção.

Com relação ao tamanho mínimo da amostra para utilizar redes neurais, esse pode variar dependendo do tipo de problema a ser resolvido e do número de entradas e saídas da rede neural. Por exemplo, se a rede neural tiver 2 entradas e 1 saída, o número de parâmetros da rede neural será de 2 (pesos das conexões entre as entradas e a camada oculta) + 1 (pesos das conexões entre a camada oculta e a saída) + 3 (vieses das camadas oculta e de saída) = 6. Nesse caso, um tamanho de amostra mínimo recomendado seria de pelo menos 60 dados.

A fim de executar as simulações com os dados fictícios, foram criadas duas amostras, denominadas Amostra A e Amostra B, com dados fictícios de 5.000 e 10.000 obras, respectivamente. As amostras variam de 75 a 125.050 m² de área total construída, sendo que a Amostra A teve um intervalo de 50 m² entre os dados de área, e a Amostra B teve um intervalo de 25 m² entre os dados de área. Apesar de ambas as amostras terem a mesma variação total de área construída, elas possuem tamanhos diferentes e, portanto, diferentes resoluções de dados. Como resultado, o intervalo entre os dados na Amostra A foi de 50 m² para manter a mesma variação total de área construída, enquanto o intervalo na Amostra B foi de 25 m², que possui uma resolução mais fina de dados. Isso significa que a Amostra B pode capturar variações menores na área construída do que a Amostra A, mas a Amostra A pode ser mais fácil de manipular e analisar devido ao menor número de

dados. Para estabelecer os valores de áreas fictícias, procurou-se trabalhar com valores condizentes com áreas reais de construção e demolição. Todavia, para validação do modelo de rede neural, foram utilizadas áreas reais de obras de edifícios verticais e suas respectivas quantidades de resíduos, também conhecidas.

Com a finalidade de obter a quantidade de RCC gerado em cada obra fictícia, de acordo com a área, utilizou-se a taxa de geração de resíduos da construção civil de 150 kg/m^2 , proposta por Pinto (1999). Já para obter a quantidade de resíduo de demolição gerado conforme a área, foi utilizada a taxa de geração de resíduo de demolição de 800 kg/m^2 , proposta pela Prefeitura Municipal de Recife (RECIFE, 2019 *apud* NAGALLI, 2021b).

Considerando essas duas taxas de geração de resíduos disponíveis na literatura, todos os dados de área da Amostra A, que variaram de 75 m^2 a 125.050 m^2 , foram multiplicados primeiramente por 150 kg/m^2 , gerando as quantidades de resíduos de construção para cada dado de área da amostra. Em seguida, a mesma Amostra A teve todos os seus dados de área multiplicados por 800 kg/m^2 , originando as quantidades de resíduos de demolição, para cada dado de área da amostra.

O mesmo processo foi utilizado na Amostra B, os dados de área foram multiplicados por 150 kg/m^2 , gerando as quantidades de resíduos de construção e, na sequência, os dados de área foram multiplicados por 800 kg/m^2 , originando as quantidades de resíduos de demolição. A diferença é que a Amostra B possui 5.000 dados, com áreas variando entre 75 m^2 a 125.050 m^2 , com intervalo de 25 m^2 cada dado de área.

Os dados de área da Amostra A foram distribuídos em uma coluna apenas, sendo 2.500 áreas variando de 75 m^2 a 125.050 m^2 e, na sequência, na mesma coluna, essas 2.500 áreas foram repetidas. As 2.500 primeiras áreas receberam uma coluna ao lado, com os dados de resíduos de construção calculados e, a partir da área de número 2.500, onde as áreas começaram a se repetir, na coluna ao lado, foram adicionados os dados de resíduos de demolição.

Dessa forma, a Amostra A teve uma coluna com 2.500 áreas, repetidas uma vez, e uma coluna com 2.500 quantidades de resíduos de construção e 2.500 quantidades de resíduos de demolição. A fim de diferenciar esses dois tipos de resíduos, acrescentou-se na amostra, uma coluna binária de "Classificação", utilizando o valor "0" para resíduos da construção, e "1" para resíduos de demolição.

Para a Amostra B, se procedeu da mesma forma. Inicialmente, 5.000 dados de áreas foram distribuídos em uma coluna, e repetidos, logo abaixo, na mesma coluna. Até o dado de área número 5.000, as quantidades de resíduos de construção foram distribuídas em uma coluna ao lado. A partir do dado de área número 5.001 até o dado número 10.000, na coluna com informações de resíduos, foram inseridas as quantidades de resíduos de demolição. Essa amostra também possui a coluna de “Classificação”, indicando “0” para resíduos da construção, e “1” para resíduos de demolição.

Para obtenção da base de dados reais, inicialmente foi realizado um levantamento das empresas de construção civil de Curitiba/PR e Região Metropolitana. Em seguida, entrou-se em contato com essas empresas a fim de solicitar informações sobre as áreas de construção e quantidades de resíduos gerados, sendo obtidos 100 dados de área total construída e quantidade de resíduos. A fim de formar uma amostra maior para utilizar na validação do modelo, os 100 dados obtidos foram unidos com os 330 dados utilizados por Nagalli (2021a), em seu estudo. A partir disso, as informações foram passadas para uma planilha do Excel, com uma coluna com 430 áreas totais construídas e uma coluna com 430 quantidades de resíduos de construção gerados em cada obra. Devido ao fato de não existirem dados reais de resíduos de demolição, essa amostra, denominada Amostra R, não possui a coluna de “Classificação”.

Com a finalidade de retirar os *outliers* ou “valores atípicos” da amostra, foi construído um diagrama *box plot*. Para tanto, utilizando o Microsoft® Excel® foram calculados a mediana (M_d), o primeiro quartil (Q_1), terceiro quartil (Q_3) e a amplitude interquartil (IQR) da Amostra R. A partir disso, foram calculados os limites inferior ($LI = Q_1 - (1,5) * IQR$) e superior ($LS = Q_3 + (1,5) * IQR$). Os valores compreendidos entre esses dois limites são chamados de “valores adjacentes”, bem como os valores que estiverem acima do limite superior ou abaixo do limite inferior estabelecidos, foram considerados *outliers* e retirados da amostra (BUSSAB; MORETTIN, 2017).

Para as simulações, as Amostras A e B, criadas a partir de taxas genéricas de geração de resíduos, incluem qualquer tipo de edificação. A Amostra R, utilizada na validação do modelo de rede neural, compreende dados de um tipo específico de obras, sendo dados de obras de edifícios verticais. Após a criação, obtenção e

tratamento dos dados, foi elaborado um planejamento das entradas e saídas no *software*, sintetizadas no Quadro 6.

Quadro 6 - Planejamento das simulações

Amostra	Número de dados	Variável de entrada	Variável de saída
A	5.000	Classificação e Área total construída	Quantidade de resíduo
B	10.000	Classificação e Área total construída	Quantidade de resíduo
R	430	Área total construída	Quantidade de resíduo

Fonte: Autoria própria (2022)

3.2.2 Execução da Simulação

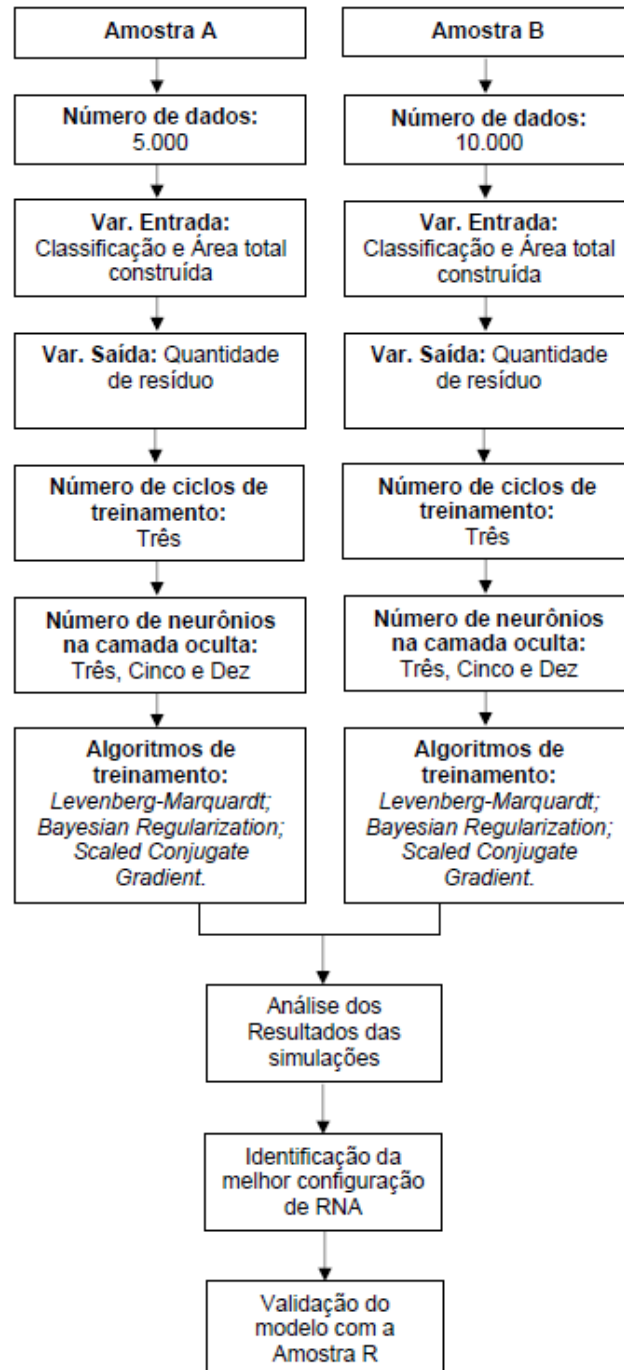
O *software* MATLAB®, versão R2022a, foi utilizado a fim de verificar a aplicabilidade das redes neurais na área de resíduos da construção civil. As redes neurais utilizadas nesta pesquisa classificam-se como redes neurais *feed-forward*, também conhecidas como *Perceptron*. Possuem apenas uma camada de neurônios, em que cada neurônio recebe um conjunto de entradas e produz uma única saída. A camada de entrada recebe as entradas da rede, a camada de saída produz as saídas finais e a camada intermediária, também conhecida como camada oculta, realiza o processamento dos dados.

Para realizar as simulações, foram importados no *software* os dados contendo as variáveis de entrada e saída e, em seguida, dentro do aplicativo, o conjunto de amostras foi dividido aleatoriamente em três subconjuntos, 70% para treinamento, 15% para validação e 15% para teste. Essas porcentagens são sugeridas pelo *software*, para que se tenha boa distribuição da amostra, mas podem ser ajustadas caso necessário. Na sequência, foi ajustado o número de neurônios que seria utilizado e selecionado o algoritmo com o qual seria feito o treinamento da rede neural. Após isso, a rede neural foi treinada e os resultados do treinamento, incluindo os gráficos de regressão e histogramas de erro, ficaram disponíveis no aplicativo. No Apêndice B encontra-se um passo a passo de como foram realizadas as simulações.

Na Figura 11 é apresentada a sequência das simulações e as configurações utilizadas para executá-las, sendo elas, quantidade de dados, variáveis de entrada e

saída, números de ciclos de treinamento, números de neurônios na camada oculta e algoritmos de treinamento utilizados.

Figura 11- Configurações e sequências das simulações



Fonte: Autoria própria (2022)

Conforme a Figura 11, inicialmente, foi importada a Amostra A de 5.000 dados, as variáveis de entrada foram a classificação do resíduo (0 para resíduos de construção, 1 para resíduos de demolição) e a área total construída da obra, e as

variáveis de saída foram as quantidades de resíduos a serem preditas. Em seguida, se avaliou qual o número mínimo de ciclos de treinamento poderia ser aplicado. Com base no estudo de Nagalli (2021) se optou por realizar três ciclos de treinamento para três, cinco e dez neurônios e para cada algoritmo disponível no *software*. Por exemplo, para o algoritmo *Levenberg-Marquardt* foram realizados três ciclos de treinamento com a configuração de rede neural com três neurônios, três ciclos com a configuração com cinco neurônios e três ciclos com a configuração com dez neurônios. Para os algoritmos *Bayesian Regularization* e *Scaled Conjugate Gradient*, se procedeu da mesma forma.

Treinar várias vezes o conjunto de dados gera resultados diferentes devido a diferentes condições iniciais e amostragens. Pode ocorrer que durante alguma fase do treinamento a generalização comece a perder a qualidade, causando *overfitting*, ou seja, a rede se especializa no conjunto de dados de treinamento e perde sua capacidade de generalização para os outros conjuntos de dados (LEE; KIM; KIM, 2016).

Após as simulações com a Amostra A, foram realizadas as simulações com a Amostra B, com 10.000 dados, que também possui duas variáveis de entrada sendo a classificação do resíduo (0 resíduo de construção, 1 resíduo de demolição) e a área total construída da obra, e uma variável de saída, que é a quantidade de resíduo. Para a Amostra B, também foram realizados três ciclos de treinamento para três, cinco e dez neurônios, com os três algoritmos disponíveis no aplicativo (*Levenberg-Marquardt*, *Bayesian Regularization* e *Scaled Conjugate Gradient*). Por fim, foram realizadas as etapas de análise dos resultados, identificação da configuração de rede neural que teve o melhor desempenho e validação do modelo computacional utilizando a Amostra R com dados reais, descritas na sequência.

3.2.3 Análise e Validação do Modelo Computacional

Para avaliação de desempenho do modelo computacional, os resultados das simulações das Amostras A e B foram analisados com base nos parâmetros R^2 , RMSE e MAPE. A partir da análise dos resultados das simulações, foi identificada qual a melhor configuração de rede neural, ou seja, quantos ciclos de treinamento, quantos neurônios na camada oculta e qual algoritmo de treinamento, teve o melhor desempenho.

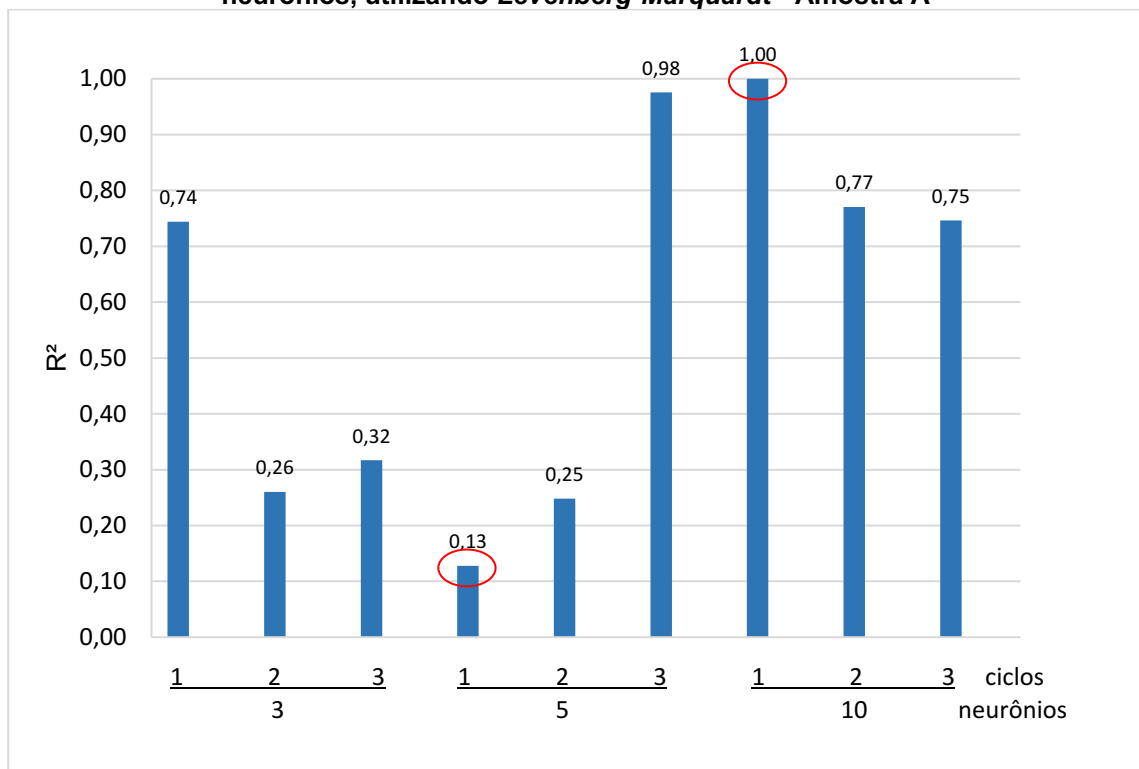
A fim de validar o modelo, a melhor configuração de rede neural obtida nas simulações com as Amostras A e B, foi aplicada à Amostra R, composta por dados reais de área total construída e quantidades de resíduos geradas. Após essa simulação, foi validada a utilização do modelo computacional, bem como foi realizada a análise de sensibilidade do modelo quanto ao seu desempenho preditivo com relação a um ou dois dados de entrada no *software* e quanto à natureza dos dados e ao tamanho das amostras. Além disso, as variáveis de saída que foram as quantidades de resíduos calculadas pela rede neural, foram comparadas com as variáveis de calibração do modelo, ou seja, com as quantidades reais de resíduos fornecidas pelas empresas.

4 RESULTADOS E DISCUSSÕES

4.1 Simulação de Redes Neurais com a Amostra A

Para as simulações utilizando a Amostra A, com duas variáveis de entrada (Classificação do resíduo e área total construída) e uma de saída (quantidade de resíduo), o algoritmo de treinamento *Levenberg-Marquardt* foi o que apresentou pior desempenho na estimativa de resíduos, com valores de R^2 variando entre 0,13 e 1,0. No Gráfico 1 são apresentados os valores de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt*.

Gráfico 1 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt* - Amostra A



Fonte: Autoria própria (2022)

A partir do Gráfico 1 verifica-se que nos ciclos de treinamento utilizando três neurônios, o primeiro ciclo apresentou o maior valor de R^2 , sendo igual a 0,74. Além disso, a capacidade preditiva não melhorou nos outros dois ciclos de treinamento com três neurônios.

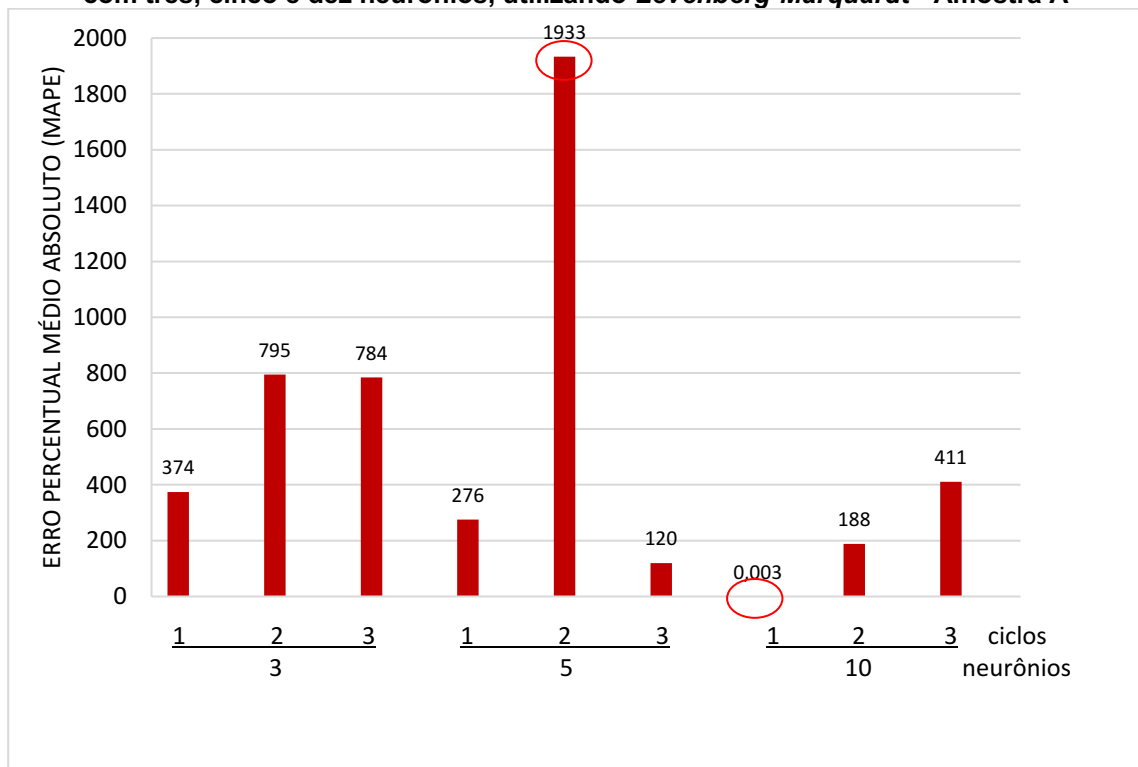
No treinamento com cinco neurônios, verificou-se que os resultados preditivos melhoraram ao longo dos três ciclos, chegando ao R^2 igual a 0,97 no terceiro ciclo. Notou-se ainda, que a RNA melhorou sua capacidade de predição ao longo dos 5.000 dados da amostra, prevendo quantidades muito diferentes das

esperadas no início, e quantidades de resíduos muito parecidas com o alvo, ao final da amostra, significando que a rede neural foi aprendendo a prever os valores de saída, ao longo do treinamento.

Para o algoritmo *Levenberg-Marquardt*, a configuração de rede neural que apresentou maior R^2 (1,0) foi o primeiro ciclo de treinamento com dez neurônios. Desde o início da amostra, até o final, os valores preditos foram muito similares aos reais, o que pode ser verificado através do erro percentual médio absoluto (MAPE), que para essa configuração, foi igual a 0,003%. Considerando o melhor resultado de R^2 (0,97) para a configuração com cinco neurônios, utilizar uma rede neural com 10 neurônios pode até ser desnecessário, uma vez que pode ocasionar *overtraining* onde a rede acaba sendo treinada por muito tempo, resultando em baixa capacidade de previsão.

No Gráfico 2 são apresentados os valores de MAPE, que é a média da diferença entre o valor estimado e o real, em percentual, para os três ciclos de treinamento das três quantidades de neurônios.

Gráfico 2 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt* - Amostra A

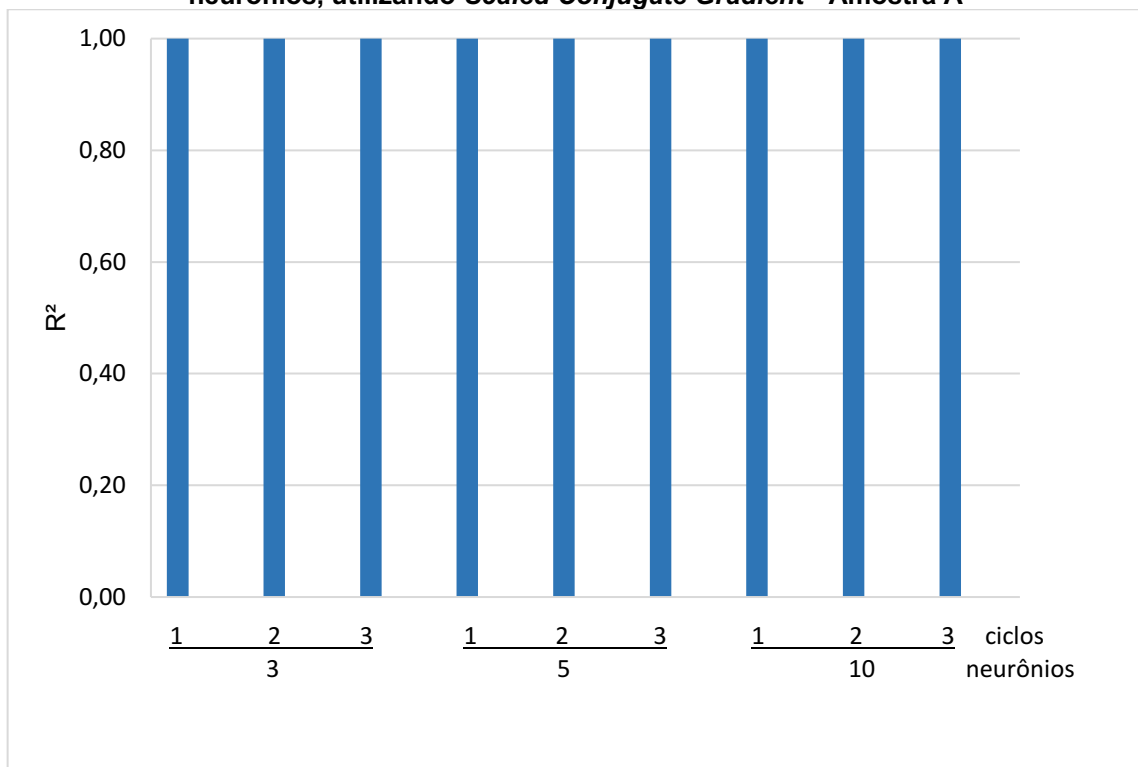


Fonte: Autoria própria (2022)

Conforme o Gráfico 2, nota-se que os valores de MAPE variaram de 374% a 1933% entre o primeiro ciclo de treinamento com três neurônios, e o segundo ciclo com cinco neurônios, o que indica percentual de erro muito alto entre as quantidades de resíduos estimadas e as reais. No terceiro ciclo de treinamento com cinco neurônios, o valor de MAPE diminuiu para 120%, ainda apresentando valor alto. Porém, aumentando para dez o número de neurônios na camada oculta, no primeiro ciclo de treinamento o valor de MAPE reduziu para 0,003%, indicando que essa configuração de rede neural teve o menor percentual de erro de todos os ciclos e quantidades de neurônios treinadas utilizando *Levenberg-Marquardt*, uma vez que treinando essa configuração por mais dois ciclos, o valor de MAPE aumentou.

Para o algoritmo *Scaled Conjugate Gradient*, todos os três ciclos de treinamento, com três, cinco e dez neurônios na camada oculta apresentaram valores de R^2 entre 0,99 e 1,0, conforme exibido pelo Gráfico 3.

Gráfico 3 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Scaled Conjugate Gradient* - Amostra A



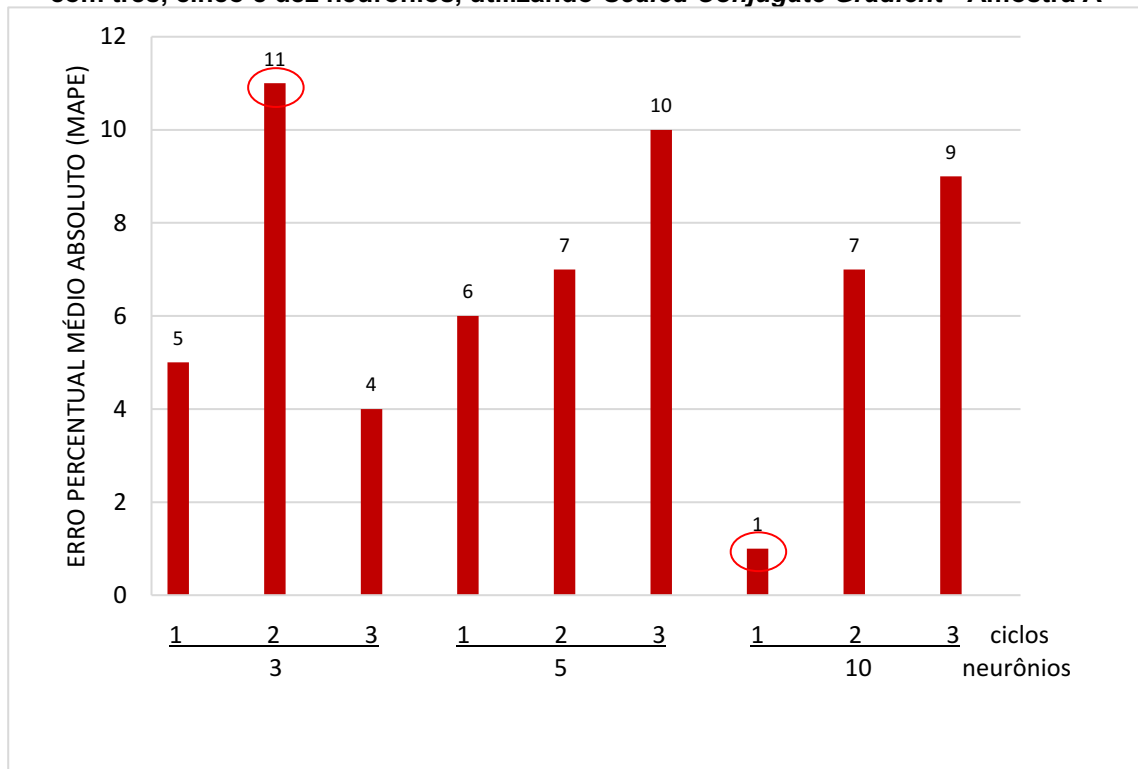
Fonte: Autoria própria (2022)

Contudo, a estimativa apresentou tanto quantidades muito superiores às esperadas, quanto quantidades próximas às esperadas, de maneira aleatória. Isso foi verificado utilizando três, cinco e dez neurônios e três ciclos de treinamento cada.

Apesar de o valor de R^2 igual a 1 indicar que a regressão linear ajustada explica perfeitamente a variação dos dados, isso não significa necessariamente que a estimativa dos dados será precisa. Nesse caso, pode ter ocorrido *overfitting*, ou seja, é possível que o modelo tenha se ajustado muito bem aos dados de treinamento, que representa 70% da amostra, mas não tenha generalizado bem para os demais dados, levando a estimativas imprecisas.

No Gráfico 4 é apresentada a média da diferença entre o valor estimado e o real, em percentual (MAPE), para os três ciclos de treinamento das três quantidades de neurônios utilizando *Scaled Conjugate Gradient*.

Gráfico 4 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Scaled Conjugate Gradient* - Amostra A



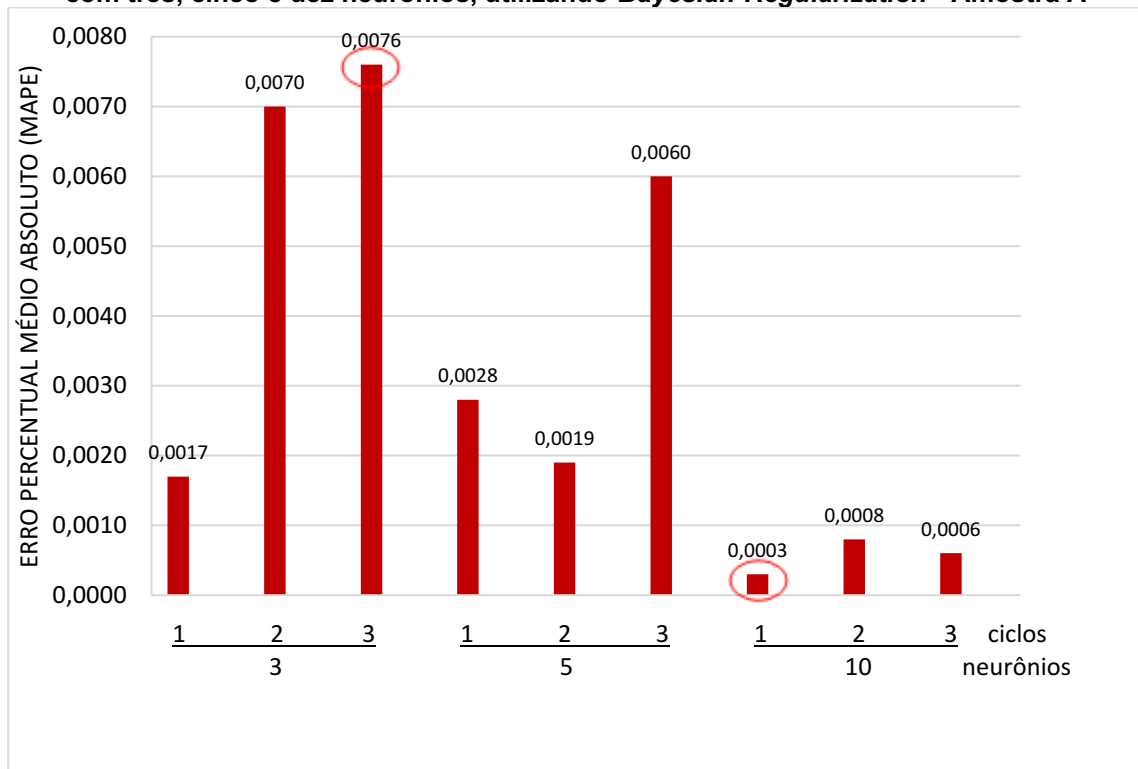
Fonte: Autoria própria (2022)

A partir do Gráfico 4 é possível verificar que a configuração de rede neural que apresentou o menor valor de MAPE (1%) foi o primeiro ciclo com dez neurônios. Além disso, o percentual médio dos erros absolutos ficou entre 1 e 11% para os três ciclos de treinamento utilizando três, cinco e dez neurônios na camada oculta e o algoritmo de treinamento *Scaled Conjugate Gradient*. Os baixos valores de MAPE podem confirmar a ocorrência de *overfitting*, o modelo pode ter conseguido bom desempenho nas estimativas do conjunto de dados de treinamento, mas não

conseguiu generalizar para os demais conjuntos (15% para teste e 15% para validação). Em outras palavras, a rede neural conseguiu boa estimativa com os dados de treinamento, que foram a maior parte da amostra, refletindo em baixo erro percentual médio absoluto.

Nas simulações utilizando o algoritmo *Bayesian Regularization* e três neurônios na camada oculta, as quantidades previstas foram muito próximas das quantidades alvo. No entanto, verificou-se que aumentando o número de neurônios para cinco e dez, a capacidade preditiva aumentou e o erro percentual médio absoluto diminuiu, como é possível observar no Gráfico 5.

Gráfico 5 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Bayesian-Regularization* - Amostra A



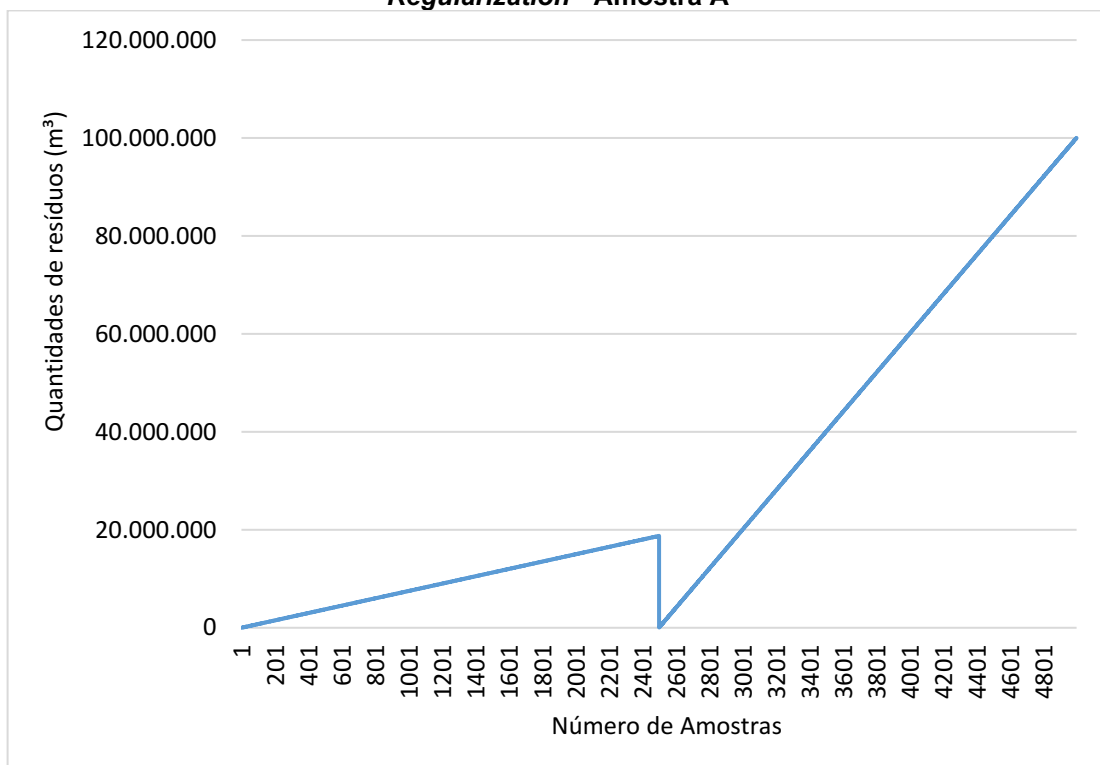
Fonte: Autoria própria (2022)

A melhor configuração de RNA utilizando esse algoritmo, foi o primeiro ciclo com dez neurônios, que apresentou valor de R^2 igual a 1,0, MAPE igual a 0,0003% e RMSE igual a 11,56 kg. De modo geral, o valor de R^2 foi igual a 1,0 para todas as configurações de redes neurais treinadas com o algoritmo *Bayesian Regularization*, e os valores de MAPE variaram entre 0,0076% (terceiro ciclo com três neurônios) e 0,003% (primeiro ciclo com dez neurônios). Após o primeiro ciclo de treinamento

com dez neurônios na camada oculta, não foi verificada melhora no desempenho preditivo, com base no MAPE.

No Gráfico 6 são apresentadas as quantidades de resíduos estimadas na simulação da Amostra A, com rede neural *feed-forward*, um ciclo de treinamento e dez neurônios na camada oculta, utilizando o algoritmo de treinamento *Bayesian Regularization*.

Gráfico 6 - Quantidades de resíduos estimadas pela rede neural *feed-forward* com dez neurônios na camada oculta, um ciclo de treinamento e algoritmo de treinamento *Bayesian Regularization* - Amostra A

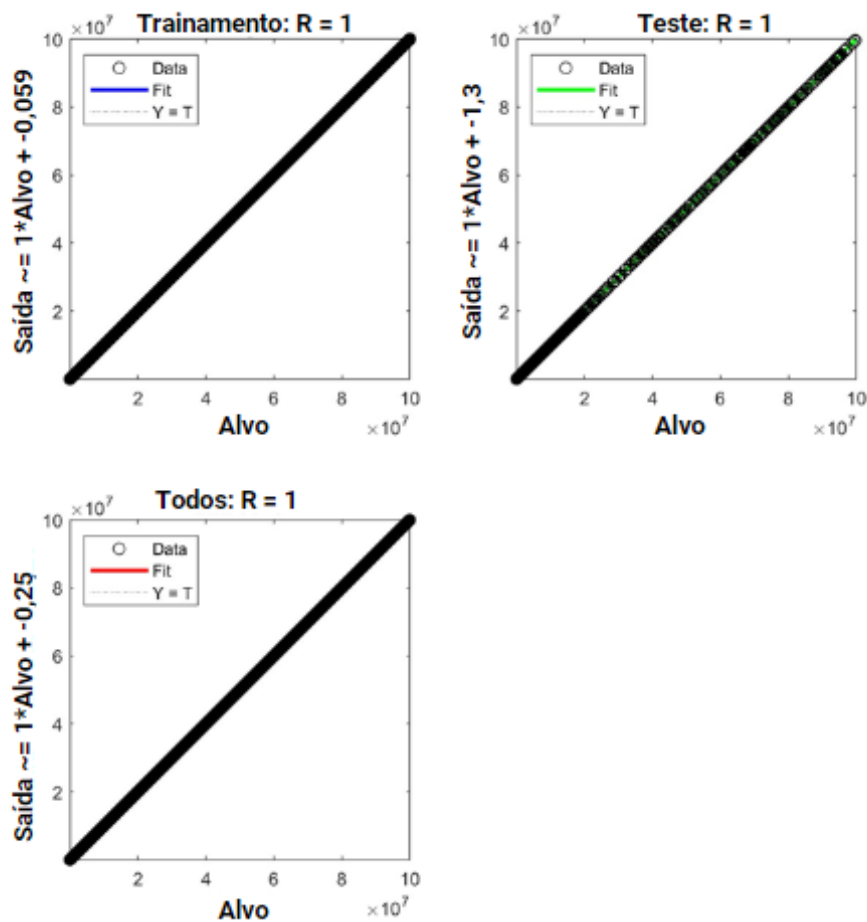


Fonte: Autoria própria (2022)

A partir do Gráfico 6, é possível verificar que as quantidades de resíduos estimadas utilizando a rede neural, ficaram entre 11.040,01 kg a 100.020.0190,01 kg, muito próximas às quantidades calculadas para a Amostra A, que foram entre 11.250,00 kg e 100.020.000,00 kg. No gráfico é possível notar ainda uma quebra na linha azul, que indica as quantidades de resíduos estimadas, na amostra 2.509, essa quebra indica onde a classificação do resíduo mudou de 0 (resíduo de construção) para 1 (resíduo de demolição). Na utilização do algoritmo de treinamento *Bayesian Regularization* as diferentes configurações de redes neurais foram capazes de estimar tanto resíduos de construção, quanto resíduos de demolição.

Em seguida, serão apresentadas as figuras exportadas do MATLAB® contendo os gráficos de regressão linear, gráfico de erro quadrático médio e histograma de erro, métricas para avaliação do modelo utilizando um ciclo de treinamento, dez neurônios, e o algoritmo *Bayesian Regularization*, considerando que este apresentou o melhor desempenho no treinamento dessa amostra, verificado pelo menor valor de MAPE. Na Figura 12 é possível verificar o gráfico de regressão com os valores de R para as amostras de treinamento e teste.

Figura 12 - Gráficos de Regressão Linear para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra A



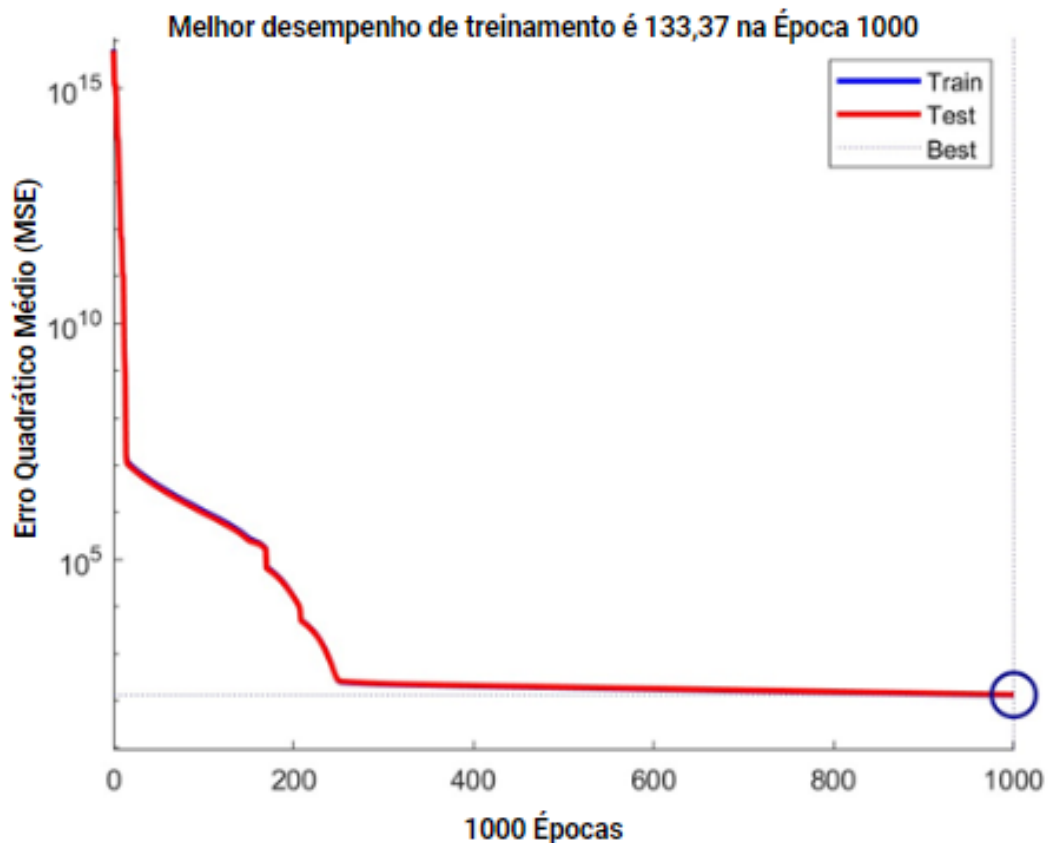
Fonte: MATLAB® (2022)

Os gráficos de regressão apresentados na Figura 12 são as previsões da rede, plotadas em relação ao alvo. Os dois gráficos apresentados no topo representam os dados de treinamento e teste. A linha tracejada, indicada na legenda como $Y = T$ representa o resultado da previsão menos as saídas, resultando nos valores alvo. A linha sólida vermelha (Fit) representa a linha de regressão linear de melhor ajuste entre as estimativas e o alvo e, o círculo (Data) indica dos dados da amostra.

O valor de R é uma indicação da relação entre as saídas e os alvos. No caso do $R = 1,0$, isso indica que existe uma relação linear exata entre os valores preditos e valores alvo, bem como, indica que o modelo de estimativa é confiável. Quanto mais próximo de 0 o valor de R, menor é a relação linear entre os dados, conseqüentemente, menos confiável é o modelo. A linha de cor preta ajustada entre os cantos inferior esquerdo e superior direito, aponta que a rede aprendeu a ajustar bem os dados, nesse caso, pode ser pela linearidade dos dados, considerando que esses foram construídos. Devido ao ajuste dos dados à linha de regressão, o símbolo indicador dos dados encobriu as linhas dos valores alvo e a linha de regressão linear, tracejada e vermelha, respectivamente, não as deixando visíveis nos gráficos.

Na Figura 13 é apresentado o gráfico do Erro Quadrático Médio com o melhor valor de desempenho no treinamento da rede neural com um ciclo de treinamento e dez neurônios, utilizando o algoritmo *Bayesian Regularization*.

Figura 13 - Gráfico do Erro Quadrático Médio para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra A

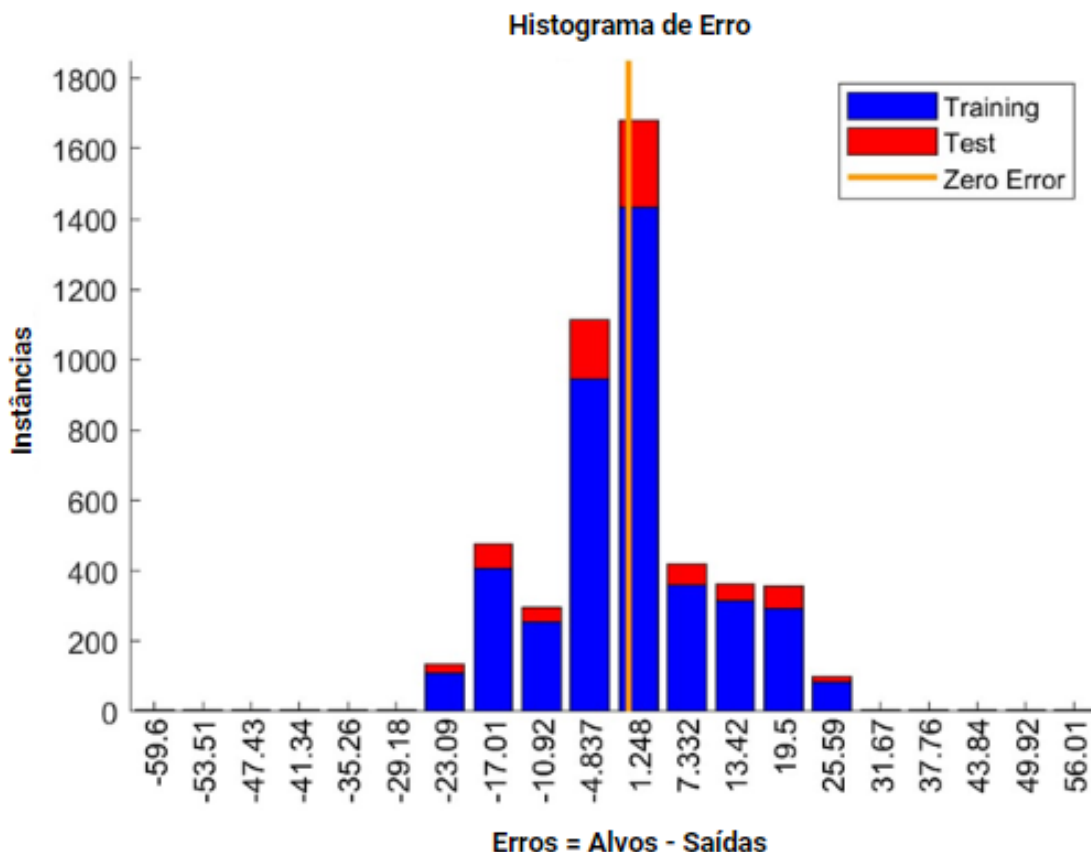


Fonte: MATLAB® (2022)

Na Figura 13 é possível verificar o desempenho da rede neural em termos de Erro Quadrático Médio, apresentado em escala logarítmica no eixo vertical. No eixo horizontal estão as épocas, que são a medida do número de vezes que todos os vetores de treinamento são usados uma vez para atualizar os pesos. No caso desse treinamento, todas as amostras passam pelo algoritmo de aprendizado simultaneamente, em uma época antes que os pesos sejam atualizados. O gráfico apresenta os erros dos conjuntos de dados de treinamento e teste, indicados pelas linhas de cores azul e vermelho, respectivamente. A linha tracejada indica a melhor performance.

É possível notar que o MSE diminuiu rapidamente à medida em que a rede foi treinada, tendo o melhor desempenho nas amostras de treinamento, na época 1000. Nota-se ainda que, o erro do conjunto de treinamento e do conjunto de teste têm características semelhantes, indicado pela sobreposição das linhas azul e vermelha. Uma verificação adicional do desempenho dessa configuração de rede neural pode ser verificada através do histograma de erro, apresentado da Figura 14.

Figura 14 - Histograma de Erro para o primeiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra A



Fonte: MATLAB® (2022)

Conforme o histograma da Figura 14, no eixo vertical encontram-se as instâncias do ciclo de treinamento e, no eixo horizontal estão os erros, resultados dos valores alvo menos os valores de saída. As barras azuis representam o conjunto de dados de treinamento, as vermelhas, o conjunto de dados de teste, e a linha sólida de cor laranja indica o valor de 0 no eixo horizontal.

O histograma de erro fornece indicações de “pontos fora da curva”, que são pontos de dados em que o ajuste é significativamente pior do que a maioria dos dados, isso acontece quando a amostra tem característica não-linear. No caso da Amostra A, a barra mais alta, próxima à linha laranja, indica que maioria dos erros estão muito próximos a 0, com poucos erros longe disso.

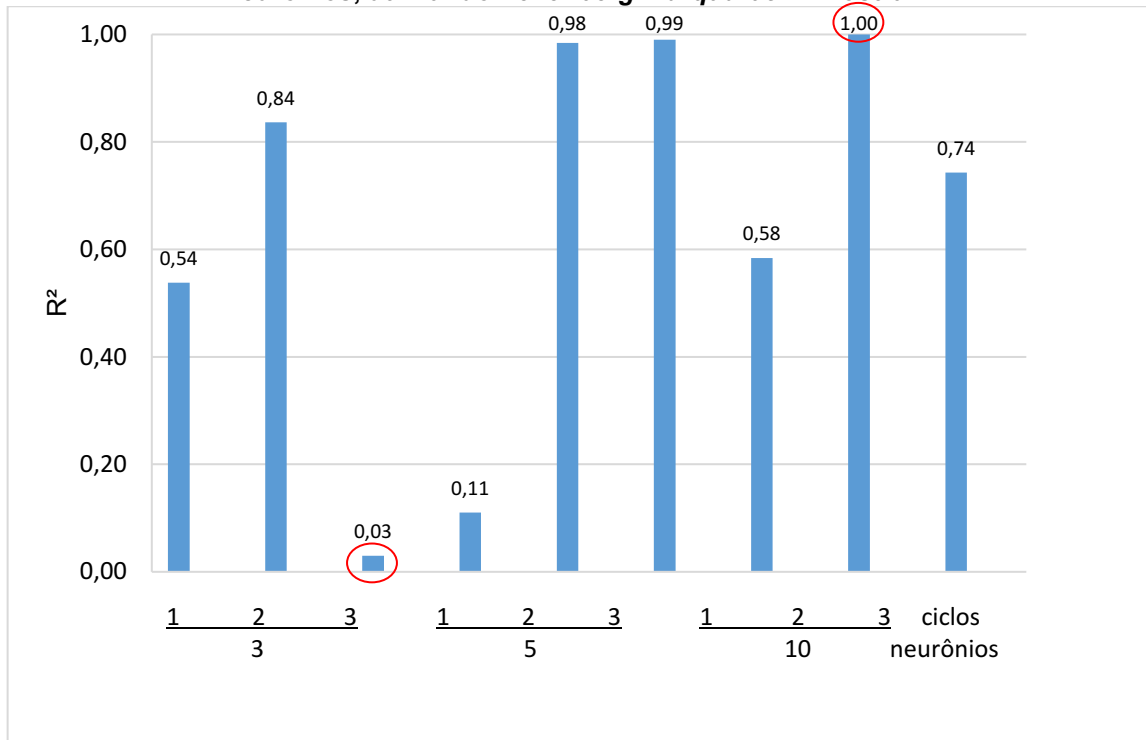
Nos treinamentos com os algoritmos *Levenberg-Marquardt* e *Scaled Conjugate Gradient*, notou-se que, quando houve mudança na amostra, do número 0 que indicava resíduo de construção, para o número 1, que indicava resíduo de demolição, a rede neural não conseguiu acertar as quantidades de resíduos na metade da amostra, que foi onde ocorreu essa mudança. Devido a isso, ao utilizar esses dois tipos de algoritmos, é recomendado fazer dois modelos separados, um para prever resíduos de construção, e um para prever resíduos de demolição. Em contrapartida, utilizando o algoritmo *Bayesian Regularization*, notou-se que a capacidade preditiva não foi alterada quando a classificação do resíduo mudou.

É importante versar sobre o tempo de treinamento, que variou de 0 segundos a 13 segundos para as diferentes configurações de redes neurais utilizadas. Treinando redes neurais com três, cinco ou dez neurônios na camada oculta, não foram verificados grandes prejuízos em tempo de processamento, uma vez que essa variação apresentada não foi significativa nesse caso analisado.

4.2 Simulação de Redes Neurais com a Amostra B

Para as simulações utilizando a Amostra B, também com duas variáveis de entrada (Classificação do resíduo e área total construída) e uma de saída (quantidade de resíduo), o algoritmo de treinamento *Levenberg-Marquardt* foi o que apresentou menor desempenho na estimativa de resíduos, com valores de R^2 variando entre 0,03 e 1,0. No Gráfico 7 são apresentados os valores de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt*.

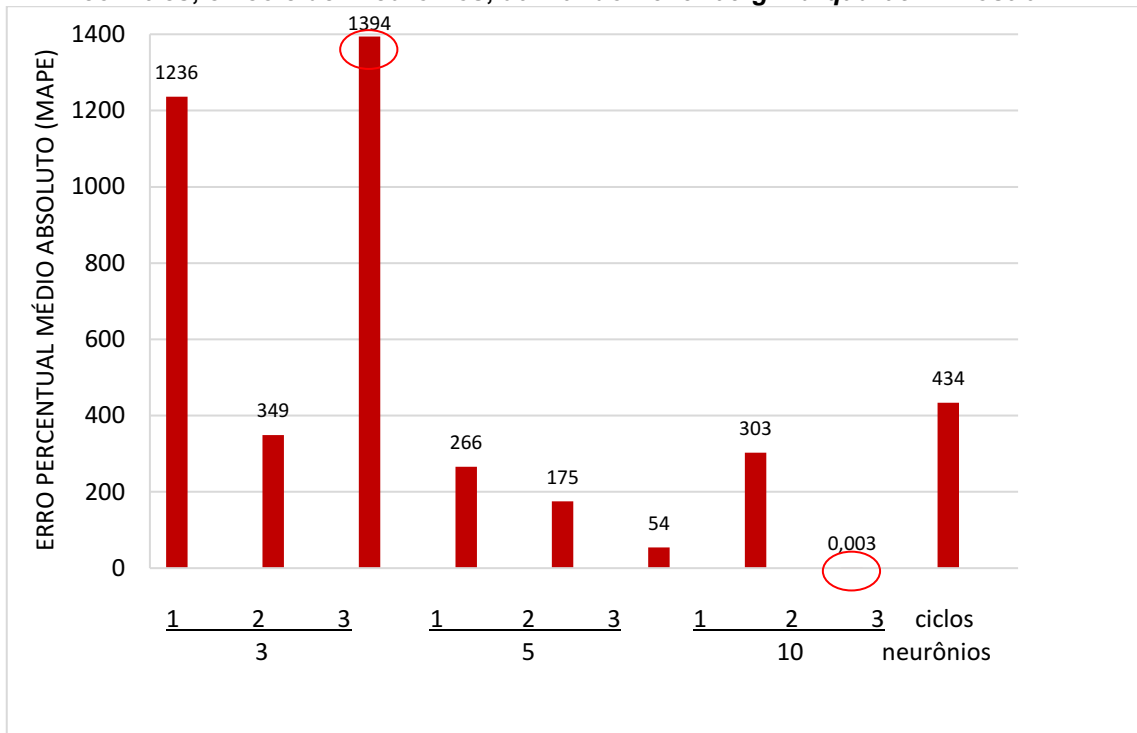
Gráfico 7 - Resultados de R^2 para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt* - Amostra B



Fonte: Autoria própria (2022)

A partir do Gráfico 7 é possível notar que no terceiro ciclo de treinamento utilizando três neurônios na camada oculta, foi onde ocorreu o menor valor de R^2 sendo igual a 0,03. Alterando a quantidade de neurônios para cinco, a rede neural melhorou seu desempenho, chegando a R^2 igual a 0,99 no terceiro ciclo de treinamento. No segundo ciclo de treinamento com dez neurônios na camada oculta, a rede neural apresentou R^2 igual a 1,0 e melhores resultados na estimativa, que podem ser verificados pelos valores de erro percentual médio absoluto, exibidos no Gráfico 8.

Gráfico 8 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Levenberg-Marquardt* - Amostra B



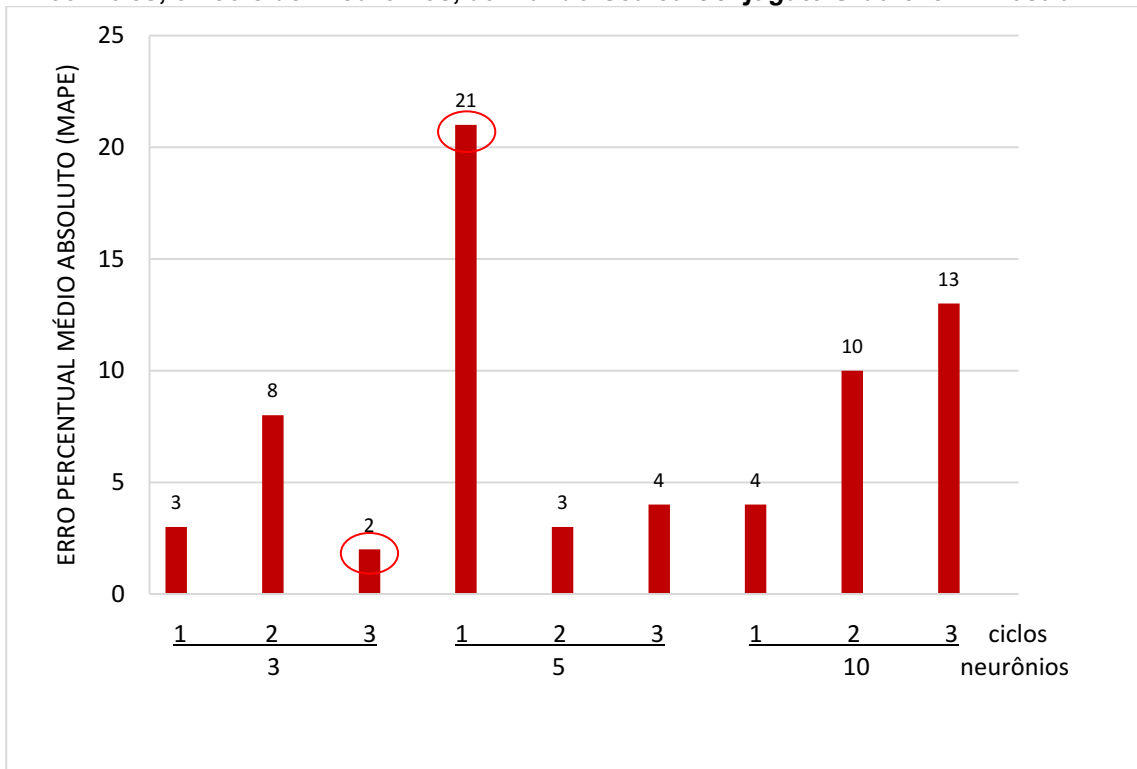
Fonte: Autoria própria (2022)

Conforme o Gráfico 8, a utilização de três neurônios na camada oculta não apresentou resultados de estimativa satisfatórios, com MAPE variando entre 349% e 1394%. Aumentando o número de neurônios para cinco, é possível notar que o MAPE diminuiu, passando de 266% no primeiro ciclo, para 54% no segundo ciclo. Contudo, foi com a configuração de rede neural com dez neurônios na camada oculta, que os resultados da estimativa foram melhores.

No segundo ciclo de treinamento, o erro percentual médio absoluto chegou a 0,003%, o que significa que as quantidades de resíduos estimadas foram muito próximas às esperadas. Além disso, notou-se que a RNA não reduziu sua capacidade preditiva quando ocorreu alteração da classificação do resíduo.

Para o algoritmo *Scaled Conjugate Gradient*, os três ciclos de treinamento, com três, cinco e dez neurônios na camada oculta apresentaram valores de R^2 entre 0,99 e 1,0. No entanto, a rede neural não conseguiu prever as quantidades de resíduos desde o início da mostra. Possivelmente, pode ter ocorrido *overfitting*, e a rede neural conseguiu prever melhor as quantidades de resíduos correspondentes aos dados do conjunto de treinamento, que correspondem a 70% da amostra, uma vez que os percentuais de erro médio absoluto foram baixos, variando entre 2% e 21%, conforme apresentado no Gráfico 9.

Gráfico 9 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Scaled Conjugate Gradient* - Amostra B

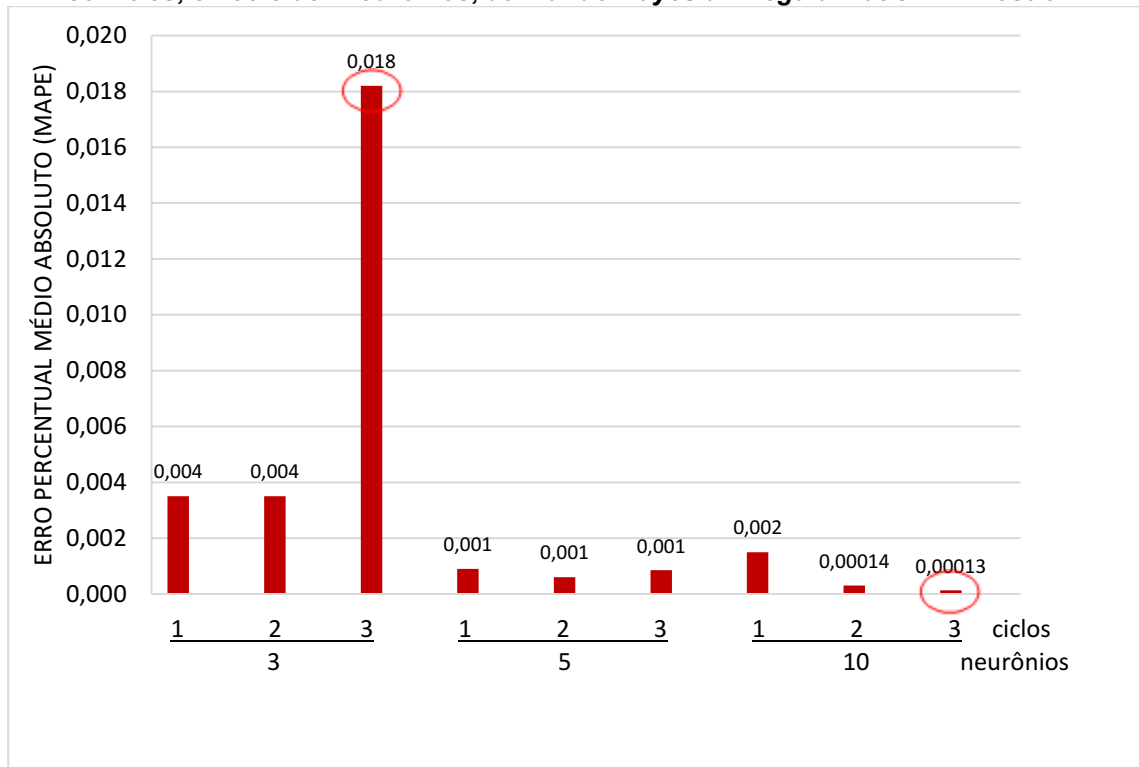


Fonte: Autoria própria (2022)

De acordo com o Gráfico 9, a configuração que apresentou melhores resultados foi a utilização de três neurônios na camada oculta e três ciclos de treinamento, com MAPE igual a 2%. Verificou-se ainda, que a capacidade preditiva da rede não melhorou utilizando mais neurônios e novos ciclos de treinamento com esse algoritmo.

Nas simulações utilizando o algoritmo *Bayesian Regularization*, os três ciclos de treinamento com três, cinco e dez neurônios na camada oculta, apresentaram valores de R^2 iguais a 1,0 e resultados preditivos próximos aos esperados. Com o aumento do número de neurônios para cinco, três ciclos de treinamento e, em seguida, para dez, com mais três ciclos de treinamento, foi possível notar que a capacidade de estimativa melhorou e o erro percentual médio absoluto diminuiu, como é possível observar no Gráfico 10.

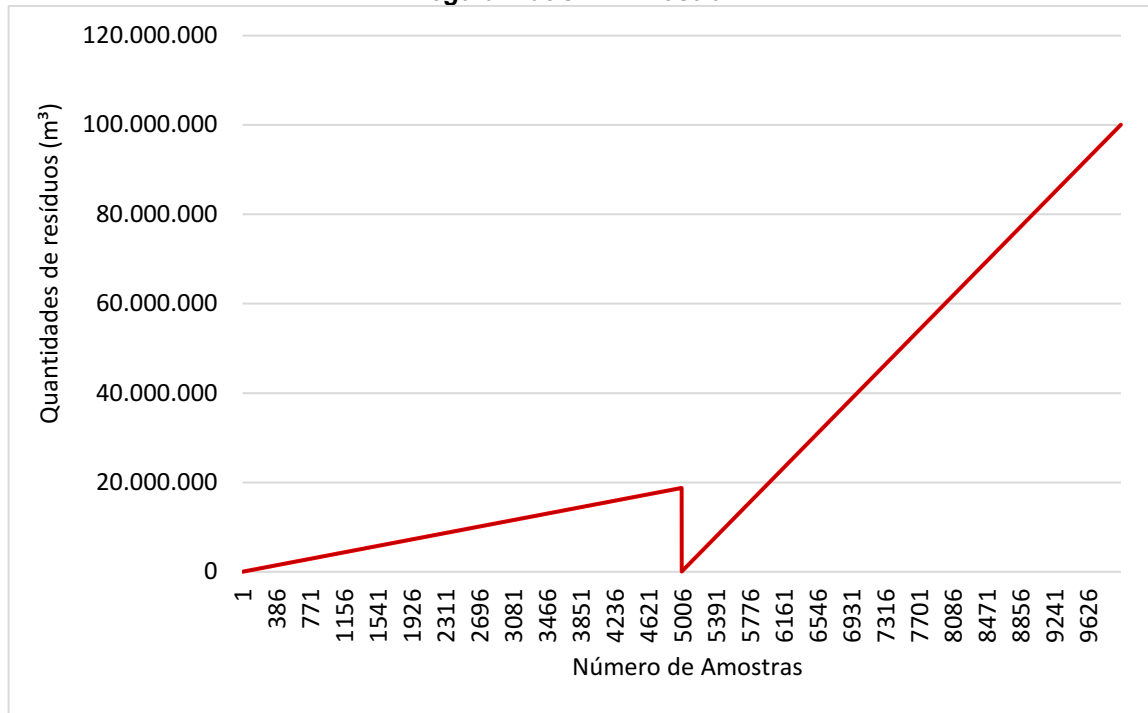
Gráfico 10 - Resultados de Erro Percentual Médio Absoluto para os três ciclos de treinamento com três, cinco e dez neurônios, utilizando *Bayesian Regularization* - Amostra B



Fonte: Autoria própria (2022)

A partir do Gráfico 10 é possível notar que no terceiro ciclo com dez neurônios foi onde se verificou a melhor capacidade de estimativa utilizando esse algoritmo, em que os resultados foram, R^2 igual a 1,0 e MAPE igual a 0,00013%. O baixo valor de MAPE significa que os resultados estimados foram muito próximos aos resultados reais. No Gráfico 11 são apresentadas as quantidades de resíduos estimadas na simulação da Amostra B, com rede neural *feed-forward*, três ciclos de treinamento e dez neurônios na camada oculta, utilizando o algoritmo de treinamento *Bayesian Regularization*.

Gráfico 11 - Quantidades de resíduos estimadas pela rede neural *feed-forward* com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization* - Amostra B

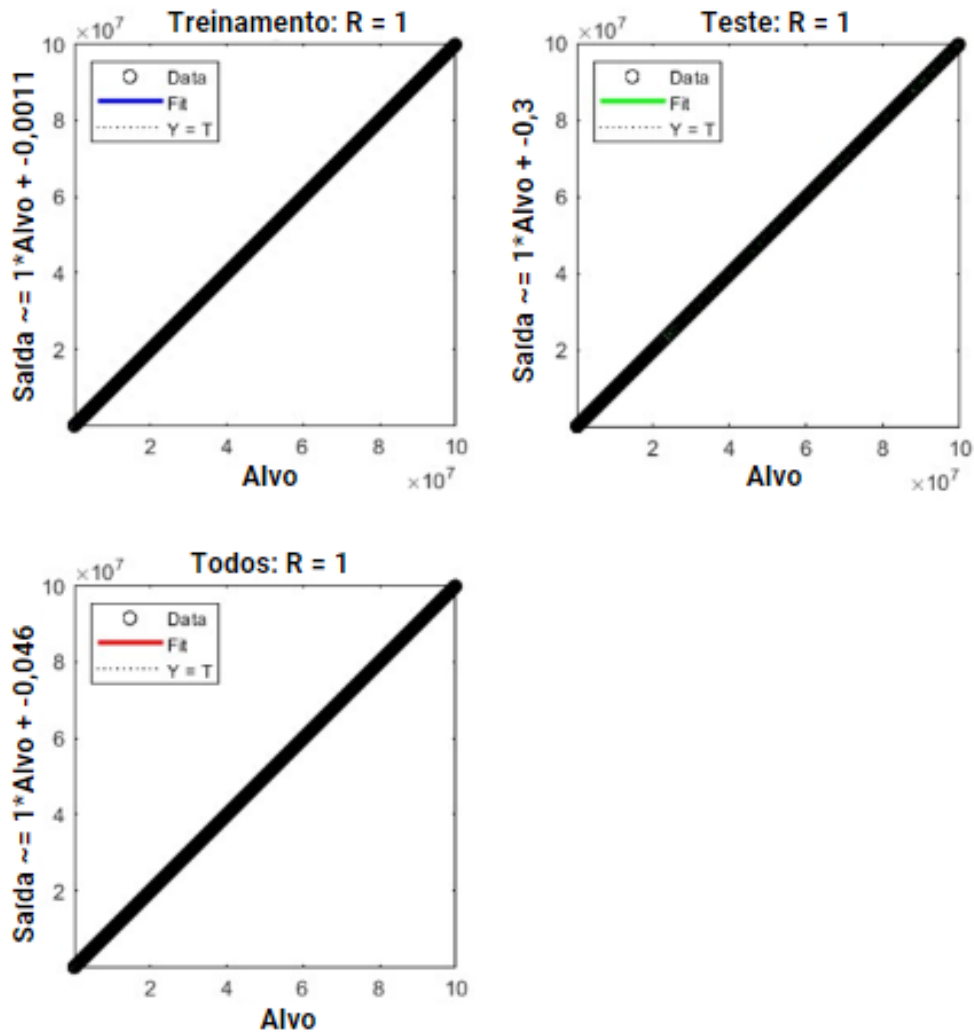


Fonte: Autoria própria (2022)

Conforme o Gráfico 11, é possível verificar que as quantidades de resíduos estimadas utilizando a rede neural, ficaram entre 11.243,60 kg a 100.039.966,66 kg, muito próximas às quantidades calculadas para a Amostra B, que foram entre 11.250,00 kg e 100.040.000,00 kg. No gráfico é possível notar ainda uma quebra na linha vermelha, que indica as quantidades de resíduos estimadas, entre as amostras 4.824 e 5.195, essa quebra indica onde a classificação do resíduo mudou de 0 (resíduo de construção) para 1 (resíduo de demolição).

Na Figura 15 são apresentados os gráficos de regressão linear com os valores de R para os conjuntos de treinamento e teste para essa configuração de rede neural, três ciclos de treinamento, dez neurônios na camada oculta e algoritmo de treinamento *Bayesian Regularization*.

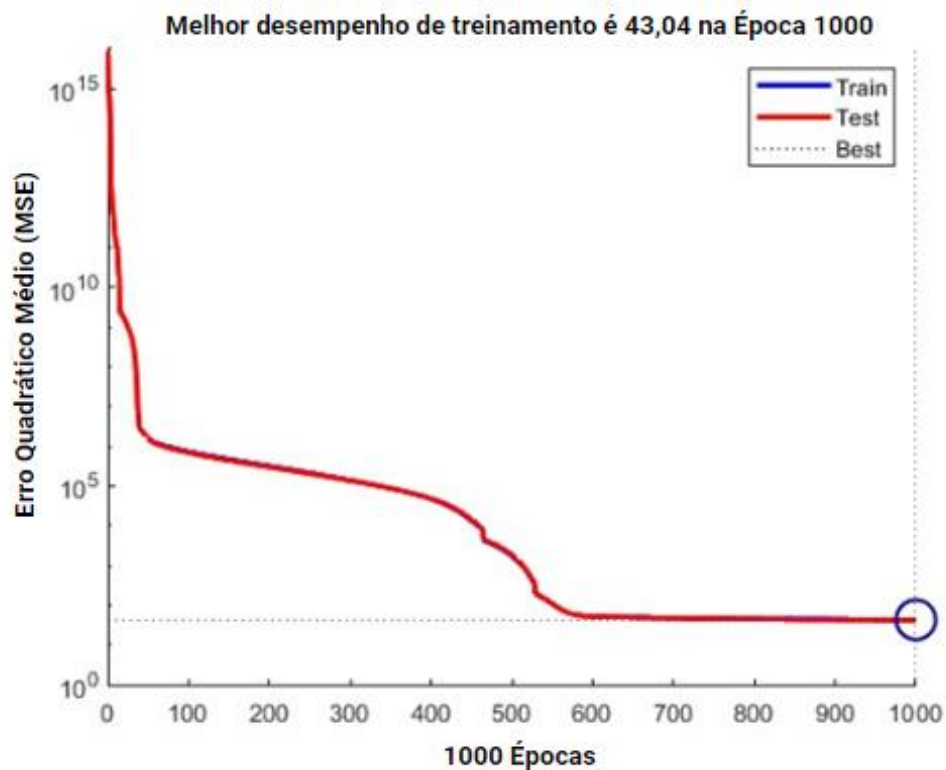
Figura 15 - Gráfico de Regressão Linear para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra B



Fonte: MATLAB® (2022)

O ajuste da linha de regressão representa o bom desempenho da rede neural para os dois conjuntos de dados. Isso fez com que as linhas indicativas ficassem alinhadas com o símbolo indicador dos dados, não estando visíveis nos gráficos. Na Figura 16 é apresentado o gráfico do Erro Quadrático Médio com a melhor performance de treinamento da rede neural com dez neurônios na camada oculta e três ciclos de treinamento, utilizando o algoritmo *Bayesian Regularization*.

Figura 16 - Gráfico do Erro Quadrático Médio para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra B

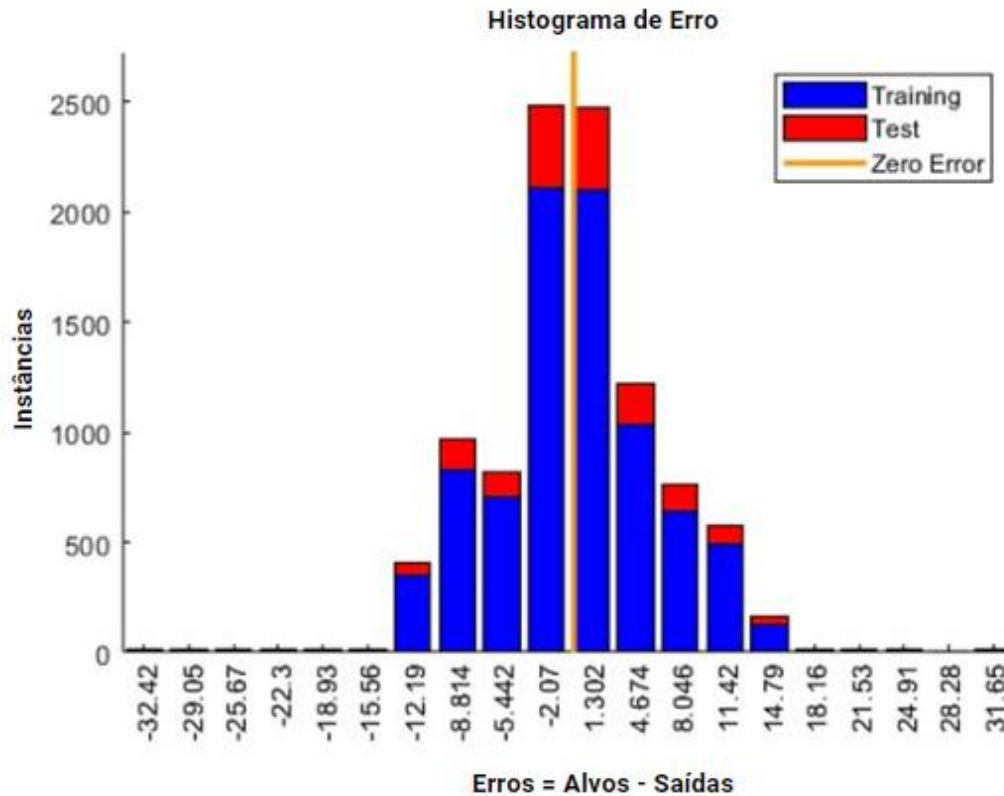


Fonte: MATLAB® (2022)

Na Figura 16 é possível verificar a performance da rede neural para a Amostra B, em termos de Erro Quadrático Médio. Para esse treinamento, nota-se que o MSE diminuiu rapidamente à medida em que a rede foi treinada, tendo o melhor desempenho nas amostras de treinamento, com o valor de MSE igual a 43,04 kg² (RMSE igual a 6,55 kg), na época 1000. Notou-se ainda que, o erro do conjunto de treinamento e do conjunto de teste têm características semelhantes, como visto na sobreposição das linhas azul e vermelha.

Na Figura 17 é apresentado o Histograma de Erro do treinamento da rede neural com dez neurônios na camada oculta e três ciclos de treinamento, utilizando o algoritmo *Bayesian Regularization*.

Figura 17 - Histograma de Erro para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra B



Fonte: MATLAB® (2022)

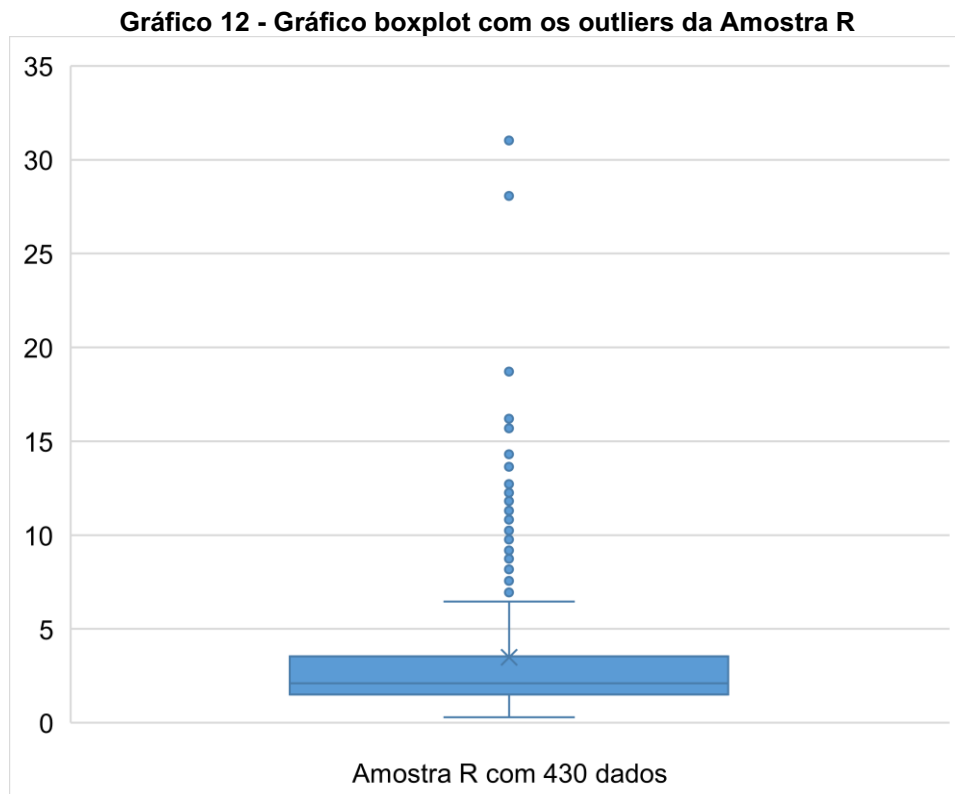
De acordo com o histograma da Figura 17, as barras mais altas, próximas da linha laranja, indicam que maioria dos erros estão muito próximos a 0, com poucos erros longe disso. Contudo, a quantidade de valores de erros negativos ficou um pouco acima do número de erros positivos, indicando que as quantidades de resíduos estimadas pela rede neural foram maiores do que as quantidades alvo.

Comparando as Amostras A e B, é possível notar que o tamanho de amostra influencia no treinamento da rede neural. A melhor configuração de treinamento da Amostra A com 5.000 dados (um ciclo de treinamento com dez neurônios, utilizando o algoritmo *Bayesian Regularization*) apresentou raiz do erro quadrático médio (RMSE) de 11,56 kg e erro percentual médio absoluto (MAPE) de 0,0003%, valores maiores do que a Amostra B (RMSE igual a 6,55 kg e MAPE igual a 0,00013%) que possui 10.000 dados. Além disso, no treinamento da Amostra A a rede neural não apresentou melhora na capacidade preditiva com o aumento do número de neurônios na camada oculta de três para cinco e dez, bem como com o aumento dos ciclos de treinamento, de um para três.

Em suma, amostras maiores tendem a apresentar melhores estimativas. Considerando que as redes neurais são modelos de aprendizagem de máquina que usam dados para aprender a relação entre as variáveis de entrada e de saída, quanto mais dados disponíveis para o modelo, mais informações ele terá para ajustar seus pesos e aprender a relação entre as variáveis. Com mais dados, o modelo tem mais exemplos para aprender e se adaptar a diferentes variações e padrões nos dados, aumentando assim a sua capacidade de generalizar para novos dados e melhorar a sua capacidade de estimativa. No entanto, o tamanho da amostra não é a única consideração na escolha do modelo ou na sua performance. Outros fatores, como a qualidade dos dados, a seleção de variáveis e a escolha de hiperparâmetros, também podem afetar a performance do modelo.

4.3 Validação do Modelo de Rede Neural com a Amostra R

A partir dos cálculos da mediana, quartis e amplitude interquartil, realizados para identificar *outliers*, 70 dados ultrapassaram o limite superior calculado e foram retirados da Amostra R. No Gráfico 12 é demonstrada a presença de *outliers* na amostra.



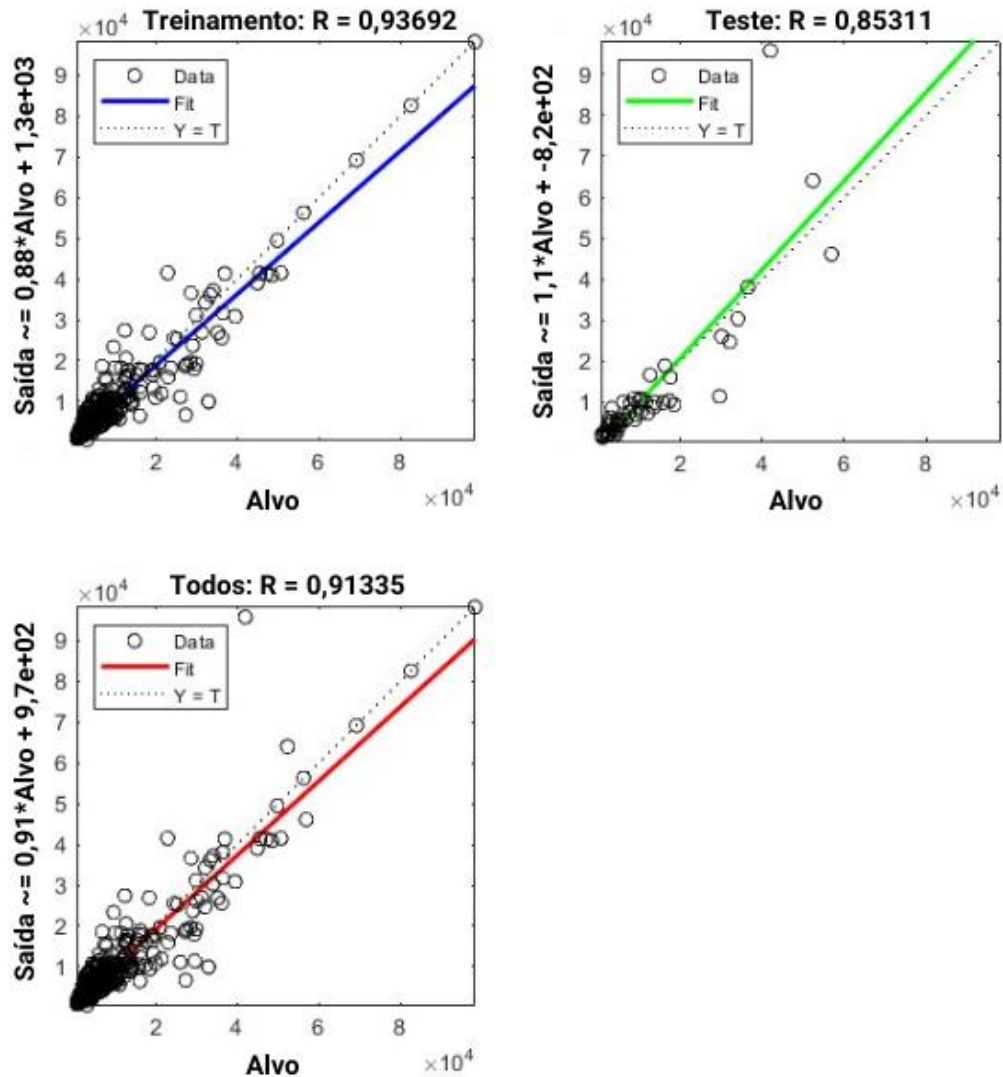
Fonte: Autoria própria (2022)

No eixo vertical do Gráfico 12 encontra-se a razão entre a área total construída e a quantidade de resíduo, essa taxa foi determinada para calcular a mediana, os quartis e a amplitude interquartil, de maneira a considerar as duas variáveis (área total construída e quantidade de resíduos), e não ocasionar a exclusão de dados que poderiam não ser *outliers*, com base em apenas uma das variáveis. Não foram verificados dados que ficaram com a taxa abaixo do limite inferior (LI = -0,94). No entanto, 70 dados apresentaram taxa acima do limite superior (LS = 5,16), esses dados considerados *outliers* são os pontos em azul que ficaram fora do gráfico *boxplot*. Devido a muitas taxas terem dado o mesmo resultado, alguns marcadores (pontos em azul) ficaram sobrepostos, contudo, após a retirada de *outliers* a Amostra R passou a ter 360 dados.

Após o tratamento dos dados, a Amostra R foi simulada a partir da melhor configuração de rede neural apresentada, com base nos parâmetros de avaliação R^2 (1,0), RMSE (6,55 kg) e MAPE (0,00013%), obtidos no treinamento da Amostra B, com a rede neural *feed-forward* com três ciclos de treinamento, dez neurônios na camada oculta e algoritmo de treinamento *Bayesian Regularization*. Os resultados apresentaram R^2 igual a 0,83, RMSE igual a 4.337,69 m³ e MAPE de 38,35%.

Na sequência, serão apresentadas as figuras contendo os gráficos de regressão linear, gráfico de erro quadrático médio e histograma de erro, alguns dos parâmetros de avaliação dessa configuração de rede neural. Na Figura 18 são apresentados os gráficos de regressão linear com os valores de R para os conjuntos de treinamento e teste para a melhor configuração de rede neural identificada, com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization*.

Figura 18 - Gráfico de Regressão Linear para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra R



Fonte: MATLAB® (2022)

Os gráficos de regressão linear da Figura 18 apresentam a relação das quantidades estimadas pela rede e as quantidades alvo. Os dois gráficos apresentados no topo representam os dados de treinamento e teste, o gráfico abaixo apresenta a plotagem dos dois gráficos de cima. O valor de R^2 para o conjunto de dados de treinamento e teste foi igual a 0,83, valor próximo ao obtido por Cha et al. (2022) (R^2 de 0,90) utilizando redes neurais com uma amostra de 160 dados, e próximo ao obtido por Coskuner (2021) (R^2 de 0,91), utilizando redes neurais na previsão de resíduos de construção.

Considerando que Coskuner (2021) utilizou amostra com 20 dados, duas variáveis de entrada e 180 simulações, pode-se dizer que a amostra R, com 360 dados, teve um bom desempenho com apenas três ciclos de treinamento. Ainda, o

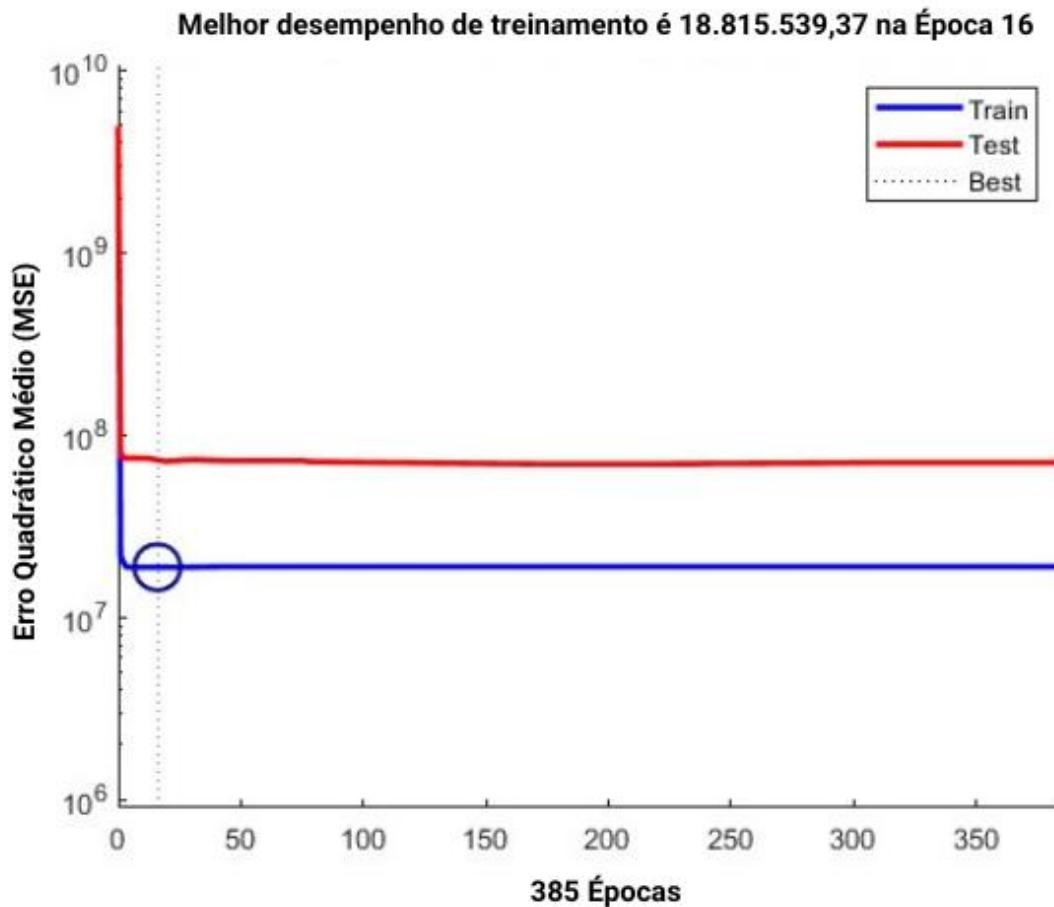
valor de R^2 (0,83) obtido nessa pesquisa, ficou acima do valor obtido por Cha; Hong e Kim (2023) que foi de 0,458 para a utilização de redes neurais em uma amostra com 782 dados de obras de demolição e seis variáveis de entrada, e superior ao valor obtido por Soni et al. (2019), de R^2 igual a 0,72.

É possível verificar que a rede neural apresentou melhor desempenho no conjunto de dados de treinamento, uma vez que os dados estão melhor distribuídos ao longo da linha de regressão e o valor de R (0,94) foi maior, comparado ao R (0,85) do conjunto de teste. Assim como observado por Nagalli (2021a), os dados do conjunto de treinamento apresentaram melhores resultados, com valores de R^2 variando entre 0,94 e 0,96, enquanto para os dados do conjunto de teste, os valores de R^2 variaram entre 0,75 e 0,91.

Na prática, isso significa que o modelo pode ser muito bom em simular os dados que ele já conhece, mas pode falhar em prever dados novos e não vistos antes. Portanto, é importante monitorar o desempenho do modelo nos dados de teste para avaliar sua capacidade de generalização e previsão para novos dados. Nesse caso, 70% dos dados da amostra pertencem ao conjunto de treinamento, é possível que diminuindo a porcentagem de dados de treinamento e aumentando a porcentagem de dados de teste, a rede neural melhore sua generalização para os dados de teste, o que conseqüentemente, irá melhorar seu desempenho preditivo nesse conjunto de dados.

Na Figura 19 é apresentado o gráfico do erro quadrático médio (MSE) em que seu valor indica a magnitude média do erro de estimativa em relação aos valores reais. Durante o treinamento de uma rede neural, o MSE é calculado para cada época de treinamento e através do gráfico representativo é possível monitorar a evolução do treinamento, com o melhor valor de desempenho no treinamento da rede neural com dez neurônios e três ciclos de treinamento, utilizando o algoritmo *Bayesian Regularization*.

Figura 19 - Gráfico do Erro Quadrático Médio para o terceiro ciclo de treinamento da rede neural com dez neurônios utilizando o algoritmo *Bayesian Regularization* - Amostra R



Fonte: MATLAB® (2022)

Na Figura 19 é possível verificar o desempenho da rede neural para a Amostra R, em termos de Erro Quadrático Médio. O ponto de melhor performance em um gráfico de MSE representa o ponto em que o modelo atingiu sua melhor capacidade de estimativa durante o treinamento, para esse caso, foi na época 16.

Para essa simulação, nota-se que o MSE não apresentou um valor baixo, todavia, é importante ressaltar que ele é uma medida de dispersão dos erros, o que pode ser difícil de interpretar em termos absolutos, porque o seu valor depende da escala dos dados e do tamanho da amostra. Portanto, é mais comum usar o MSE em comparação com outros modelos ou para avaliar a melhoria da performance do modelo em amostras com tamanhos diferentes. Por exemplo, a raiz do erro quadrático médio (RMSE) de Cha et al. (2022) foi de 100,80 kg, nesse estudo foi de 4.337,69 m³, comparando com a utilização do algoritmo máquina de vetor suporte (SVM) no estudo de Cha; Hong e Kim (2023), o valor de RMSE foi de 1.020,30, para uma amostra de 782 dados, bem como no estudo de Abbasi et. al (2013), utilizando máquina de vetor suporte, o valor de RMSE foi de 2.070.

Nesse caso, possivelmente, o tamanho da amostra de 360 dados influenciou nas simulações, visto que Akanbi et al. (2020) realizaram simulações utilizando RNA's *feed-forward* em uma amostra com 2.228 dados de obras de construção e obtiveram resultados de erro absoluto médio entre 17,93 m³ e 19,04 m³, bem como nas Amostras fictícias A e B já apresentadas, os valores de MSE diminuíram ao longo do treinamento e foram menores, como apresentado na Tabela 2.

Tabela 2 - Valores de RMSE e quantidade de dados das Amostras A, B e R

Amostra	Quantidade de dados	Valor de RMSE
A	5.000	11,56 kg
B	10.000	6,55 kg
R	360	4.337,67 m ³

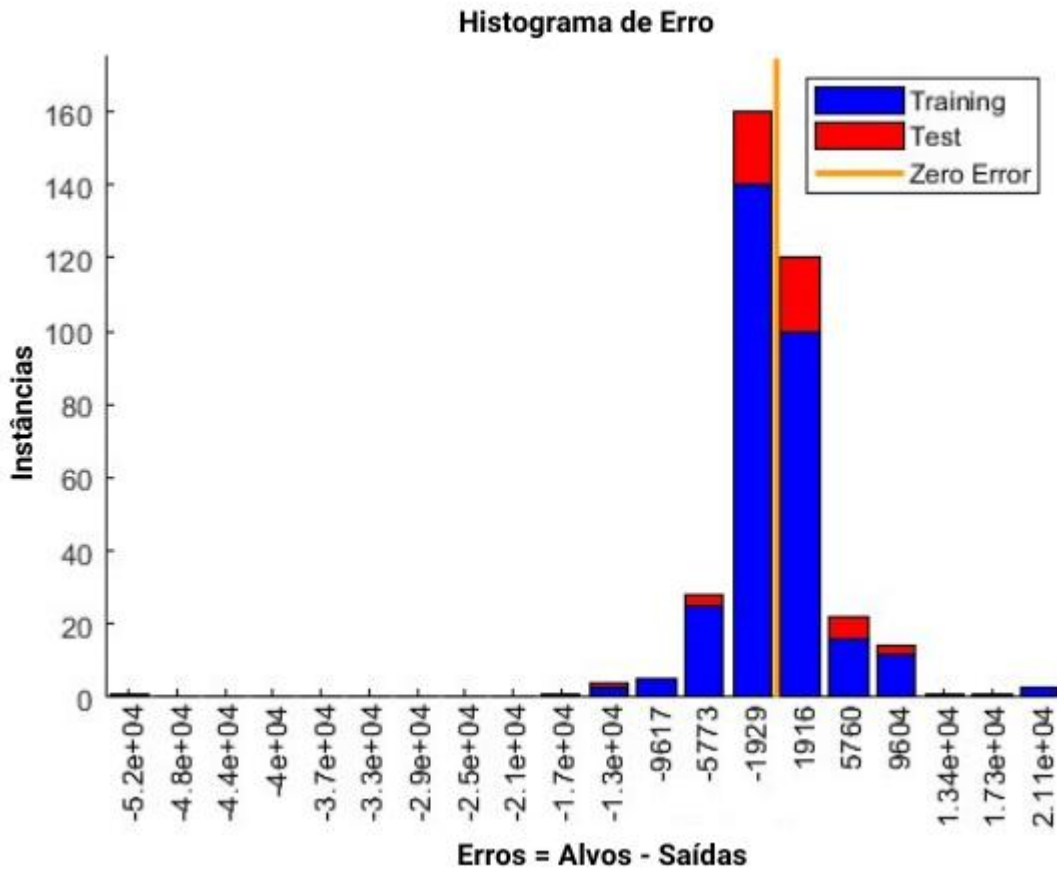
Fonte: Aatoria própria (2022)

A fórmula para calcular o MSE inclui a diferença entre o valor estimado e o valor real, elevada ao quadrado. Portanto, quanto maior a amostra, mais dados serão considerados no cálculo do MSE, o que pode reduzir a influência de valores atípicos ou erros aleatórios nos dados. Por outro lado, em amostras pequenas, a variância do MSE é maior, o que pode fazer com que o modelo seja mais sensível a valores atípicos ou a erros aleatórios nos dados. Isso pode levar a uma estimativa menos precisa do desempenho do modelo.

Assim sendo, é importante considerar o tamanho da amostra ao avaliar o desempenho de um modelo de estimativa. Amostras maiores podem fornecer estimativa mais precisa do desempenho do modelo e ajudar a reduzir a incerteza em torno das estimativas, mas também é importante garantir que a amostra seja representativa e suficiente para cobrir todas as variações importantes nos dados. Por isso, recomenda-se que o tamanho da amostra a ser simulada utilizando redes neurais seja maior do que 5.000 dados, visto que de 5.000 para 10.000 dados o MSE diminuiu consideravelmente.

Na Figura 20 é apresentado o Histograma de Erro da rede neural com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization*.

Figura 20 – Histograma de Erro da rede neural com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization* – Amostra R



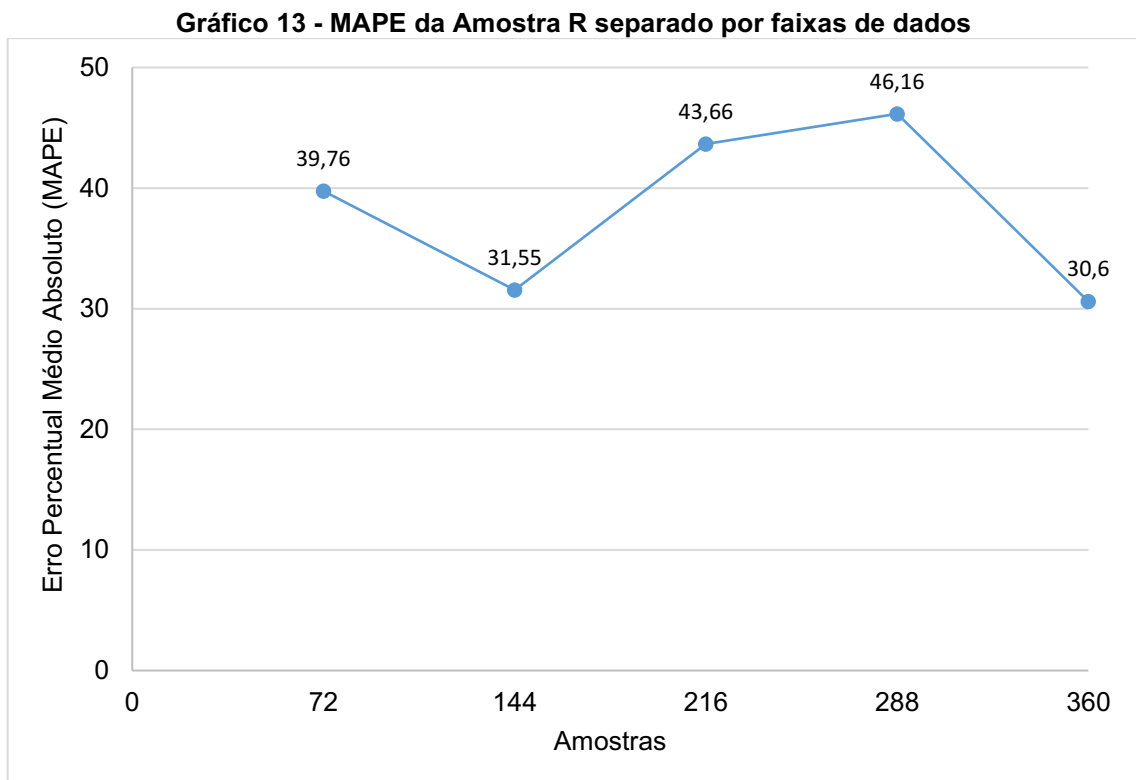
Fonte: MATLAB® (2022)

Conforme o histograma da Figura 20, na altura das barras azuis é possível verificar que o conjunto de treinamento apresentou mais valores de erros positivos e próximos de zero (linha laranja), confirmando melhor desempenho com relação ao conjunto de teste. Essa configuração de rede neural não apresentou resultados muito próximos de zero quando comparada com os histogramas das Amostras A e B, isso pode se justificar pelo fato de que a Amostra R contém valores variados e não segue uma escala crescente como as amostras com dados fictícios.

O erro percentual médio absoluto da Amostra R foi igual a 38,35%, o que indica que, em média, a diferença entre os valores preditos e os reais, equivale a 38,35%, bem como as estimativas têm pouca imprecisão e podem ser confiáveis. A melhor configuração de rede neural proposta por Nagalli (2021a) também apresentou como algoritmo de treinamento o *Bayesian Regularization*, porém, foram utilizados dois neurônios e dois ciclos de treinamento, a RNA apresentou MAPE igual a 56,7%.

O MAPE é uma medida relativa e pode variar dependendo do escopo e da natureza do conjunto de dados. Além disso, pode ser influenciado por valores extremos ou valores muito pequenos nos dados. Nesse caso, aumentar o espaço amostral de 330 para 360 dados, bem como realizar o tratamento dos dados retirando os *outliers*, foram fatores que contribuíram com os resultados preditivos da rede neural. Ainda, aumentar o número de neurônios e de ciclos de treinamento, também foi favorável, visto que o MAPE teve redução de aproximadamente 20% de um estudo para outro.

No Gráfico 13 é apresentado o erro percentual médio absoluto, dividido em cinco faixas, cada uma com 72 dados da amostra. Essa divisão permite visualizar o desempenho da capacidade preditiva da rede neural ao longo da amostra.



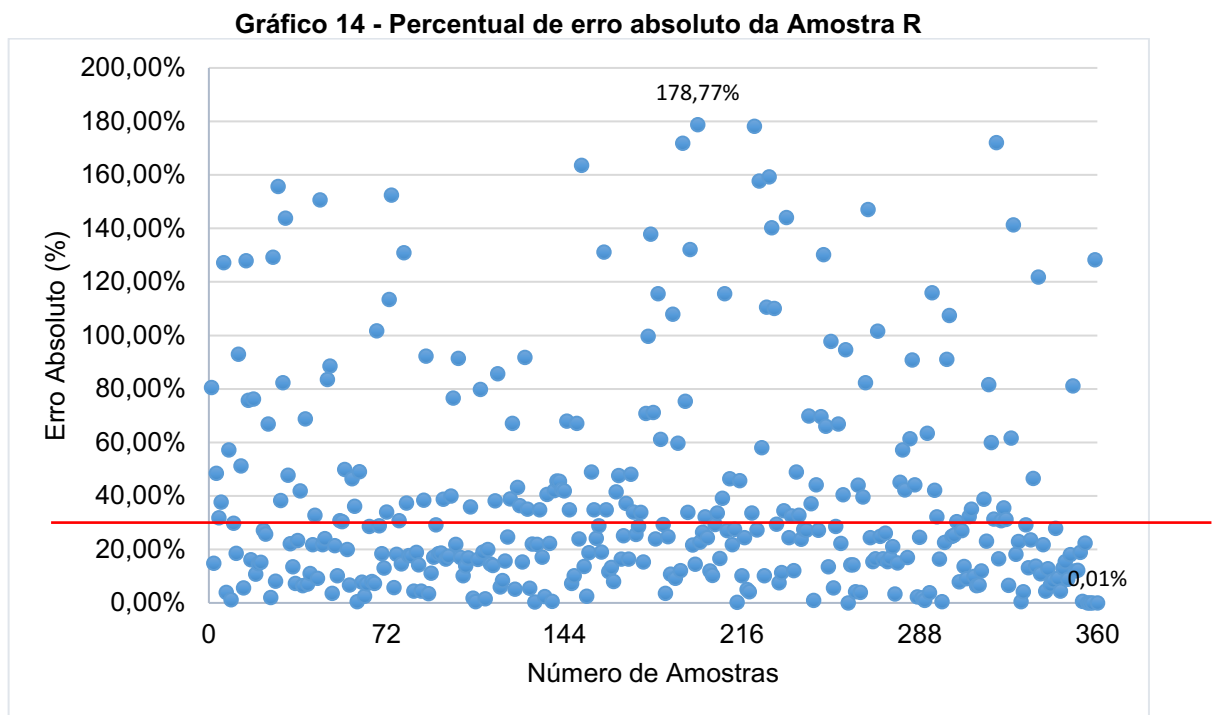
Fonte: Autoria própria (2022)

Conforme o Gráfico 13, da amostra de número 72 até a amostra de número 144, o valor de MAPE reduziu de 39,76% para 31,55%, da amostra 144 até a amostra 216 o MAPE aumentou para 43,66% e entre as amostras 216 e 288, foi a faixa onde se verificou o maior erro percentual médio absoluto, igual a 46,16%, o que significa que nessa faixa da amostra os resultados estimados tiveram valores mais distantes dos resultados esperados. Na última faixa da amostra, entre os dados

de número 288 até 360, foi onde se verificou o menor percentual de MAPE, igual a 30,60%, significando que os valores preditos ficaram mais próximos aos reais, comparando com a faixa de dados anterior.

Além do MAPE, é interessante apresentar os percentuais de erro absoluto, que representam a diferença absoluta entre os valores estimados e os reais, em porcentagem. A diferença entre o MAPE e o percentual de erro absoluto é que o MAPE é calculado como a média dos erros absolutos divididos pelo número total e amostras, multiplicado por 100, ou seja, ele fornece uma medida de erro em termos percentuais. Já o erro absoluto é simplesmente a diferença entre o valor real e o estimado, dividida pelo valor estimado, multiplicada por 100.

No Gráfico 14 é apresentado o percentual de erro absoluto para os 360 dados, dividindo a amostra em cinco faixas com 72 dados cada uma. Os percentuais de erro absoluto variaram entre 0,01% e 178,77%.

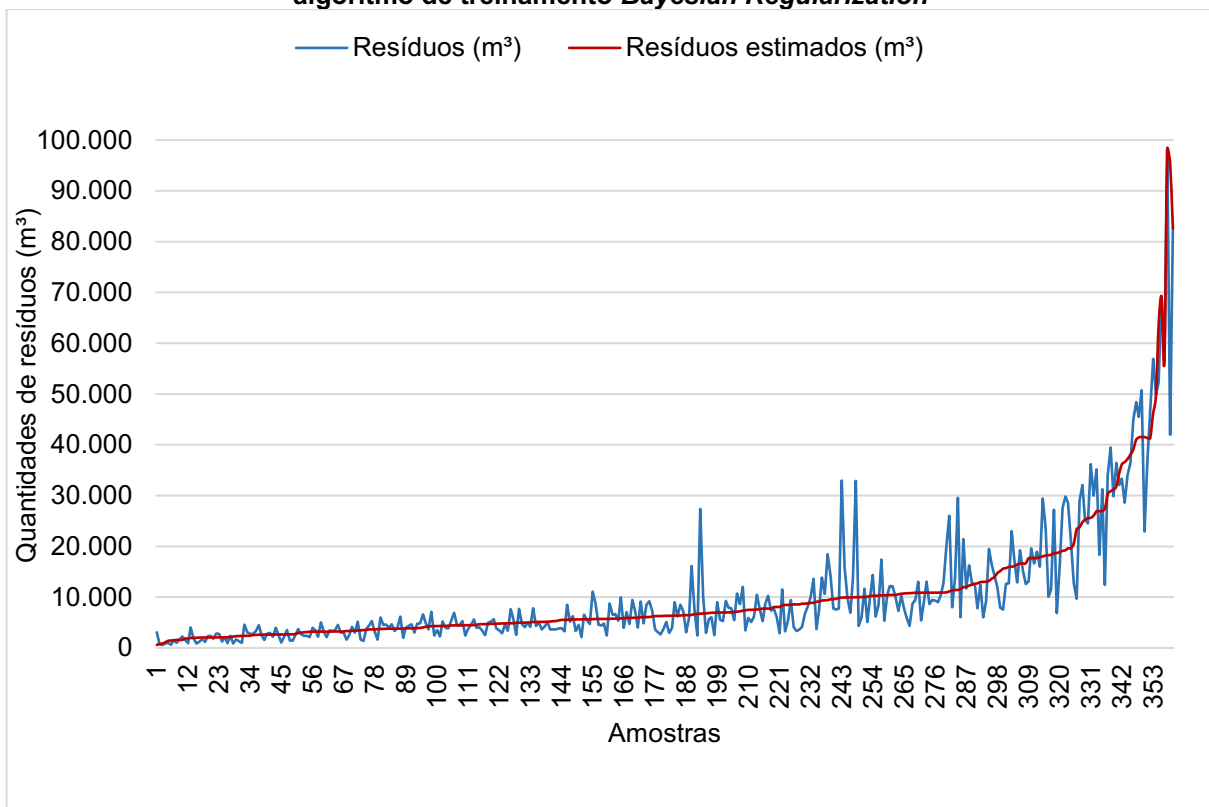


A partir do Gráfico 14 é possível notar que a maior parte dos dados da amostra apresentou percentuais de erro absoluto até 30%. A diferença entre a quantidade de resíduos calculada pelo modelo de rede neural e a informada pelas construtoras foi inferior ou igual a 30% em 210 casos (58,33% da amostra). Em 69 dados (19,17% da amostra), a diferença foi menor ou igual a 10%.

Comparando com o estudo de Kern et al. (2015) que utilizaram modelo de regressão linear para prever as quantidades de resíduos de 18 edifícios, e obtiveram uma diferença entre as quantidades reais e as estimadas pelo modelo, menor do que 30% em cerca de 83% da amostra, e inferior a 5% em 5 edifícios, pode-se afirmar que o modelo de rede neural teve desempenho excelente em relação ao critério de diferença entre as quantidades reais e estimadas e ao tamanho da amostra. No entanto, essa análise fica restrita a essa amostra, sendo importante considerar que a quantidade de dados e a metodologia empregada nos dois estudos são completamente diferentes.

As quantidades de resíduos reais e as quantidades estimadas utilizando a rede neural *feed-forward* com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization*, são apresentadas no Gráfico 15, que em seu eixo vertical exibe as quantidades de resíduos em m³, e em seu eixo horizontal, as 360 amostras.

Gráfico 15 - Quantidades de resíduos reais e quantidades de resíduos estimadas pela rede neural *feed-forward* com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization*



Fonte: Autoria própria (2022)

Analisando as linhas azul (quantidades reais) e vermelha (quantidades estimadas) do Gráfico 15, é possível notar que as quantidades reais apresentam grandes flutuações, indicando muitas variações nas quantidades de resíduos. Em contrapartida, visualizando a linha que apresenta as quantidades estimadas, é possível que o modelo de estimativa tenha capturado uma tendência geral dos dados, uma vez que as estimativas são crescentes, e a Amostra R foi simulada com os dados dispostos de maneira aleatória. Ainda, é possível que tenha ocorrido *overfitting*, o modelo pode estar sobreajustado aos dados, criando uma previsão que se encaixa perfeitamente nos valores reais, mas não é generalizável para outras situações. Isso pode levar a uma previsão crescente mesmo quando os dados reais são aleatórios.

As quantidades reais de resíduos variaram entre 572 m³ e 98.280 m³ e as quantidades de resíduos estimadas, variaram entre 600,10 m³ e 98.218,92 m³. Analisando as quantidades reais e estimadas, disponíveis no Apêndice A, pode-se afirmar que em casos onde a variável de entrada teve amplitude maior do que a variável de saída, a rede neural apresentou dificuldade em ajustar seus pesos adequadamente para prever corretamente a saída. Isso pode ocorrer devido à forma como as redes neurais são projetadas para realizar suas previsões.

Em muitos casos, a rede neural é treinada para aprender uma relação entre as variáveis de entrada e de saída por meio de um processo de ajuste de pesos. Por exemplo, se a variável de entrada variar de 0 a 1.000, enquanto a variável de saída variar apenas de 0 a 10, a rede neural pode acabar dando muito mais peso à variável de entrada do que à variável de saída. Isso pode levar a previsões excessivamente influenciadas pela variável de entrada, resultando em erros significativos quando a saída é prevista. Uma maneira de lidar com esse problema, além de fazer a retirada dos *outliers*, é normalizar as variáveis de entrada e de saída antes do treinamento da rede neural, de modo que elas tenham amplitudes semelhantes. Isso pode ajudar a evitar que a rede neural dê mais peso a uma variável em detrimento de outra e, assim, melhorar a precisão das previsões.

Nesse estudo, o melhor resultado da previsão foi verificado para a área 138.824,46 m² em que a quantidade de resíduos real foi igual a 82.597,50 m³ e a quantidade estimada, igual a 82.609,38 m³, nesse caso, o percentual de erro absoluto foi de 0,01%. O pior resultado da estimativa foi verificado na área 12.022,00 m², em que a quantidade real de resíduos foi de 2.487,00 m³, nesse caso, a

quantidade estimada foi de 6.933,03 m³, apresentando um percentual de erro de 178,77%. No estudo de Kern et al. (2015), o menor percentual de diferença entre a quantidade real de resíduos e a quantidade calculada pelo modelo de regressão linear, foi de 1,53%, para uma área de 696,57 m² em que a quantidade de resíduos real foi igual a 2.164 m³ e a quantidade estimada, igual a 2.197,10 m³. O pior resultado da previsão foi para área de 506 m², em que a quantidade real de resíduos foi de 712 m³ e a quantidade prevista foi de 1.164,10 m³, apresentando percentual de erro de 63,50%. Pode-se verificar que o modelo de rede neural utilizado na amostra apresentou maior variação nos resultados de previsão, com pior resultado de erro absoluto de 178,77%, enquanto o modelo de regressão linear do estudo de Kern et al. (2015) apresentou pior resultado de 63,50%. Porém, ambos os modelos apresentaram a maior parte dos resultados de erro absoluto bastante baixos, indicando boa capacidade de estimativa dos modelos para as amostras analisadas.

Em vista disso, afirma-se que a rede neural estabeleceu uma tendência entre a área e a quantidade de resíduo, e quando esses dois dados possuem muita diferença numérica, a rede neural não consegue ter um bom desempenho preditivo. Vale ressaltar que a qualidade das estimativas depende da qualidade dos dados, tamanho da amostra, da escolha das variáveis de entrada, do algoritmo de previsão, bem como do pré-processamento dos dados, entre outros fatores.

Para esse estudo, a configuração ideal de RNA é a rede neural *feed-forward*, com dez neurônios na camada oculta, três ciclos de treinamento e algoritmo de treinamento *Bayesian Regularization*, uma vez que essa configuração apresentou os melhores resultados para todas as métricas analisadas (R², RMSE e MAPE), nas simulações com a Amostra B (10.000 dados). Da mesma forma, utilizando essa configuração de rede neural na validação do modelo com a Amostra R (360 dados), a rede neural comprovou sua eficiência, sendo capaz de estimar as quantidades de resíduos com precisão em mais de 60% dos casos.

É importante ressaltar que o processo de estimativa da quantidade de resíduos gerados pode ser complexo, pois há grande variação na geração e nas taxas de geração desses resíduos. Além disso, não necessariamente existe uma relação direta e plena entre as variáveis de entrada (área total construída) utilizadas e as variáveis de saída (quantidade de resíduos), já que esta última é dependente de outros fatores, conforme descrito na literatura.

O modelo proposto nesse estudo possui configurações semelhantes ao proposto por Nagalli (2021a) que também utilizou o algoritmo *Bayesian Regularization* e dez neurônios na camada oculta. Comparando esses estudos com a pesquisa de Lu (2021) é possível verificar que o algoritmo de treinamento *Bayesian Regularization* funciona melhor com amostras maiores, uma vez que Lu (2021) utilizou o algoritmo *Levenberg-Marquardt* em uma amostra com 43 conjuntos de dados e o melhor resultado obtido para R^2 (0,92) foi alcançado com quinze neurônios na camada oculta e trinta e cinco ciclos de treinamento. A utilização do algoritmo *Bayesian Regularization* proporciona bons resultados preditivos com menor número de neurônios e menos ciclos de treinamento, o que também contribui para que não ocorra *overtraining*.

Com relação aos parâmetros utilizados para avaliar os modelos de redes neurais, o valor de R^2 igual a 0,83 nessa pesquisa, foi um valor considerado satisfatório para modelos de estimativa, ficando acima do valor obtido por Hu (2021), R^2 igual a 0,75, em sua pesquisa com redes neurais utilizando dados de 206 obras, igual ao valor obtido por Abbasi e El Hanandeh (2016), R^2 de 0,83 para dados do conjunto de treinamento e próximo ao obtido por Cha et al. (2022), R^2 igual a 0,90 para treinamento com redes neurais de retropropagação.

Apesar do valor de R^2 ter sido satisfatório, indicando que a rede neural pode explicar 83% das variações nas quantidades de resíduos, o valor de RMSE igual a 4.337,69 m³ ficou acima dos valores obtidos por Abbasi et al. (2013) e Abbasi e El Hanandeh (2016) utilizando o algoritmo máquina de vetor suporte (RMSE 2.070 e RMSE 300,70 kg, respectivamente), e redes neurais (RMSE 498,43 kg).

Nesse estudo, o valor de MAPE foi de 38,35%, o que significa que o modelo de redes neurais proposto é capaz de estimar a quantidade de resíduos com precisão em mais de 60% dos casos. Esse valor foi menor do que o valor obtido por Nagalli (2021) de 56,7%, porém maior do que os valores alcançados por Abbasi e El Hanandeh (2016), de 0,07% para dados do conjunto de treinamento e Hu (2021), de 8,22% para simulações com redes neurais, 5,37% para simulações com máquina de vetor suporte e 10,31% para modelo de regressão linear. O valor de MAPE foi aceitável comparado aos demais estudos, indicando que o modelo apresenta pouca variação percentual em relação aos dados reais.

A partir dessa pesquisa, verifica-se que a melhora na capacidade preditiva pode ser alcançada aumentando o tamanho da amostra, para no mínimo 5.000

dados. Embora os resultados obtidos sejam restritos às amostras analisadas, as RNA's demonstraram excelente potencial de aplicação na estimativa de resíduos. Além disso, a perspectiva é que com a ampliação do banco de dados, até mesmo considerando diferentes variáveis de entrada e mantendo o tratamento dos dados da amostra, a fim de retirar *outliers*, a rede neural melhore ainda mais sua capacidade de estimar quantidades de resíduos.

5 CONCLUSÃO

Com relação à investigação de redes neurais para estimar a geração de resíduos da construção civil, os resultados mostram a eficiência das RNA's. Diferentes configurações de redes neurais foram testadas com duas amostras A e B com 5.000 e 10.000 dados, respectivamente. Cada amostra foi simulada três vezes para três quantidades de neurônios na camada oculta (três, cinco e dez) e para os três algoritmos de treinamento disponíveis no *software* MATLAB® (*Levenberg-Marquardt*, *Bayesian Regularization* e *Scaled Conjugate Gradient*).

A partir das simulações, foi possível verificar que a configuração da rede neural que apresentou o melhor resultado preditivo para a Amostra A foi a rede neural *feed-forward* com duas variáveis de entrada (classificação do resíduo e área total construída), dez neurônios na camada oculta, uma variável de saída (quantidade de resíduos) e um ciclo de treinamento com o algoritmo de treinamento *Bayesian Regularization*, apresentando valores de R^2 igual a 1,0, RMSE igual a 11,56 kg e MAPE igual a 0,0003%. Para a Amostra B a configuração de rede neural que apresentou o melhor desempenho foi a rede neural *feed-forward* com duas variáveis de entrada (classificação do resíduo e área total construída), dez neurônios na camada oculta, uma variável de saída (quantidade de resíduos) e três ciclos de treinamento com o algoritmo de treinamento *Bayesian Regularization*, apresentando valores de R^2 igual a 1,0, RMSE igual a 6,55 kg e MAPE igual a 0,00013%.

Com a configuração de rede neural que apresentou os melhores resultados definida, o modelo de rede neural foi validado, utilizando essa configuração de RNA, na Amostra R, composta por 360 dados. Na validação do modelo com a Amostra R, o resultado de R^2 foi igual a 0,83, MAPE igual a 38,35% e RMSE de 4.337,69 m³. O resultado de R^2 0,83 indica que a variação na quantidade de resíduos pode ser explicada em 83% pela rede neural, o que é uma boa medida de ajuste do modelo. Isso sugere que as variáveis de entrada incluídas no modelo têm uma relação forte com as variáveis de saída e que o modelo é capaz de explicar grande proporção da variação observada. Além disso, os valores estimados foram próximos aos valores esperados, uma vez que o resultado de MAPE aponta que em média, os valores estimados apresentam diferença absoluta de 38,35% em relação aos valores reais.

Ainda, para a Amostra R, a rede neural teve o maior percentual de erro verificado entre as amostras 216 e 288, chegando a um MAPE de 46,16% nessa

faixa da amostra, já o menor percentual de erro verificado entre as amostras 288 e 360, onde a amplitude entre as áreas totais e as quantidades de resíduos não foi significativa, nessa faixa o MAPE foi de 30,60%. Além da rede neural apresentar um resultado satisfatório de MAPE, o RMSE igual a 4.377,69 m³ é relativamente baixo, quando comparado com os valores observados, indicando que o modelo é capaz de fazer estimativas precisas.

De acordo com esses resultados, é possível concluir que, dispondo de uma base de dados confiável e com tamanho significativo para realizar simulações com redes neurais, a seguir o exemplo dessa pesquisa, maior ou igual a 5.000 dados, com pelo menos duas ou mais variáveis de entrada, as RNA's desempenham bons resultados na estimativa de resíduos da construção civil. Como foi verificado, a Amostra B com 10.000 apresentou melhores resultados preditivos.

Vale a pena ressaltar que é interessante realizar um tratamento prévio nos dados, a fim de retirar *outliers* e demais dados que prejudiquem a homogeneidade da amostra. Se, ainda assim, os resultados das estimativas não forem satisfatórios, ficando muito diferentes dos resultados esperados, se sugere utilizar outras configurações de redes neurais, aumentando o número de neurônios na camada oculta, ou o número de ciclos de treinamento, porém, mantendo a utilização do algoritmo de treinamento *Bayesian Regularization*, visto que esse apresentou o melhor desempenho para as amostras simuladas.

Essa pesquisa teve uma limitação quanto à obtenção de informações sobre obras reais de construção. A maneira como esses dados chegam e são organizados em órgãos municipais e empresas, dificultou a construção de uma amostra real ampla e robusta para ser aplicada às redes neurais. Além disso, a confiabilidade dos dados é um fator que pode ter influenciado nos resultados da Amostra R.

Como sugestão para trabalhos futuros, propõe-se:

- Aumentar a quantidade de dados, bem como de informações que possam servir como variáveis de entrada do modelo e investigar a influência dessas variáveis na previsão de resíduos;
- Estudar a precisão da estimativa de resíduos em diferentes regiões geográficas;
- Investigar a influência de variáveis ambientais e climáticas na geração de resíduos de construção civil e incorporar essas variáveis nos modelos de redes neurais;

- Integrar modelos de previsão de resíduos de construção civil em sistemas de gerenciamento de resíduos existentes.

REFERÊNCIAS

- ABBASI, M. et al. Forecasting municipal solid waste generation by hybrid support vector machine and partial least square model. **International Journal Environmental**. Res. 2013, 7, 27–38. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0956053X16302501>>. Acesso em: 24 ago. 2021.
- ABBASI, M.; EL HANANDEH, A. Forecasting municipal solid waste generation using artificial intelligence modelling approaches. **Waste Management**, [S. l.], v. 56, p. 13–22, 2016. DOI: 10.1016/j.wasman.2016.05.018. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0956053X16302501>>. Acesso em: 30 ago. 2021.
- ABRELPE. Panorama Dos Resíduos Sólidos no Brasil 2018/2019. **Panorama dos Resíduos Sólidos no Brasil 2018/2019**, [S. l.], p. 68, 2019. Disponível em: www.abrelpe.org.br. Disponível em: <<https://abrelpe.org.br/download-panorama-2018-2019/>>. Acesso em: 05 mar. 2022.
- ABRELPE. Panorama dos Resíduos Sólidos no Brasil 2022. **Panorama dos Resíduos Sólidos no Brasil 2022**, [S. l.], p. 60, 2022. Disponível em: <<https://abrelpe.org.br/download-panorama-2022/>>. Acesso em: 10 dez. 2022.
- AKANBI, O. O. et al. Deep learning model for Demolition Waste Prediction in a circular economy. **Journal of Cleaner Production**, [S. l.], v. 274, 2020. DOI: <https://doi.org/10.1016/j.jclepro.2020.122843> Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0959652620328882>>. Acesso em: 23 jan. 2023.
- AKINADE, O. O. et al. Designing out construction waste using BIM technology: Stakeholders' expectations for industry deployment. **Journal of Cleaner Production**, [S. l.], v. 180, p. 375–385, 2018. DOI: 10.1016/j.jclepro.2018.01.022. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0959652618300283>>. Acesso em: 10 mai. 2021.
- ALI ABDOLI, M. et al. Longterm forecasting of solid waste generation by the artificial neural networks. **Environmental Progress & Sustainable Energy**, [S. l.], v. 31, n. 4, p. 628–636, 2012. DOI: 10.1002/ep.10591. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/ep.10591>. Acesso em: 06 abr. 2021.
- BIJU, B. P. **Utilização do sistema de informação geográfica (SIG) na indicação de possíveis áreas aptas à disposição de resíduos de construção e de demolição**. 2015. Universidade Tecnológica Federal do Paraná, [S. l.], 2015. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/1328>. Acesso em: 18 jun. 2021.
- BRASIL. Lei Federal Nº 12.305 de 02 de Agosto de 2010. Brasil, 2010. **Institui a Política Nacional de Resíduos Sólidos; altera a Lei nº 9.605, de 12 de fevereiro de 1998; e dá outras providências**. Disponível em:

http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12305.htm. Acesso em: 26 jun. 2021.

BRASIL. Conselho Nacional do Meio Ambiente. Resolução nº 307, de 05 de julho de 2002. **Estabelece diretrizes, critérios e procedimentos para a gestão dos Resíduos da Construção Civil**. Brasília, DF, 02 jan. 2003. Disponível em: https://cetesb.sp.gov.br/licenciamento/documentos/2002_Res_CONAMA_307.pdf. Acesso em: 02 jun. 2021.

BRASIL. Conselho Nacional do Meio Ambiente. Resolução nº 431, de 24 de maio de 2011. **Altera o art. 3º da Resolução nº 307, de 5 de julho de 2002, do Conselho Nacional do Meio Ambiente - CONAMA, estabelecendo nova classificação para o gesso**. Brasília, DF. Disponível em: <https://www.legisweb.com.br/legislacao/?id=114767>. Acesso em: 02 jun. 2021.

BRASIL. Conselho Nacional do Meio Ambiente. Resolução nº 448, de 18 de janeiro de 2012. **Altera os arts. 2º, 4º, 5º, 6º, 8º, 9º, 10, 11 da Resolução nº 307, de 5 de julho de 2002, do Conselho Nacional do Meio Ambiente - CONAMA**. Brasília, DF. Disponível em: <https://www.legisweb.com.br/legislacao/?id=116060>. Acesso em: 02 jun. 2021.

BRASIL. Conselho Nacional do Meio Ambiente. Resolução nº 469, de 29 de julho de 2015. **Altera a Resolução CONAMA n 307, de 05 de julho de 2002, que estabelece diretrizes, critérios e procedimentos para a gestão dos resíduos da construção civil**. Brasília, DF. Disponível em: <http://www.ctpconsultoria.com.br/pdf/Resolucao-CONAMA-469-de-29-07-2015.pdf>. Acesso em: 02 jun. 2021.

BRASIL, Salesforce. **Deep Learning e Machine Learning: conheça as diferenças**. 2018. Disponível em: <https://www.salesforce.com/br/blog/2018/4/Machine-Learning-e-Deep-Learning-aprenda-as-diferencas.html>. Acesso em: 26 mai. 2022.

CHA, G.; HONG, W.; KIM, Y. Performance Improvement of Machine Learning Model Using Autoencoder to Predict Demolition Waste Generation Rate. **Environmetal Research and Public Health**. [S. l.], p. 1–21, 2023. Disponível em: https://www.researchgate.net/publication/368616480_Performance_Improvement_of_Machine_Learning_Model_Using_Autoencoder_to_Predict_Demolition_Waste_Generation_Rate. Acesso em: 26 fev. 2023.

CHA, G.; CHOI, S., W.; HONG, W.; PARK, C. Development of Machine Learning Model for Prediction of Demolition Waste Generation Rate of Buildings in Redevelopment Areas. **Environmetal Research and Public Health**. [S. l.], p. 1–17, 2022. Disponível em: <https://www.mdpi.com/1660-4601/20/1/107>. Acesso em: 26 jan. 2023.

COSKUNER, G. et al. Application of artificial intelligence neural network modeling to predict the generation of domestic, commercial and construction wastes. **Waste Management**. [S. l.], v. 39, n. 3, p. 499–507, 2021. DOI: 10.1177/0734242X20935181. Disponível em: https://www.researchgate.net/publication/342468609_Application_of_artificial_intellig

ence_neural_network_modeling_to_predict_the_generation_of_domestic_commercial_and_construction_wastes. Acesso em: 05 jun. 2021.

DAOUD, A. O. et al. Quantifying materials waste in the Egyptian construction industry: A critical analysis of rates and factors. **Ain Shams Engineering Journal**, [S. l.], n. xxxx, 2021. DOI: 10.1016/j.asej.2021.02.039. Disponível em: <https://doi.org/10.1016/j.asej.2021.02.039>. Acesso em: 10 fev. 2022.

DAVIS, P. et al. The classification of construction waste material using a deep convolutional neural network. **Automation in Construction**, [S. l.], v. 122, n. December 2020, p. 103481, 2021. DOI: 10.1016/j.autcon.2020.103481. Disponível em: <https://doi.org/10.1016/j.autcon.2020.103481>. Acesso em: 19 mar. 2022.

DIAS, M. F. Universidade do Vale do Rio Dos Sinos, UNISINOS, Brazil. **Modelo para estimar a geração de resíduos na produção de obras residenciais verticais**, [S. l.], p. 111, 2013. DOI: 10.1163/9789004382091_008. Disponível em: [http://www.repositorio.jesuita.org.br/handle/UNISINOS/4129](https://www.repositorio.jesuita.org.br/handle/UNISINOS/4129). Acesso em: 21 mar. 2022.

ELSEVIER. **What is Scopus Preview?** 2021. Disponível em: https://service.elsevier.com/app/answers/detail/a_id/15534/supporthub/scopus/#tips. Acesso em: 14 set. 2021.

EPA. **Construction and Demolition Debris: Material-Specific Data**. 2020. Disponível em: <https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/construction-and-demolition-debris-material>. Acesso em: 13 set. 2021.

FILHO GERALDO, P. R. Avaliação do gerenciamento de resíduos sólidos: estudo de caso em um consultório odontológico. **Cippus**, [S. l.], v. 9, n. 1, 2021. DOI: 10.18316/cippus.v9i1.7693.

FINOCCHIO, M. A. F. Noções De Redes Neurais Artificiais. **Universidade Tecnológica Federal do Paraná**, [S. l.], p. 1–29, 2014. Disponível em: <http://paginapessoal.utfpr.edu.br/mafinocchio/labsi-laboratorio-de-seguranca-e-iluminacao/redes-neurais-artificiais/NOCaO%20DE%20REDES%20NEURAIIS%20ARTIFICIAIS.pdf/view>. Acesso em: 28 abr. 2022.

FIRAT, M. et al. Comparative analysis of neural network techniques for predicting water consumption time series. **Journal of Hydrology**, [S. l.], v. 384, n. 1–2, p. 46–51, 2010. DOI: 10.1016/j.jhydrol.2010.01.005. Disponível em: <http://dx.doi.org/10.1016/j.jhydrol.2010.01.005>. Acesso em: 03 out. 2021.

GÁLVEZ-MARTOS, J. L. et al. Construction and demolition waste best management practice in Europe. **Resources, Conservation and Recycling**, [S. l.], v. 136, n. May, p. 166–178, 2018. DOI: 10.1016/j.resconrec.2018.04.016. Disponível em: <https://doi.org/10.1016/j.resconrec.2018.04.016>. Acesso em: 15 nov. 2021.

GHORBANI, B. et al. Thermal and mechanical properties of demolition wastes in geothermal pavements by experimental and machine learning techniques.

Construction and Building Materials, [S. l.], v. 280, p. 122499, 2021. DOI: 10.1016/j.conbuildmat.2021.122499. Disponível em: <https://doi.org/10.1016/j.conbuildmat.2021.122499>. Acesso em: 29 jan. 2022.

GIL, A. C. Como elaborar projetos de pesquisa. São Paulo: Atlas, 2002. Disponível em: <<https://docente.ifrn.edu.br/mauriciofacanha/ensino-superior/redacao-cientifica/livros/gil-a.-c.-como-elaborar-projetos-de-pesquisa.-sao-paulo-atlas-2002./view>>. Acesso em: 12 set. 2021.

GUERRA, B. C. et al. BIM-based automated construction waste estimation algorithms: The case of concrete and drywall waste streams. **Waste Management**, [S. l.], v. 87, p. 825–832, 2019. DOI: 10.1016/j.wasman.2019.03.010. Disponível em: <https://doi.org/10.1016/j.wasman.2019.03.010>. Acesso em: 13 mar. 2022.

HAYKIN, S. **Neural Networks and Learning Machines**. 3º ed. Ontario: Pearson, 2009.

HU, H. et al. Estimation of construction waste generation based on an improved on-site measurement and SVM-based prediction model: A case of commercial buildings in China. **Waste Management**. [S. l.], v. 126, p. 791-799, 2021. DOI: <https://doi.org/10.1016/j.wasman.2021.04.012>. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0956053X21002105>>. Acesso em: 30 jan. 2023.

JÚNIOR OLIVEIRA, C. **Métricas para Regressão: Entendendo as métricas R², MAE, MAPE, MSE e RMSE**. 2021. Disponível em: <https://medium.com/data-hackers/prevendo-números-entendendo-métricas-de-regressão-35545e011e70#:~:text=O erro quadrático médio>. Acesso em: 4 out. 2022.

KERN, A. P. et al. Waste generated in high-rise buildings construction: A quantification model based on statistical multiple regression. **Waste Management**, [S. l.], v. 39, n. February, p. 35–44, 2015. DOI: 10.1016/j.wasman.2015.01.043. Disponível em: <http://dx.doi.org/10.1016/j.wasman.2015.01.043>. Acesso em: 12 jul. 2022.

LI, J. et al. A model for estimating construction waste generation index for building project in China. **Resources, Conservation and Recycling**, [S. l.], v. 74, p. 20–26, 2013. DOI: 10.1016/j.resconrec.2013.02.015. Disponível em: <http://dx.doi.org/10.1016/j.resconrec.2013.02.015>. Acesso: 02 dez. 2021.

LIU, K. et al. Prediction of carbonation depth for recycled aggregate concrete using ANN hybridized with swarm intelligence algorithms. **Construction and Building Materials**, [S. l.], v. 301, n. January, p. 124382, 2021. DOI: 10.1016/j.conbuildmat.2021.124382. Disponível em: <https://doi.org/10.1016/j.conbuildmat.2021.124382>. Acesso em: 24 nov. 2021

LIU, W. et al. Profile of developments in biomass-based bioenergy research: A 20-year perspective. **Scientometrics**, [S. l.], v. 99, n. 2, p. 507–521, 2014. DOI: 10.1007/s11192-013-1152-z. Disponível em: https://www.researchgate.net/publication/259635938_Profile_of_developments_in_bi

omass-based_bioenergy_research_A_20-year_perspective. Acesso em: 06 out. 2021.

LLATAS, C. A model for quantifying construction waste in projects according to the European waste list. **Waste Management**, [S. l.], v. 31, n. 6, p. 1261–1276, 2011. DOI: 10.1016/j.wasman.2011.01.023. Disponível em: <http://dx.doi.org/10.1016/j.wasman.2011.01.023>. Acesso em: 12 dez. 2021.

LU, W. et al. Estimating construction waste generation in the Greater Bay Area, China using machine learning. **Waste Management**, [S. l.], v. 134, n. February, p. 78–88, 2021. DOI: 10.1016/j.wasman.2021.08.012. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0956053X21004372>. Acesso em: 15 dez. 2021.

LU, W. et al. The S-curve for forecasting waste generation in construction projects. **Waste Management**, [S. l.], v. 56, p. 23–34, 2016. DOI: 10.1016/j.wasman.2016.07.039. Disponível em: <http://dx.doi.org/10.1016/j.wasman.2016.07.039>. Acesso em: 15 dez. 2021.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre Aprendizado de Máquina. **Sistemas inteligentes: fundamentos e aplicações**, [S. l.], p. 89–114, 2003. Disponível em: <https://scholar.google.com.br/citations?user=OZ-WHMsAAAAJ&hl=pt-BR>. Acesso em: 12 set. 2021.

NAGALLI, A. Estimation of construction waste generation using machine learning. **Proceedings of Institution of Civil Engineers: Waste and Resource Management**, [S. l.], v. 174, n. 1, p. 22–31, 2021. a. DOI: 10.1680/jwarm.20.00019. Disponível em: <https://www.icevirtuallibrary.com/doi/10.1680/jwarm.20.00019>. Acesso em: 02 fev. 2022.

NAGALLI, A. **Aspectos quantitativos da geração de resíduos da construção civil**. 1. ed. São Paulo: Oficina de Textos, 2021. b. Disponível em: <https://www.ofitexto.com.br/livro/aspectos-quantitativos-da-geracao-de-residuos-da-construcao-civil/>. Acesso em: 10 mar. 2022.

NAGALLI, A. **Resíduos Sólidos de Construção Civil**. 2. ed. São Paulo: Oficina de Textos, 2022. Disponível em: <https://www.lojaofitexto.com.br/residuos-de-construcao-civil/p>. Acesso em: 15 nov. 2022.

NAGALLI, A; DE CARVALHO, K. Q. Model for estimating construction waste generation in masonry building. **Proceedings of Institution of Civil Engineers: Waste and Resource Management**, [S. l.], v. 172, n. 1, p. 28–36, 2018. DOI: 10.1680/jwarm.18.00016. Disponível em: <https://www.icevirtuallibrary.com/doi/abs/10.1680/jwarm.18.00016?src=recsys&journalCode=jwarm>. Acesso em: 20 nov. 2021.

OPOKU, A. Biodiversity and the built environment: Implications for the Sustainable Development Goals (SDGs). **Resources, Conservation and Recycling**, [S. l.], v. 141, n. October 2018, p. 1–7, 2019. DOI: 10.1016/j.resconrec.2018.10.011. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0921344918303768>. Acesso em: 25 abr. 2022.

PALA, L. O. de O. **Revisitando a estimação do Coeficiente de Determinação**. 2019. Universidade Federal de Alfenas, [S. l.], 2019. Disponível em: [http://bdttd.unifal-mg.edu.br:8080/bitstream/tede/1419/5/Dissertação de Luiz Otávio de Oliveira Pala.pdf](http://bdttd.unifal-mg.edu.br:8080/bitstream/tede/1419/5/Dissertação%20de%20Luiz%20Otávio%20de%20Oliveira%20Pala.pdf). Acesso em: 12 ago. 2021.

PINTO, T. P. **Metodologia para a gestão diferenciada de resíduos sólidos da construção urbana**. 1999. Escola Politécnica da Universidade de São Paulo, [S. l.], 1999. Disponível em: <https://repositorio.usp.br/item/001061434>. Acesso em: 15 out. 2022.

PREFEITURA MUNICIPAL DE RECIFE. **Diretrizes para elaboração do Plano de Gerenciamento de Resíduos da Construção Civil PGRCC** 2019. p. 6. Disponível em: <https://www2.recife.pe.gov.br/sites/default/files/diretrizes-pgrcc.pdf>. Acesso em: 15 out. 2022.

RAY, S. et al. Predicting the strength of concrete made with stone dust and nylon fiber using artificial neural network. **Heliyon**, [S. l.], v. 8, n. 3, p. e09129, 2022. DOI: 10.1016/j.heliyon.2022.e09129. Disponível em: <https://doi.org/10.1016/j.heliyon.2022.e09129>. Acesso em: 12 out. 2022.

SANTOS, R. P.; TUBINO, R. Potential evaluation of the use of construction and demolition waste (CDW) in the recovery of degraded soils by mining in Brazil. **Resources, Conservation & Recycling Advances**, [S. l.], v. 12, p. 200060, 2021. DOI: 10.1016/j.rcradv.2021.200060. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2667378921000092>. Acesso em: 01 mar. 2022.

SCHAMNE, A. N. **Avaliação do potencial de aplicação dos preceitos da logística reversa de resíduos sólidos ao setor da construção civil em Curitiba, Paraná**. Dissertação apresentada ao Programa de Pós- Graduação em Engenharia Civil da Universidade Tecnológica Federal do Paraná, como requisito parcial para obtenção do grau de Mestre em Engenharia Civil. Área de Concentração: Meio Ambiente. Linha de pesquisa: Su, [S. l.], p. 157, 2016. Disponível em: <https://repositorio.utfpr.edu.br/jspui/handle/1/2700>. Acesso em: 26 mai. 2022.

SILVA, R. C. **Potencial de recuperação de materiais e componentes de edificações: análise crítica em um processo de reabilitação**. 2020. 205 f. Tese (Área de concentração: Construção civil, Linha de pesquisa: Gestão e sustentabilidade da construção) - Programa de Pós-Graduação em Engenharia Civil, Universidade Tecnológica Federal do Paraná. Curitiba, 2020. Disponível em: <<http://repositorio.utfpr.edu.br/jspui/handle/1/5044>>. Acesso em: 16 mar. 2021.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais Para Engenharia e Ciências Aplicadas**. 2ª ed. São Paulo.

SINIR. **Sistema Nacional de Informações sobre os Resíduos Sólidos**. 2020.

Disponível em: <https://sinir.gov.br/>. Acesso em: 13 set. 2021.

SONI, U. et al. Forecasting municipal solid waste generation using artificial intelligence models - A case study in India. **SN Applied Sciences**. 2019, 1, 1-10. Disponível em: <https://link.springer.com/article/10.1007/s42452-018-0157-x>. Acesso em: 02 mar. 2023.

SOUZA, F. S. et al. Mapping and recycling proposal for the construction and demolition waste generated in the Brazilian Amazon. **Resources, Conservation and Recycling**, [S. l.], v. 176, p. 105896, 2022. DOI: 10.1016/J.RESCONREC.2021.105896. Disponível em: <https://www.repositorio.ufop.br/handle/123456789/15524>. Acesso em: 15 nov. 2022.

SU, H. N.; LEE, P. C. Mapping knowledge structure by keyword co-occurrence: A first look at journal papers in Technology Foresight. **Scientometrics**, [S. l.], v. 85, n. 1, p. 65–79, 2010. DOI: 10.1007/s11192-010-0259-8. Disponível em: <https://link.springer.com/article/10.1007/s11192-010-0259-8>. Acesso em: 25 nov. 2021.

TABOADA, G. L. et al. Exploratory data analysis and data envelopment analysis of construction and demolition waste management in the European economic area. **Sustainability (Switzerland)**, [S. l.], v. 12, n. 12, 2020. DOI: 10.3390/su12124995. Disponível em: <https://www.mdpi.com/2071-1050/12/12/4995>. Acesso em: 12 ago. 2021.

TAM, V. W. Y. et al. A prediction model for compressive strength of CO₂ concrete using regression analysis and artificial neural networks. **Construction and Building Materials**, [S. l.], v. 324, n. February, p. 126689, 2022. DOI: 10.1016/j.conbuildmat.2022.126689. Disponível em: <https://doi.org/10.1016/j.conbuildmat.2022.126689>. Acesso em: 13 dez. 2021.

TEIXEIRA, E. C. et al. Modelling waste generated during construction of buildings using regression analysis. **Waste Management and Research**, [S. l.], v. 38, n. 8, p. 857–867, 2020. DOI: 10.1177/0734242X19893012. Disponível em: <https://journals.sagepub.com/doi/abs/10.1177/0734242X19893012?journalCode=wmra>. Acesso em: 13 dez. 2021.

SÁEZ, P. V.; OSMANI, M. A diagnosis of construction and demolition waste generation and recovery practice in the European Union. **Journal of Cleaner Production**, [S. l.], v. 241, 2019. DOI: 10.1016/j.jclepro.2019.118400. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0959652619332706>. Acesso em: 02 ago. 2021.

WU, Z. et al. Quantifying construction and demolition waste : An analytical review. **Waste Management**, [S. l.], v. 34, n. 9, p. 1683–1692, 2014. DOI: 10.1016/j.wasman.2014.05.010. Disponível em: <http://dx.doi.org/10.1016/j.wasman.2014.05.010>. Acesso em: 26 out. 2022.

ZHANG, N. et al. Differences of methods to quantify construction and demolition waste for less-developed but fast-growing countries: China as a case study.

Environmental Science and Pollution Research, [S. l.], v. 26, n. 25, p. 25513–25525, 2019. DOI: 10.1007/s11356-019-05841-4. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31264149/>. Acesso em: 04 jul. 2022.

APÊNDICE A - Dados da Amostra R

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
906	3.085,00	600,10	80,55%
1401	701,00	805,30	14,88%
1505	572,00	849,66	48,54%
2108	847,00	1.116,60	31,83%
2700	1.012,00	1.393,92	37,74%
2840	643,28	1.461,68	127,22%
2896	1.550,00	1.489,01	3,93%
3068	1.001,00	1.573,79	57,22%
3098	1.608,00	1.588,70	1,20%
3098	2.263,30	1.588,71	29,81%
3555	1.535,00	1.820,41	18,59%
3598	954,50	1.842,81	93,07%
3809	4.007,00	1.952,82	51,26%
3842	1.864,00	1.970,21	5,70%
3896	877,00	1.998,76	127,91%
3971	1.160,00	2.038,59	75,74%
3978	1.758,00	2.042,32	16,17%
3998	1.165,00	2.052,99	76,22%
4002	2.304,00	2.055,12	10,80%
4004	2.412,00	2.056,19	14,75%
4004	1.784,00	2.056,19	15,26%
4009	2.828,00	2.058,86	27,20%
4022	2.781,00	2.065,80	25,72%
4035	1.241,60	2.072,90	66,95%
4045	2.121,00	2.078,11	2,02%
4056	909,00	2.084,00	129,26%
4218	2.364,10	2.171,52	8,15%
4326	872,00	2.230,08	155,74%
4478	1.673,00	2.313,43	38,28%
4509	1.278,00	2.330,54	82,36%
4545	964,00	2.350,45	143,82%
4603	4.565,00	2.382,63	47,81%
4609	3.067,00	2.385,97	22,21%
4617	2.764,50	2.390,44	13,53%
4879	2.739,00	2.537,37	7,36%
4889	3.317,00	2.543,03	23,33%
4978	4.464,00	2.593,51	41,90%
4996	2.444,00	2.603,75	6,54%
5012	1.548,00	2.612,86	68,79%
5021	2.819,00	2.617,99	7,13%
5025	2.944,00	2.620,27	11,00%
5055	2.165,00	2.637,39	21,82%
5063	3.930,10	2.641,89	32,78%
5063	2.418,00	2.641,96	9,26%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
5063	1.054,00	2.641,96	150,66%
5065	2.169,00	2.643,11	21,86%
5089	3.500,00	2.656,83	24,09%
5094	1.449,00	2.659,69	83,55%
5126	1.420,00	2.678,03	88,59%
5225	2.640,00	2.734,96	3,60%
5566	3.734,00	2.933,39	21,44%
5649	2.706,00	2.982,21	10,21%
5839	2.367,00	3.094,70	30,74%
5868	2.385,60	3.111,93	30,45%
5993	2.125,00	3.186,60	49,96%
5999	3.986,00	3.190,19	19,97%
6003	3.422,50	3.192,48	6,72%
6019	2.187,00	3.202,18	46,42%
6021	5.017,00	3.203,37	36,15%
6023	3.220,00	3.204,57	0,48%
6026	2.151,00	3.206,37	49,06%
6031	3.480,00	3.209,37	7,78%
6058	3.311,00	3.225,57	2,58%
6065	3.479,00	3.229,78	7,16%
6066	4.523,00	3.230,38	28,58%
6087	3.003,00	3.243,00	7,99%
6092	3.023,00	3.246,00	7,38%
6208	1.644,00	3.315,92	101,70%
6222	2.580,00	3.324,38	28,85%
6309	4.143,00	3.377,06	18,49%
6309	2.988,00	3.377,06	13,02%
6383	5.185,20	3.421,78	34,01%
6453	1.622,80	3.464,59	113,49%
6503	1.384,00	3.495,17	152,54%
6619	3.786,00	3.566,18	5,81%
6733	4.448,00	3.636,24	18,25%
6734	5.257,00	3.636,86	30,82%
6744	3.179,00	3.643,02	14,60%
6835	1.602,00	3.699,15	130,91%
6888	5.959,00	3.731,91	37,37%
6914	4.555,00	3.748,00	17,72%
6920	4.558,00	3.751,72	17,69%
6962	3.956,00	3.777,74	4,51%
6967	4.664,00	3.780,84	18,94%
6970	3.312,00	3.782,70	14,21%
6972	3.961,00	3.783,94	4,47%
6986	6.159,00	3.792,62	38,42%
7014	1.981,00	3.810,00	92,33%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
7071	3.984,00	3.845,42	3,48%
7082	4.334,00	3.852,26	11,12%
7088	4.656,00	3.856,00	17,18%
7143	3.010,30	3.889,96	29,22%
7154	4.776,00	3.897,09	18,40%
7209	4.834,00	3.931,39	18,67%
7356	6.577,00	4.023,29	38,83%
7565	4.973,00	4.154,59	16,46%
7707	3.612,00	4.243,89	17,49%
7722	7.092,00	4.253,35	40,03%
7779	2.429,00	4.289,31	76,59%
7787	3.522,00	4.294,36	21,93%
7812	2.252,00	4.310,30	91,40%
7826	5.208,00	4.318,99	17,07%
7854	3.935,00	4.336,68	10,21%
7880	3.821,00	4.353,12	13,93%
8000	5.330,00	4.429,04	16,90%
8005	6.914,00	4.432,28	35,89%
8008	4.515,00	4.434,11	1,79%
8008	4.455,00	4.434,11	0,47%
8012	5.300,00	4.436,64	16,29%
8034	2.475,00	4.450,58	79,82%
8082	3.763,00	4.480,99	19,08%
8085	4.553,00	4.482,89	1,54%
8088	5.604,00	4.484,79	19,97%
8203	3.976,00	4.557,73	14,63%
8375	4.092,00	4.667,25	14,06%
8381	3.381,00	4.670,75	38,15%
8383	2.515,00	4.672,03	85,77%
8400	4.985,00	4.682,83	6,06%
8485	5.169,00	4.736,85	8,36%
8502	5.635,00	4.747,66	15,75%
8540	3.827,00	4.771,82	24,69%
8554	3.441,00	4.780,73	38,93%
8606	2.880,00	4.813,63	67,14%
8620	4.585,00	4.822,70	5,18%
8648	3.380,00	4.840,51	43,21%
8655	7.623,00	4.844,96	36,44%
8666	5.733,00	4.851,96	15,37%
8835	2.585,00	4.959,46	91,86%
8845	7.654,00	4.965,82	35,12%
8866	4.719,00	4.979,18	5,51%
8867	4.081,00	4.979,82	22,02%
8868	4.959,00	4.980,45	0,43%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
8979	4.144,00	5.051,05	21,89%
9056	7.823,00	5.100,01	34,81%
9065	4.360,00	5.105,73	17,10%
9076	4.994,00	5.112,72	2,38%
9092	3.644,00	5.122,89	40,58%
9099	4.193,00	5.127,34	22,28%
9145	5.122,00	5.156,57	0,67%
9197	3.653,00	5.189,61	42,06%
9309	3.612,10	5.260,60	45,64%
9356	3.637,00	5.290,55	45,46%
9565	3.809,00	5.422,99	42,37%
9702	3.882,00	5.509,63	41,93%
9725	3.288,00	5.524,16	68,01%
9756	8.511,00	5.543,74	34,86%
9784	5.182,00	5.561,41	7,32%
9846	6.236,40	5.600,75	10,19%
9858	3.355,00	5.608,08	67,16%
9870	4.532,00	5.615,64	23,91%
9870	2.130,00	5.615,89	163,66%
9891	6.519,00	5.628,87	13,65%
9899	5.494,00	5.633,91	2,55%
9899	4.742,00	5.633,91	18,81%
9944	11.098,00	5.662,18	48,98%
9988	8.736,00	5.689,92	34,87%
10002	4.588,00	5.698,72	24,21%
10002	4.421,00	5.698,72	28,90%
10004	4.788,00	5.699,98	19,05%
10011	2.468,00	5.704,38	131,13%
10034	8.773,00	5.718,83	34,81%
10098	6.523,00	5.759,02	11,71%
10100	6.647,00	5.760,27	13,34%
10104	5.336,00	5.762,78	8,00%
10200	9.972,00	5.822,95	41,61%
10254	3.966,00	5.856,74	47,67%
10297	7.036,00	5.883,62	16,38%
10343	4.726,00	5.912,34	25,10%
10361	9.439,00	5.923,38	37,25%
10387	7.108,00	5.939,78	16,44%
10402	4.015,00	5.949,13	48,17%
10548	9.151,00	6.039,95	34,00%
10550	4.809,00	6.041,19	25,62%
10601	8.510,00	6.072,74	28,64%
10622	9.211,00	6.085,85	33,93%
10762	7.296,00	6.172,42	15,40%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
10888	3.657,00	6.250,02	70,91%
10920	3.140,00	6.269,49	99,67%
10920	2.635,00	6.269,49	137,93%
10982	3.683,00	6.307,72	71,27%
11001	5.099,00	6.319,36	23,93%
11014	2.934,00	6.327,32	115,66%
11033	3.932,00	6.338,95	61,21%
11065	8.997,00	6.358,52	29,33%
11065	6.140,00	6.358,52	3,56%
11099	8.488,00	6.379,29	24,84%
11195	7.212,00	6.437,80	10,73%
11220	3.103,00	6.453,00	107,96%
11223	5.907,00	6.454,83	9,27%
11298	16.127,14	6.500,23	59,69%
11566	7.576,00	6.662,00	12,06%
11634	2.465,00	6.702,74	171,92%
11689	27.340,00	6.735,62	75,36%
11767	10.262,00	6.782,11	33,91%
11799	2.929,00	6.801,14	132,20%
11882	5.630,00	6.850,38	21,68%
12005	6.045,00	6.923,03	14,52%
12022	2.487,00	6.933,03	178,77%
12066	9.010,10	6.958,66	22,77%
12067	5.505,00	6.959,49	26,42%
12098	5.277,00	6.977,69	32,23%
12110	9.242,00	6.984,72	24,42%
12123	7.941,00	6.992,34	11,95%
12128	7.799,00	6.995,27	10,31%
12303	5.485,00	7.097,35	29,40%
12343	10.731,00	7.120,57	33,64%
12498	8.651,20	7.210,26	16,66%
12667	12.006,60	7.306,85	39,14%
12841	3.435,00	7.405,51	115,59%
13002	5.897,00	7.496,46	27,12%
13003	5.117,00	7.497,03	46,51%
13045	6.176,00	7.520,55	21,77%
13130	10.450,00	7.568,00	27,58%
13187	7.617,50	7.599,74	0,23%
13430	5.305,00	7.733,73	45,78%
13444	8.625,00	7.741,40	10,24%
13456	10.246,00	7.747,96	24,38%
13466	7.378,00	7.753,32	5,09%
14006	7.715,00	8.044,27	4,27%
14044	6.035,00	8.064,41	33,63%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
14048	2.899,00	8.066,52	178,25%
14614	11.501,00	8.361,32	27,30%
14791	3.279,00	8.451,25	157,74%
14831	5.360,00	8.471,74	58,05%
14899	9.467,00	8.506,06	10,15%
15005	4.064,00	8.559,39	110,61%
15023	3.305,00	8.568,60	159,26%
15025	3.566,00	8.569,41	140,31%
15304	4.145,00	8.707,88	110,08%
15309	6.730,00	8.710,54	29,43%
15434	8.156,00	8.771,98	7,55%
15615	9.998,00	8.860,25	11,38%
15676	13.572,00	8.889,81	34,50%
16066	3.718,00	9.076,78	144,13%
16198	7.347,00	9.139,29	24,39%
16604	13.842,80	9.329,48	32,60%
16671	10.656,00	9.360,50	12,16%
16793	18.461,10	9.416,83	48,99%
17099	14.236,00	9.557,39	32,86%
17203	7.754,00	9.604,84	23,87%
17350	7.582,00	9.671,69	27,56%
17664	7.717,00	9.813,76	27,17%
17892	32.976,10	9.916,30	69,93%
17921	15.806,10	9.929,44	37,18%
17923	9.832,00	9.930,36	1,00%
17931	6.891,00	9.933,95	44,16%
18008	13.683,00	9.968,54	27,15%
18041	32.891,00	9.983,18	69,65%
18063	4.340,00	9.993,23	130,26%
18079	6.020,00	10.000,51	66,12%
18182	11.626,00	10.046,62	13,58%
18205	5.084,00	10.056,94	97,82%
18618	9.691,00	10.242,11	5,69%
18635	14.360,40	10.249,59	28,63%
18663	6.150,00	10.262,30	66,87%
18745	8.422,00	10.299,10	22,29%
18876	17.386,20	10.358,10	40,42%
18878	5.320,00	10.358,67	94,71%
18888	10.371,00	10.363,35	0,07%
19023	12.153,00	10.424,10	14,23%
19025	12.154,00	10.425,01	14,23%
19034	10.005,00	10.429,06	4,24%
19319	7.328,00	10.557,81	44,07%
19538	10.253,00	10.657,30	3,94%

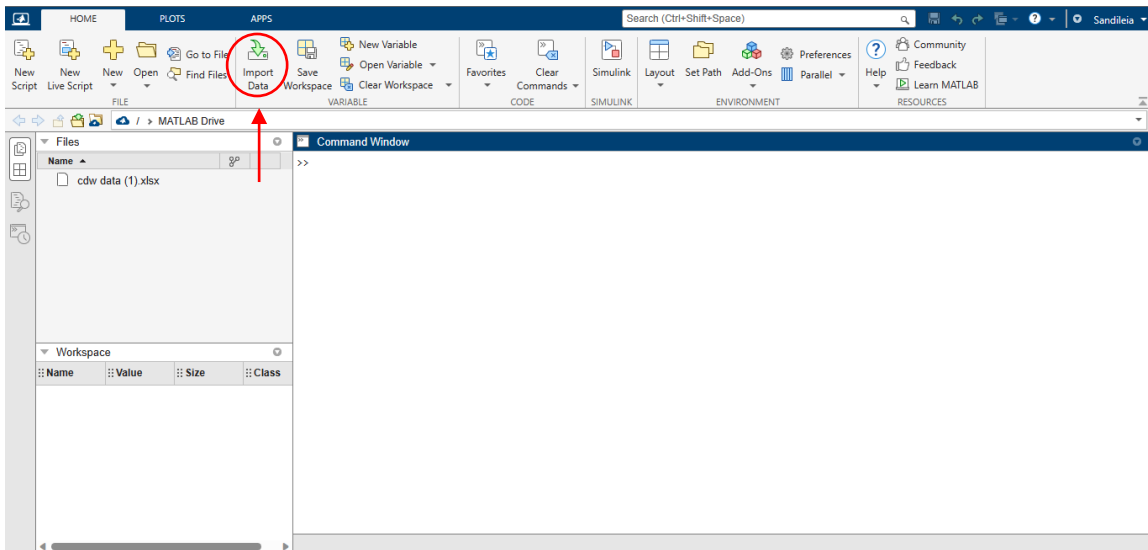
Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
19674	7.679,00	10.719,39	39,59%
19767	5.900,00	10.762,00	82,41%
19802	4.362,00	10.778,07	147,09%
19867	8.688,00	10.807,96	24,40%
20005	9.420,00	10.871,67	15,41%
20006	13.021,00	10.872,13	16,50%
20008	5.391,00	10.873,06	101,69%
20010	8.705,00	10.873,98	24,92%
20011	13.025,00	10.874,45	16,51%
20013	8.625,00	10.875,37	26,09%
20017	9.425,00	10.877,22	15,41%
20024	9.304,90	10.880,55	16,93%
20025	8.984,00	10.880,93	21,11%
20064	10.545,00	10.899,01	3,36%
20065	12.829,00	10.899,47	15,04%
20159	19.964,70	10.943,31	45,19%
20536	26.010,06	11.120,27	57,25%
21013	7.977,00	11.349,36	42,28%
21089	13.718,00	11.386,44	17,00%
21163	29.577,00	11.422,71	61,38%
21478	6.065,00	11.579,05	90,92%
22219	21.436,76	11.960,76	44,20%
22312	11.743,00	12.010,05	2,27%
22801	16.262,30	12.275,93	24,51%
23097	12.693,00	12.442,12	1,98%
23409	12.497,00	12.621,86	1,00%
23578	7.785,00	12.721,23	63,41%
23965	12.475,00	12.954,34	3,84%
24048	6.021,00	13.005,36	116,00%
24084	9.161,00	13.027,61	42,21%
24344	19.472,00	13.190,35	32,26%
25033	16.312,00	13.639,72	16,38%
25525	14.055,00	13.977,09	0,55%
26581	12.024,00	14.748,20	22,66%
27070	7.914,60	15.126,85	91,13%
27667	7.520,00	15.606,60	107,53%
27777	12.566,00	15.697,07	24,92%
28080	12.717,00	15.949,46	25,42%
28117	22.984,40	15.980,25	30,47%
28280	17.511,00	16.118,55	7,95%
28631	12.921,00	16.419,90	27,08%
28806	19.203,00	16.572,24	13,70%
28890	15.248,00	16.645,84	9,17%
28890	12.580,00	16.645,84	32,32%

Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
30004	13.070,00	17.648,51	35,03%
30010	19.618,00	17.654,03	10,01%
30042	16.612,00	17.683,49	6,45%
30054	18.967,00	17.694,55	6,71%
30298	16.007,00	17.920,34	11,95%
30383	29.401,10	17.999,43	38,78%
30603	23.690,00	18.204,96	23,15%
30670	10.059,00	18.267,82	81,61%
30719	11.450,00	18.313,65	59,94%
31099	27.208,00	18.672,74	31,37%
31099	6.862,00	18.672,74	172,12%
31318	16.203,60	18.880,45	16,52%
31564	27.604,00	19.115,72	30,75%
31665	29.809,60	19.212,29	35,55%
32088	28.532,00	19.618,56	31,24%
32089	21.020,00	19.619,52	6,66%
33065	12.723,00	20.561,00	61,60%
36012	9.679,00	23.358,34	141,33%
36420	29.021,90	23.733,07	18,22%
37512	32.084,20	24.716,95	22,96%
38122	25.376,00	25.254,40	0,48%
38451	24.512,20	25.541,20	4,20%
38534	36.180,40	25.613,55	29,21%
39051	30.018,00	26.058,53	13,19%
40021	35.184,00	26.882,97	23,59%
40055	18.362,00	26.911,67	46,56%
40102	31.281,00	26.951,32	13,84%
40691	12.375,00	27.447,18	121,80%
44071	34.030,00	30.330,36	10,87%
44661	39.472,00	30.858,08	21,82%
45033	29.877,00	31.196,75	4,42%
45632	36.405,20	31.752,56	12,78%
48209	32.106,00	34.295,16	6,82%
50003	33.357,00	36.162,89	8,41%
50437	28.614,00	36.613,75	27,96%
51008	34.067,00	37.199,27	9,19%
51884	36.451,10	38.067,61	4,43%
52950	44.922,30	39.046,69	13,08%
55788	48.415,00	40.944,06	15,43%
57562	45.523,40	41.454,28	8,94%
58732	50.717,90	41.524,17	18,13%
59054	22.909,00	41.514,96	81,22%
60987	36.909,70	41.339,96	12,00%
61457	47.084,20	41.301,33	12,28%

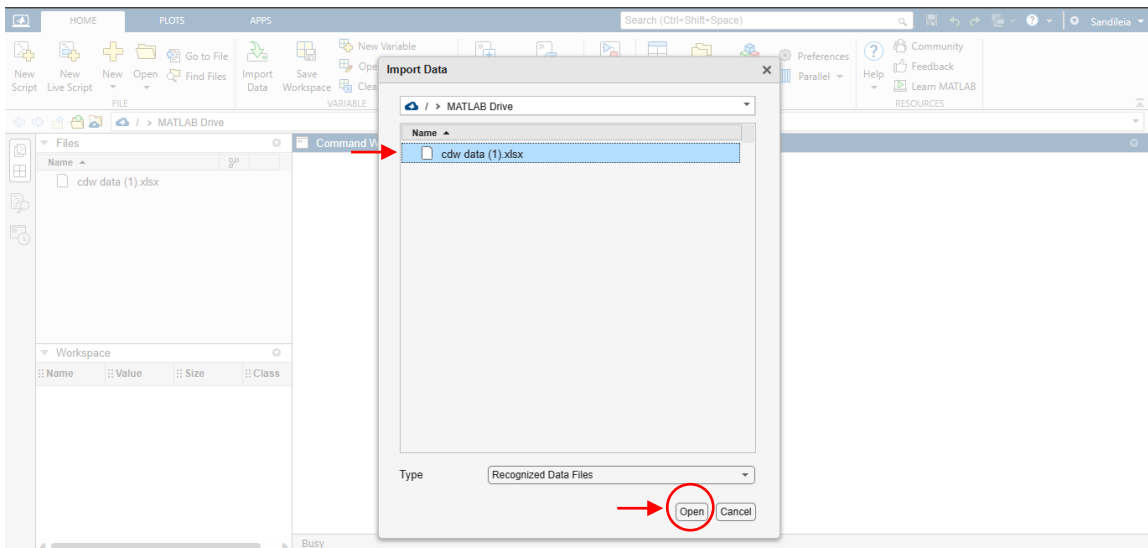
Amostra R			
MAPE		RMSE	
38,35%		4.337,69 m ³	
Área (m²)	Resíduos (m³)	Resíduos estimados (m³)	Erro Absoluto
68351	56.892,40	46.111,90	18,95%
70035	49.786,00	49.462,32	0,65%
76013	52.322,00	64.021,13	22,36%
78776	69.214,00	69.178,19	0,05%
97662	56.254,00	56.289,58	0,06%
121313	98.279,50	98.218,92	0,06%
131089	41.948,00	95.775,62	128,32%
138824	82.597,50	82.609,38	0,01%

APÊNDICE B - Passo a passo para realizar simulações com Redes Neurais Artificiais no MATLAB®

1 - Para iniciar, é necessário importar os dados para o MATLAB®, clicando em *Import Data*.



2 - O arquivo a ser importado deve ser selecionado.



3 - Em seguida é necessário selecionar *Numeric Matrix* e selecionar os dados de entrada que serão importados.

The screenshot shows the MATLAB Import Wizard interface. The 'Name' field is 'cdwData_1_S2' and the 'Type' is 'Numeric Matrix'. A red arrow points to the 'Type' dropdown. The spreadsheet preview shows the following data:

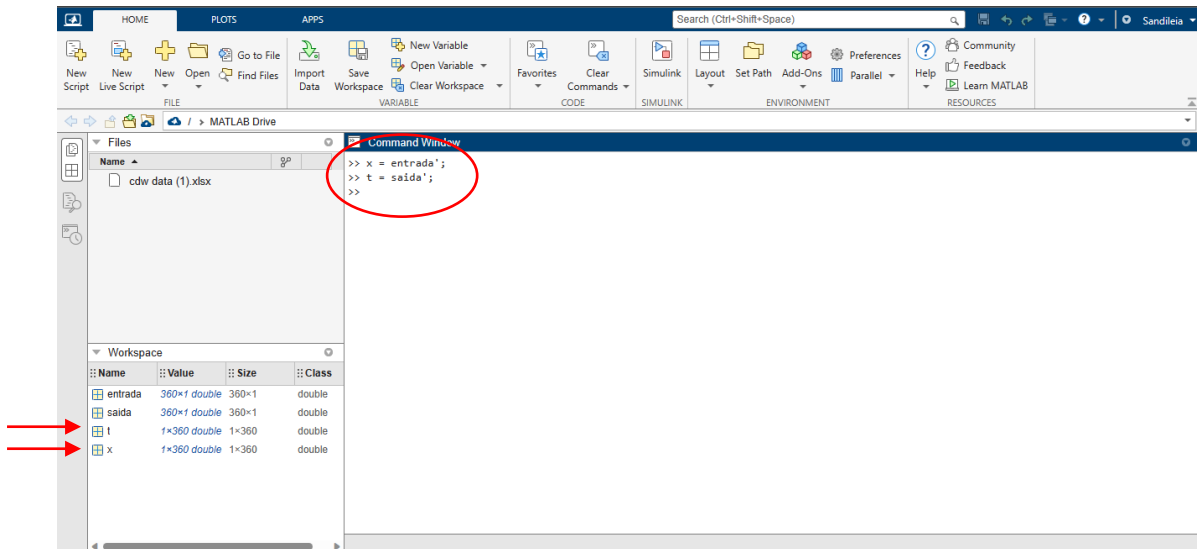
	N	O	P	Q	R	S	T	U	V	W	X	Y
Var14	Var15	Var16	Var17	Var18	Var19	Var20	Var21	Var22	Var23	Var24	Var25	
Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
1												
2	Valores se...											
3	Area M²	Residuo M²	Taxa	Mediana	L-Sup	L-Inf			área	res		
4	906	3085	0.29368	2.111	5.1581	-0.93615			906	3085		572
5	11689	27340	0.42754	2.111	5.1581	-0.93615			1401.21	701		643.28
6	17891.74	32976.1	0.54257	2.111	5.1581	-0.93615			1505	572		701
7	18040.61	32891	0.5485	2.111	5.1581	-0.93615			2108	847		847
8	11297.79	16127.14	0.70055	2.111	5.1581	-0.93615			2700	1012		872
9	21163	29577	0.71552	2.111	5.1581	-0.93615			2840	643.28		877
10	20535.95	26010.06	0.78954	2.111	5.1581	-0.93615			2896	1550		909
11	9943.9	11098	0.89601	2.111	5.1581	-0.93615			3068	1001		954.5

4 - Na sequência é necessário selecionar os dados de saída que serão importados.

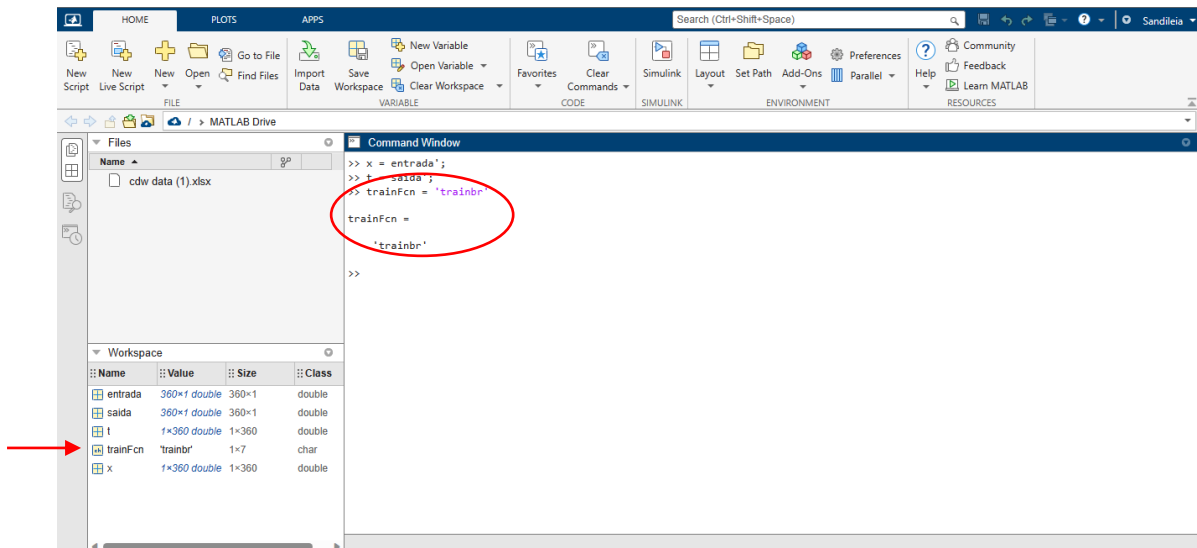
The screenshot shows the MATLAB Import Wizard interface. The 'Name' field is 'cdwData_1_S2' and the 'Type' is 'Numeric Matrix'. The spreadsheet preview shows the following data:

	N	O	P	Q	R	S	T	U	V	W	X	Y
Var14	Var15	Var16	Var17	Var18	Var19	Var20	Var21	Var22	Var23	Var24	Var25	
Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
1												
2	Valores se...											
3	Area M²	Residuo M²	Taxa	Mediana	L-Sup	L-Inf			área	res		
4	906	3085	0.29368	2.111	5.1581	-0.93615			906	3085		572
5	11689	27340	0.42754	2.111	5.1581	-0.93615			1401.21	701		643.28
6	17891.74	32976.1	0.54257	2.111	5.1581	-0.93615			1505	572		701
7	18040.61	32891	0.5485	2.111	5.1581	-0.93615			2108	847		847
8	11297.79	16127.14	0.70055	2.111	5.1581	-0.93615			2700	1012		872
9	21163	29577	0.71552	2.111	5.1581	-0.93615			2840	643.28		877
10	20535.95	26010.06	0.78954	2.111	5.1581	-0.93615			2896	1550		909
11	9943.9	11098	0.89601	2.111	5.1581	-0.93615			3068	1001		954.5

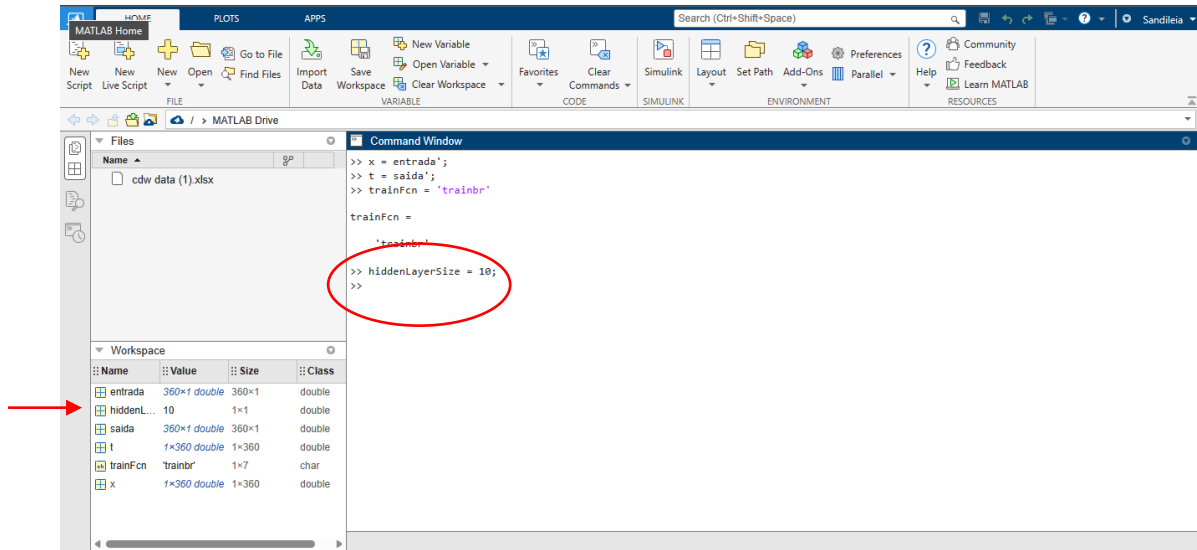
5 - Na *Command Window*, para o arquivo de entrada é necessário digitar `x = nomedoarquivo'`; e pressione *Enter*, para o arquivo de saída, digite `t = nomedoarquivo'`; e pressione *Enter*.



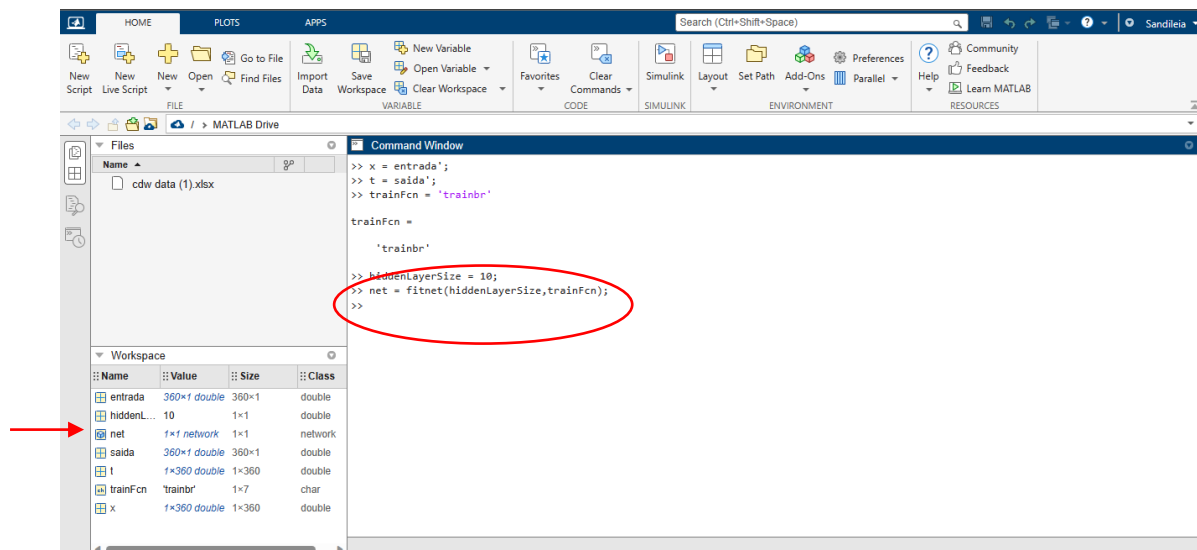
6 - Para utilizar o algoritmo *Bayesian Regularization* é necessário digitar o comando `trainFcn = 'trainbr'`. Caso seja utilizado o algoritmo *Levenberg-Marquardt* o comando é `trainFcn = 'trainlm'`; ou *Scaled Gradient*, o comando é `trainFcn = 'trainscg'`.



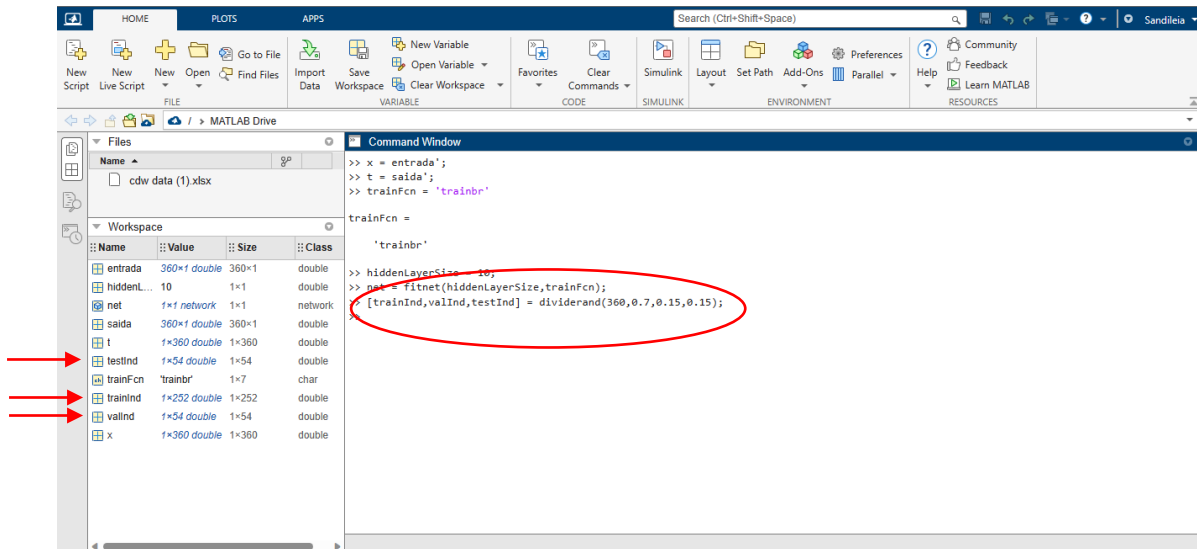
7 - Em seguida, deve ser digitado `hiddenLayerSize = 10;`, dez é a quantidade de neurônios na camada oculta.



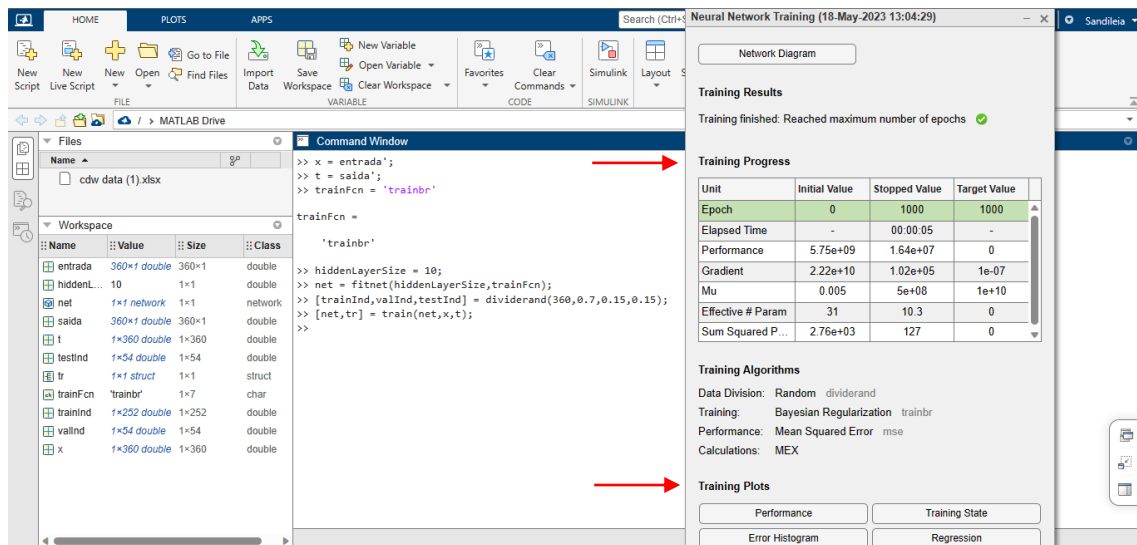
8 - Para criar a rede neural é necessário digitar `net = fitnet(hiddenLayerSize,trainFcn);`.



9 - Para dividir a amostra em treinamento, validação e teste, é necessário digitar `[trainInd,valInd,testInd] = dividerand(360,0.7,0.15,0.15);`. Os números indicam o tamanho da amostra e as porcentagens utilizadas, nesse caso 70%, 15% e 15%.

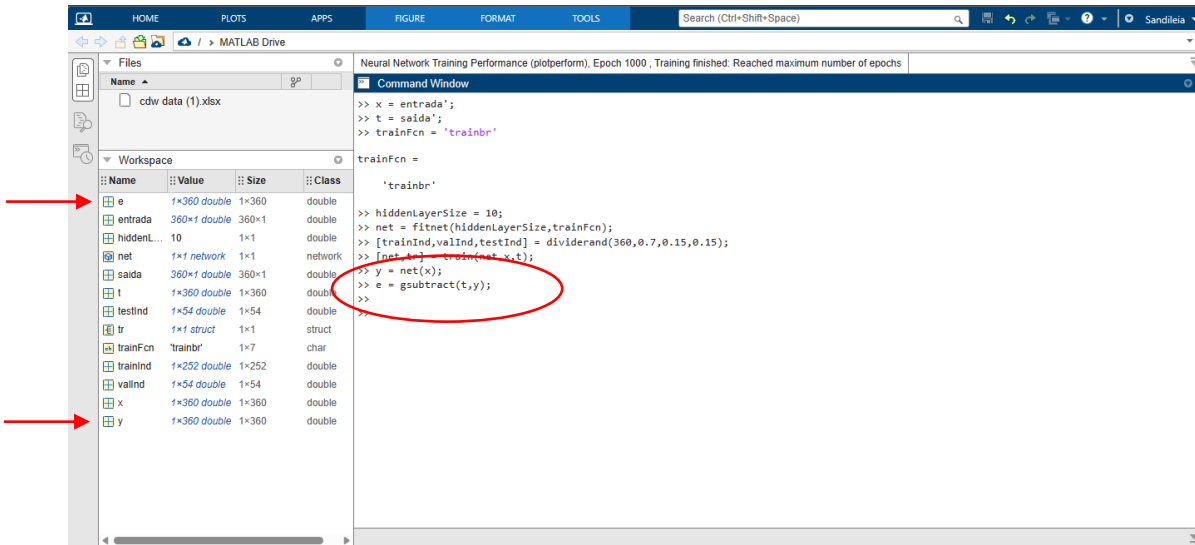


10 - Para treinar a rede neural é necessário digitar `[net,tr] = train(net,x,t);`. A rede será treinada em poucos segundos, e abrirá a tela a seguir.



Na tela *Neural Network Training*, é possível verificar o progresso do treinamento e obter os gráficos de regressão linear, histograma de erro, erro quadrático médio e diagrama da rede neural.

11 - Em seguida, o comando `y = net(x);` informa os valores estimados pela rede neural, e o comando `e = gsubtract(t,y);` informa os erros entre os valores esperados e os valores estimados.



Todos os dados que estão no *Workspace* podem ser abertos dentro do programa ou exportados, bem como as imagens podem ser salvas no *desktop*. Este passo a passo representa um ciclo de treinamento, que pode ser repetido quantas vezes for necessário. Também é possível alterar o algoritmo de treinamento, quantidade de neurônios na camada oculta ou divisão dos conjuntos de dados, a cada ciclo.