

## Uso da técnica de vizinhança e rede neural artificial na estimativa de necessidade de calagem

### Use of the neighborhood technique and artificial neural network in estimating the need for liming

DOI:10.34117/bjdv8n8-001

Recebimento dos originais: 21/06/2022

Aceitação para publicação: 29/07/2022

#### **Carlos Eduardo Bento Barbosa**

Mestrado

Instituição: Faculdade Anhanguera de Anápolis e Faculdade Metropolitana de Anápolis

Endereço: Av. Fernando Costa, nº 49, Vila Jaiara St. Norte, Anápolis

E-mail: carloskun.ce@gmail.com

#### **Ivandro José de Freitas Rocha**

Mestre

Instituição: Faculdade Evangélica de Goianésia (FACEG)

Endereço: Av. Brasil, Nº 1000, Covoá, Goianésia - GO, CEP: 76385-608

E-mail: ivandro.rocha@faceg.edu.br

#### **Guiliano Rangel Alves**

Mestre

Instituição: Universidade Estadual de Goiás (UEG)

Endereço: Universidade Estadual de Goiás, Campus Central Anápolis, BR 153, KM 99, Zona Rural, Anápolis - GO, CEP: 75.132-903

E-mail: guiliano@gmail.com

#### **Sabrina Dias de Oliveira**

Bacharel em Engenharia Agrícola

Instituição: Universidade Estadual de Goiás

Endereço: Rodovia BR-153, Quadra Área, Km 99, Fazenda Barreiro do Meio, Anápolis - GO, CEP: 75132-400

E-mail: sabrina9632oliveira@gmail.com

#### **Francisco Ramos de Melo**

Doutor (PhD)

Instituição: Universidade Estadual de Goiás (UEG)

Endereço: Universidade Estadual de Goiás, Campus Central Anápolis, BR 153, KM 99, Zona Rural, Anápolis - GO, CEP: 75132-903

E-mail: francisco.melo@ueg.br

#### **Sueli Martins de Freitas Alves**

Doutora

Instituição: Universidade Estadual de Goiás (UEG)

Endereço: Universidade Estadual de Goiás, Campus Central Anápolis, BR 153, KM 99, Zona Rural, Anápolis - GO, CEP: 75.132-903

E-mail: suelifreitas@ueg.br

## RESUMO

O emprego de novas tecnologias tem sido de grande importância nas investigações agronômicas, como a Inteligência Artificial (IA), o que pode representar uma estratégia eficiente para simulação e otimização de processos importantes na agricultura. O estudo tem como objetivo reduzir a quantidade de amostras necessárias para um possível mapeamento da necessidade de calagem do solo. A coleta dos dados foi obtida a partir de um esquema de grade amostral de 60x60 m. Os dados da coleta foram utilizados para treinamento da ANN, empregando o software Octave. Foi realizada uma adaptação da técnica de vizinhança para possibilitar a redução da quantidade de amostras. O desempenho nas fases de validação externa, calibração e treinamento foram utilizados como critério para definir a melhor ANN, no qual a ANN1 com exceção do teste e validação externa que se obteve desempenho muito bom, apresentou desempenho ótimo nas demais fases de calibração, tendo 6 neurônios na camada oculta e apresentando uma correlação acima de 89%. Por meio do Teste-t, foi mensurado que não houve diferença significativa das estimativas de necessidade de calagem usando ANN e o cálculo pelo método tradicional de necessidade de calagem. A técnica vizinhança adaptada para inserção de dados na ANN possibilitou diminuir em 30% o número de amostras a serem coletadas para um possível mapeamento da necessidade de calagem a taxa variada.

**Palavras-chave:** inteligência artificial, acidez do solo, agricultura de precisão.

## ABSTRACT

The use of new technologies has been of great importance in agronomic investigations, such as Artificial Intelligence (AI), which can represent an efficient strategy for simulation and optimization of important processes in agriculture. The study aims to reduce the amount of samples needed for a possible mapping of the need for soil liming. Data collection was obtained from a sampling grid scheme of 60x60 m. The data collection was used for ANN training, using the Octave software. An adaptation of the neighborhood technique was performed to allow the reduction of the amount of samples. The performance in the phases of external validation, calibration and training were used as criteria to define the best ANN, in which the ANN1 with the exception of the test and external having 6 neurons in the hidden layer and showing a correlation above 89%. Through the t-test, it was measured that there was no significant difference in the estimates of need for liming using ANN and the calculation by the traditional method of need for liming. The neighborhood technique adapted for data insertion in the ANN made it possible to reduce by 30% the number of samples to be collected for a possible mapping of the need for liming at varied rate.

**Keywords:** artificial intelligence, soil acidity, precision agriculture.

## 1 INTRODUÇÃO

O solo é considerado um corpo natural, resultante da decomposição e desintegração das rochas, cuja formação está inteiramente ligada ao processo de intemperismo físico e químico (SILVA et al., 2020). O bioma Cerrado é formado principalmente pelos Latossolos que são considerados solos altamente intemperizados, e estima-se que cerca de 43,39% da área total da região seja ocupada pelo mesmo (IBGE,

2021; SANO et al., 2019). Além desses solos serem altamente lixiviados, possuem elevada acidez e saturação por alumínio, baixa capacidade de retenção de cátions e alta capacidade de fixação de fósforo (LOPES e GUILHERME, 2016).

A acidez do solo é uma das características que mais restringe a produtividade das culturas, isso devido a mesma limitar a disponibilidade de nutrientes essenciais ao desenvolvimento das plantas (VIEIRA e WEBWE, 2017). A natureza da relação entre o rendimento da produção e o pH do solo tem implicações importantes para a sustentabilidade e a eficiência da produção agrícola (HOLLAND et al., 2019). É uma realidade bem conhecida que a acidificação do solo diminui o pH do solo, causando efeitos adversos nas plantas e microrganismos do solo (KUNHIKRISHNAN, et al., 2016).

Existem diversas explicações para a acidificação do solo como o material de origem, uso intensivo de fertilizantes, decomposição da matéria orgânica, dissociação de  $H_2CO_3$  devido a água da chuva, retirada de nutrientes devido a cultura, entre outros (GALINDO et al., 2017). Um método eficiente e viável, que possibilita a correção deste problema, consiste na adição de calcário ao solo por meio da prática da calagem (MELO et al., 2019).

O manejo correto de grandes áreas, possui um elevado custo de produção, fazendo com que muitos produtores acabem não realizando. A partir deste pressuposto, diversas estratégias vêm sendo desenvolvidas ao longo das últimas décadas com o objetivo de otimizar o uso de recursos, aumentar a produtividade e diminuir os impactos ambientais (FERRAZ et al., 2012). Com isso tem-se a agricultura de precisão (AP) que apresenta um conjunto de ferramentas que possibilita a otimização dos processos produtivos (OLIVEIRA et al., 2020).

Dentre as ferramentas da AP, tem-se a amostragem georreferenciada que propicia a caracterização da variabilidade espacial do solo. No entanto, um dos fatores limitantes para avaliar a variabilidade espacial é a quantidade de amostras de solo necessárias para a análise geoestatística, pois o número de pontos e conseqüentemente a distância entre eles, são fundamentais para gerar um semivariograma de confiança (MENDES et al., 2019). Logo, os produtores são limitados pelos recursos gastos na amostragem e mão de obra de trabalho, o que os leva a fazer apenas a coleta de amostras para a análise convencional. Entretanto, a análise convencional não permite a geração de semivariogramas (SOUZA et al., 2014).

Todavia, o emprego de novas tecnologias tem sido de grande importância nas investigações agrônomicas, como a Inteligência Artificial (IA), que pode representar uma estratégia eficiente para simulação e otimização de processos importantes na agricultura (ÇELEBI et al., 2017). As redes neurais artificiais (ANNs) são sistemas modelados matematicamente para reproduzir a inteligência de seres vivos, a fim de resolver situações que necessitem do armazenamento de conhecimento e que adquiram conhecimento através da experiência, tornando-as capazes de solucionar problemas muito complexos em diversas áreas do conhecimento (SOARES et al., 2015; BRAGA, 2014). Tal capacidade de resolver problemas tem sido evidenciada nas ciências agrárias. Bittar et al. (2018) desenvolveu ANNs para estimar os atributos do solo com um menor número de amostras em relação ao tradicionalmente usado para a análise geoestatística, o autor obteve resultados satisfatórios, sugerindo pesquisas futuras. Estimativas válidas de produtividade das culturas do milho (*Zeamays L.*) e da soja (*Glycine max L.*) foram obtidas por meio de aplicações de redes neurais artificiais (LEAL et al., 2015; ALVES et al., 2018). Meneses et al. (2020) conseguiram estimar a evapotranspiração potencial por Penman-Monteith usando redes neurais com inputs de temperaturas do ar.

O estado de Goiás se destaca por grandes produções agrícolas, estudos voltados para a potencialização da produção possuem grande importância, sendo a otimização do tempo e a redução dos custos de produção uma maneira de contribuir economicamente para o estado. Dessa maneira, o emprego de redes neurais artificiais demonstram ser uma interessante maneira de aprimorar os fatores de produção relativo a solos no estado de Goiás.

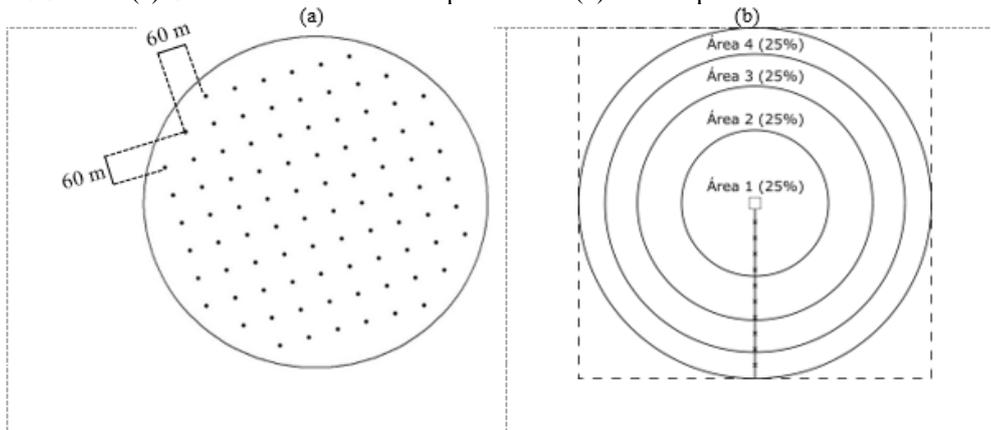
Desta forma, objetivou-se com este trabalho adaptar a técnica de vizinhança para reduzir a quantidade de amostras necessárias para viabilizar um possível estudo de mapeamento da necessidade de calagem a taxa variada em áreas de pivô central utilizando Redes Neurais Artificiais (ANN).

## 2 MATERIAL E MÉTODOS

O experimento foi desenvolvido em uma área comercial de tomate industrial irrigado por pivô central correspondente a 35 ha localizada no município de Vila Propício com de altitude 686 m, cujas coordenadas geográficas são: latitude 15°24'12" S e longitude 48°53'59.69" W. O solo pode ser classificado como um latossolo vermelho de textura argilosa, com consistência plástica e pegajosa.

A amostragem de solo para determinação de atributos químicos foi feita segundo grade amostral de 60x60m (Figura 1a) conforme metodologia de Bittar et al. (2018). Cada ponto foi georreferenciado utilizando aparelho de Sistema de Posicionamento Global (GPS), da marca Garmin modelo Etrex Legend RoHs (erro  $\pm 3m$ ), com o datum ajustado ao sistema SIRGAS 2000. Como sugerido por Freitas (2018) a área da lavoura foi dividida em 4 áreas (Figura 1b), com cada um desses representando 25% da lavoura, com o intuito de controlar o limite espacial para os pontos amostrais.

FIGURA 1. (a) Grade amostral da área experimental. (b) Área experimental dividida em áreas.



Foram determinados os atributos físicos e químicos do solo, no qual em cada ponto da grade amostral georreferenciada devidamente estaqueado, foram coletadas amostras de solo (composta por 5 subamostras) na camada de 0,00 a 0,20 m, com um trado de rosca, coletadas num raio de 3m, em torno de cada ponto para a determinação dos seguintes atributos: pH, potássio (K), fósforo (P), cálcio (Ca), magnésio (Mg), alumínio trocável (Al), acidez potencial (H+Al), matéria orgânica (MO), saturação por alumínio (m%), capacidade de troca catiônica (CTC) e saturação por bases (V%).

Para estimar a necessidade de calagem, foi utilizado o método de saturação por bases, conforme a Equação 1 adotada por Primavesi et al. (2004):

$$NC = \frac{(v_2 - v_1) \times T}{100} \times f \quad (1)$$

No qual:

NC = Necessidade de calagem  $\text{ton ha}^{-1}$ ;

$V_2$  = Saturação por bases recomendada para a cultura do tomate (V%);

$V_1$  = Saturação por bases atuais do solo (V%);

T = Capacidade de Troca de Cátions (CTC)  $\text{pH } 7,0 \text{ cmolc dm}^{-3}$  [ $\text{Ca}^{2+} + \text{Mg}^{2+} + \text{K} + (\text{H} + \text{Al}^{3+})$ ];

f = Fator de correção 100/PRNT.

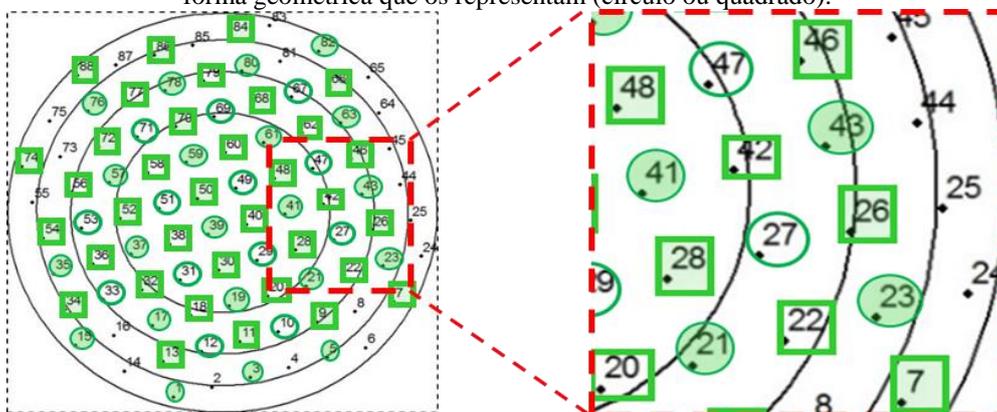
Para realizar a preparação e treinamento das ANNs foi utilizado o software GNU Octave versão 6.4.0 e uma adaptação do programa desenvolvido por Rocha (2019). Para se realizar o treinamento das ANNs, foi gerado um arquivo xls para inserção de dados na rede. O arquivo xls foi composto por 26 linhas, cada uma delas representando um ponto que a ANN estimou. Em 87 colunas estão as variáveis de entrada além de a última coluna ser preenchida pelos dados da variável de saída.

O treinamento foi realizado de acordo com a metodologia de Braga et al. (2014), que utiliza um algoritmo de retropropagação de erros (*back-propagation*), com topologia do tipo *Perceptron* de Múltiplas Camadas (MLP) para a aprendizagem das ANNs.

O conjunto de variáveis foi separado em dois tipos de variáveis, variáveis de entrada e variáveis de saída. As variáveis de entrada foram os dados inseridos na camada de entrada da ANN. Para escolha das variáveis de entrada a serem inseridas foi utilizada uma adaptação da técnica de vizinhança, normalmente utilizada no processamento digital de imagens (GONZALES, 2010).

A técnica da vizinhança define quais serão os pontos que fornecerão os dados para compor as variáveis de entrada e quais pontos irão compor as variáveis de saída. Como pode-se observar na Figura 2, os pontos destacados com os círculos e quadrados preenchidos pela cor verde representam os pontos que são referências para serem inseridos na camada de entrada da ANN, já os pontos destacados por quadrados e círculos sem preenchimento, serão os pontos a terem sua necessidade de calagem estimados pela ANN, ou seja, dão origem aos dados que compõem as variáveis de saída.

Figura 2. Esquema representativo da adaptação da técnica de vizinhança dos pontos, onde os pontos preenchidos da cor verde são pontos utilizados para estimar os pontos sem preenchimentos, equivalentes a forma geométrica que os representam (círculo ou quadrado).



Com este arranjo para inserir os dados na ANN, consegue-se diminuir o número de amostras coletadas em cerca de 30%, uma vez que de 88 pontos foram utilizados apenas 62 pontos para estimativa.

Sendo assim, a camada de entrada é composta pelas seguintes variáveis:

- Como sugerido por Rocha et al. (2022) a variável área foi definida como as variáveis A1, A2 e A3 para a área, sendo que estas assumem valores de 0 ou 1.
- Os valores de DX, DY e DA tanto dos pontos preenchidos de verde quanto aos pontos sem preenchimento (Figura 2). Nesta metodologia, a rede pode ser aplicada em outros pivôs, não ficando restrita a somente à área de estudo. Os valores de DX, DY e DA foram obtidos pela diferença das coordenadas UTM: X (longitude m) e Y (latitude m), e altitude (m) de cada ponto em relação ao centro do pivô, respectivamente.
- Os atributos químicos do solo em cada um dos pontos: pH ( $\text{CaCl}_2$ ), potássio ( $\text{K cmolc/dm}^3$ ), fósforo ( $\text{P cmolc/dm}^3$ ), cálcio ( $\text{Ca cmolc/dm}^3$ ), magnésio ( $\text{Mg cmolc/dm}^3$ ), alumínio trocável ( $\text{Al cmolc/dm}^3$ ), acidez potencial ( $\text{H+Al } \%$ ), matéria orgânica (MO%), saturação por alumínio (m%), capacidade de troca catiônica (CTC  $\text{cmolc/dm}^3$ ), saturação por bases (V%).

Todas as variáveis foram disponibilizadas para a rede conforme a vizinhança dos pontos aos quais se desejou estimar, escolhidos aleatoriamente conforme a separação dos conjuntos.

As variáveis de saída foram os dados inseridos na camada de saída da ANN. Para compor a variável de saída, tem-se a necessidade de calagem ( $\text{ton ha}^{-1}$ ) calculada referentes aos pontos que foram estimados, destacados por círculos e quadrados sem preenchimento na Figura 2.

De acordo com Alves et al. (2018), os dados devem ser separados de forma aleatória em três conjuntos tanto para as variáveis de entrada quanto para as de saída:

- Conjunto de treinamento: 17 pontos, sendo 65% do total de pontos amostrais (26 pontos);
- Conjunto de validação: 6 pontos, sendo 25% do total de pontos amostrais;
- Conjunto de teste: 3 pontos, sendo 10% do total de pontos amostrais.

Para que os dados possam ser inseridos na ANN torna-se necessário um pré-processamento, onde essas variáveis foram normalizadas para valores numa faixa entre -1 e 1, o que garante a equalização dos dados (ZARIFNESHAT et al., 2012). A transformação foi realizada por meio da Equação 2 (SRINIVASAN et al., 1994):

$$y = \frac{(x - x_{\min})(d_2 - d_1)}{x_{\max} - x_{\min}} + d_1 \quad (2)$$

No qual:

y = resultado normalizado;

x = valor a ser normalizado;

$x_{\min}$  = valor mínimo de x;

$x_{\max}$  = valor máximo de x;

$d_1$  = limite inferior do resultado da normalização (-1);

$d_2$  = limite superior do resultado da normalização (1).

A seleção da melhor ANN foi feita analisando os dados dos treinamentos consolidados e das validações externas, por meio das planilhas eletrônicas geradas pelas funções de treinamento e validação. Para determinar a rede com melhor desempenho foram utilizados os seguintes critérios:

- Maior índice de desempenho (id) no treinamento e na validação externa;
- Maior coeficiente de determinação ( $R^2$ );
- Menor erro quadrático médio (EQM).

Para as redes com melhor desempenho, foi executada a função de geração de gráficos, gerando dois arquivos: um com o gráfico de treinamento e outro com os gráficos do R linear dos conjuntos de treinamento, validação, teste e geral.

Após o treinamento e validação das ANNs, foi necessário realizar a desnormalização dos dados, para efeitos de comparação dos resultados estimados pelas ANNs, conforme Equação 3:

$$x = \frac{(x_{\max} - x_{\min}) * (y - d_1)}{d_2 - d_1} + x_{\min} \quad (3)$$

No qual:

y = resultado desnormalizado;

x = valor normalizado;

$x_{\min}$  = valor mínimo de x;

$x_{\max}$  = valor máximo de x;

$d_1$  = limite inferior do resultado da normalização (-1);

$d_2$  = limite superior do resultado da normalização (1).

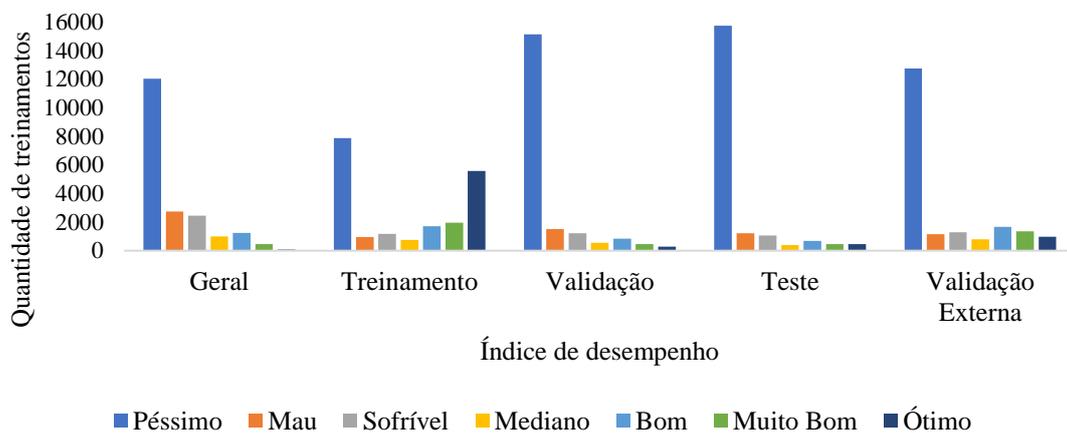
Foram realizados 1.000 treinamentos para cada número de neurônios da camada oculta que variaram de 1 a 20, totalizando 20.000 ANNs. Para verificação da eficiência das estimativas obtidas pelas ANNs, foram utilizados os seguintes indicadores: Erro médio relativo (P%), Coeficiente de correlação (r) (SCHNEIDER, 1998), Coeficiente de determinação ( $R^2$ ), Índice de concordância (c) (WILLMONT, 1981) e Índice de desempenho (Id) conforme Camargo e Sentelhas 1997. Com objetivo de testar se há

indícios de diferenças dos valores originais em relação aos valores estimados por meio das ANNs, foi realizado o Teste-t.

### 3 RESULTADOS E DISCUSSÃO

Pode-se observar na Figura 3 o Id de todas as ANNs em cada fase de calibração, além da validação externa, nota-se que o número de Id classificados como “Péssimo” se destaca com mais de 15.000 redes para fases de calibração, validação e teste, mais de 12.000 para validação externa e mais de 7.000 para o treinamento. Freitas (2019) encontrou resultados parecidos para estimativas do índice de área foliar do tomate industrial, por meio de ANNs.

FIGURA 3. Índice de Desempenho (id) das fases de treinamento (Geral, Treinamento, validação e teste) e validação externa dos dados de Necessidade de Calagem.



O primeiro critério avaliado para a escolha da rede foi o desempenho na validação externa, uma vez que este se remete a capacidade da ANN de generalizar o problema (BRAGA et al., 2012). Na fase de validação externa 980 redes obtiveram um desempenho ótimo e pouco mais de 1300 redes obtiveram um desempenho muito bom. O segundo critério avaliado foi o desempenho nas fases de calibração do treinamento, tendo como meta encontrar desempenhos entre ótimo e muito bom. Com o segundo critério foi possível filtrar as redes e encontrar apenas duas que atendem a estes pressupostos, como pode-se observar na Tabela 1.

TABELA 1. Resultados dos indicadores estatísticos para estimativas da necessidade de calagem das redes selecionadas.

RNA	Neur.	Calib.	EQM	P(%)	R <sup>2</sup>	r	c	id	Desempenho
1	6	Geral	0.025	25.055	0.908	0.953	0.974	0.929	Ótimo
		Treinamento	0.006	13.354	0.986	0.993	0.994	0.987	Ótimo
		Validação int.	0.040	34.639	0.865	0.930	0.937	0.871	Ótimo
		Teste	0.091	67.587	0.795	0.891	0.912	0.813	Muito bom
		Validação ext.	0.096	179.099	1.000	0.999	0.828	0.828	Muito bom
2	12	Geral	0.048	51.882	0.820	0.906	0.948	0.858	Ótimo
		Treinamento	0.036	39.332	0.907	0.952	0.967	0.921	Ótimo
		Validação int.	0.045	55.330	0.811	0.900	0.925	0.833	Muito bom
		Teste	0.108	108.886	0.991	0.995	0.813	0.809	Muito bom
		Validação ext.	0.011	30.163	0.998	0.992	0.954	0.952	Ótimo

ANN: rede neural artificial; Neur.: número de neurônios na camada oculta; Calib.: fase de calibração; EQM: erro quadrático médio; P(%): erro médio relativo; R<sup>2</sup>: coeficiente de determinação; id: índice de desempenho classificado de acordo com os critérios de classificação de Camargo e Sentelhas (1997); Trein.: Fase de treinamento; Val. Int.: fase de validação interna; Val. Ext.: fase de validação externa.

Também pode-se observar na Tabela 1, os valores referentes a todos os indicadores estatísticos que são fundamentais para a escolha da melhor rede. Com exceção do teste e validação externa que se obteve desempenho muito bom, a ANN1 obteve desempenho ótimo nas demais fases de calibração, tendo 6 neurônios na camada oculta. A ANN1 também obteve um baixo erro aleatório (EQM) em todas as fases de calibração, que com exceção da validação externa obteve vantagem sobre a ANN2, ou seja, com valores de EQM mais baixos, mostrando que o grau de dispersão dos dados em relação à média foi menor na ANN1.

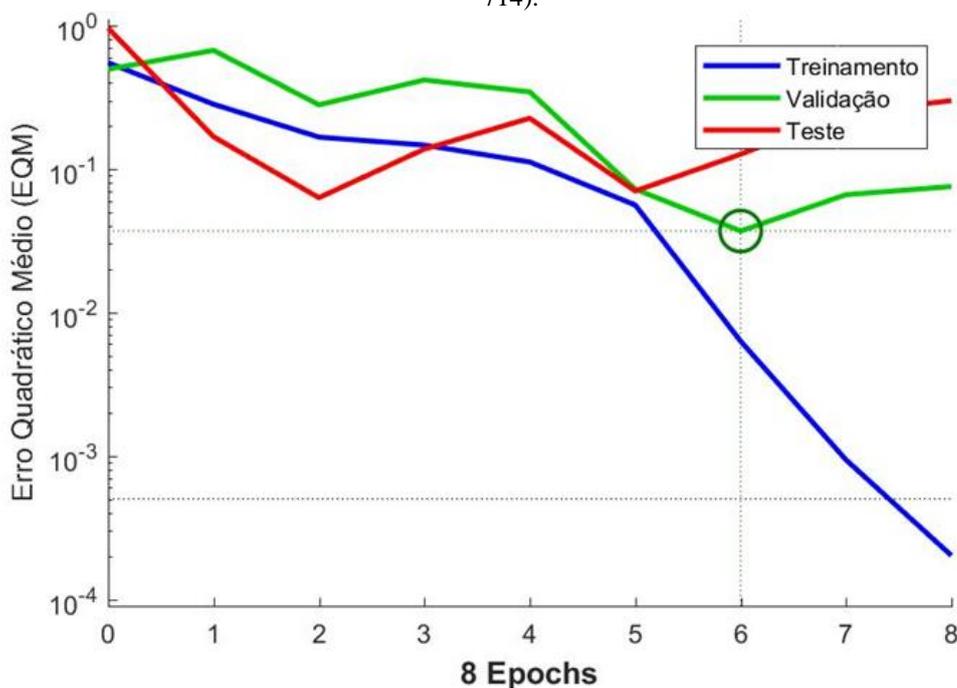
Os coeficientes de correlação (r) foram todos próximos de 1, indicando que além de uma relação de linearidade muito forte, existe uma tendência crescente, para todos os conjuntos analisados (Tabela 1). Neste indicador a ANN1 também mostrou vantagens quando comparada a ANN2 que com exceção do teste na fase de calibração, obteve coeficientes r mais próximos de 1 em todas as demais fases.

Por fim, outro fator importante para escolha da ANN é o número de neurônios na camada oculta, que por sua vez quanto menor este número maior é a capacidade de generalização do problema, pois segundo Braga et al. (2012) redes com grande número de neurônios na camada oculta podem resultar em *overfitting*. Quanto maior a estrutura, maior o número de parâmetros livres ajustáveis e, conseqüentemente, maior a flexibilidade da rede, gerando redes que memorizam os dados de entrada. superdimensionadas.

A rede selecionada têm 6 neurônios e foi encontrada no treinamento 714. Ela necessitou de 8 ciclos para chegar à convergência estipulada no programa de treinamento

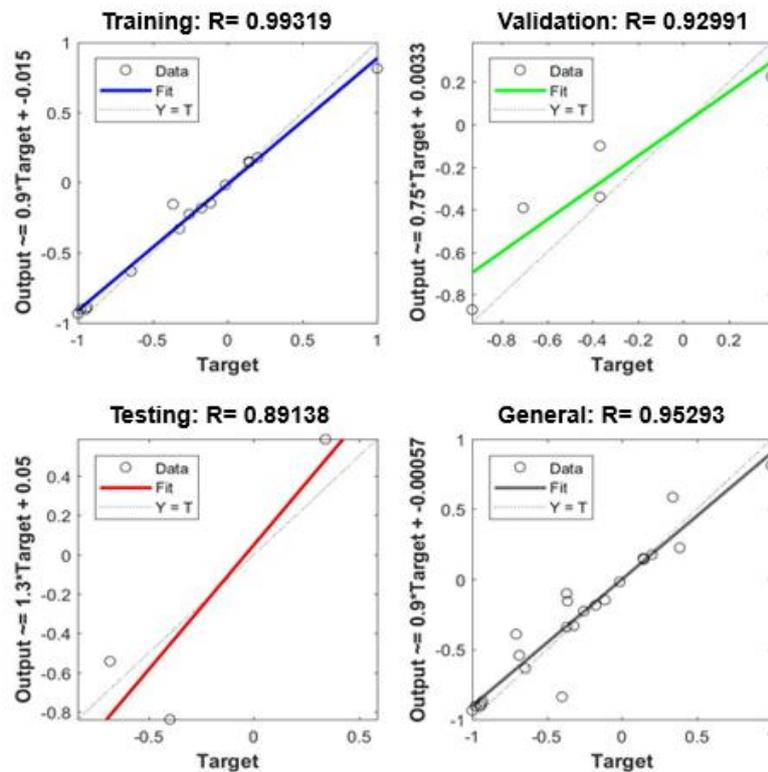
(Figura 4), interrompendo o treinamento também no ciclo 8. O treinamento da rede neural ajustou os valores dos 6 neurônios da camada oculta e os valores dos *bias* de cada um, assim como do neurônio de saída. Com um EQM baixo, pode-se evidenciar a alta precisão da rede.

FIGURA 4. Representação gráfica da fase de treinamento da rede selecionada (6 neurônios e treinamento 714).



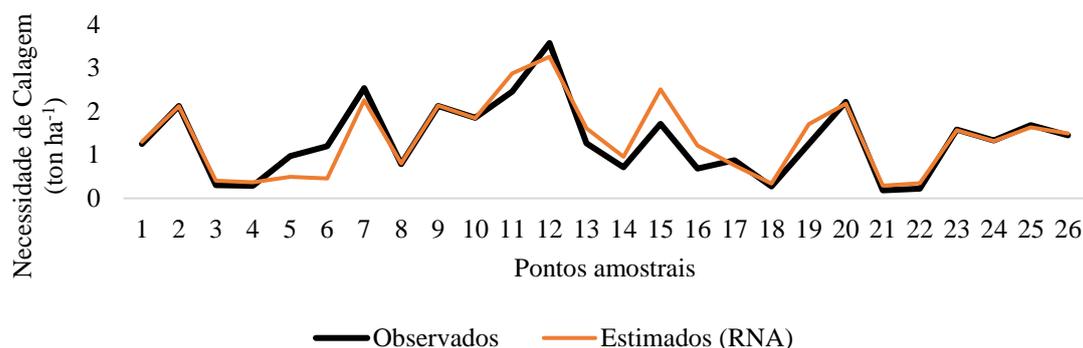
Na Figura 5 pode-se observar o gráfico de dispersão em todas as fases de treinamento, onde os dados de necessidade de calagem dos 26 pontos observados experimentalmente (calculados pelo método tradicional de saturação por bases) estão no eixo X, e os dados de necessidade de calagem dos 26 pontos estimados pela rede estão no eixo Y, além de evidenciar a capacidade que a ANN tem de estimar os dados com acurácia, uma vez que todos os coeficientes de correlação da fase de calibração foram maiores que 0,89, sendo assim também, mostrando uma forte tendência crescente.

FIGURA 5. Representação gráfica do coeficiente de correlação (r) da fase de treinamento (treinamento, validação, teste e geral) da rede selecionada com (6 neurônios e treinamento 714).



A Figura 6 apresenta duas linhas que representam os dados observados de necessidade de calagem e os dados estimados de necessidade de calagem, mais uma vez evidenciando o quanto a rede teve um baixo erro aleatório, uma vez que os valores estimados praticamente acompanham os valores observados (calculados pelo método tradicional de saturação por bases) para os 26 pontos amostrados. As estimativas realizadas por essa rede neural artificial encontraram erros aleatórios (EQM) entre 0,06% e 0,9% e podem ser consideradas muito boas, quando comparadas a outros trabalhos, como a estimativa de atributos químicos do solo apresentada por Freitas (2017), que encontrou erros aleatórios entre 8% e 16%.

FIGURA 6. Comparativo geral da rede selecionada para os valores observados (calculados pelo método tradicional de saturação por bases) e estimados para redução de amostras necessárias para coleta de amostras (2 neurônios e treinamento 574).



Por meio da técnica de seleção de variáveis a ANN utilizada para as estimativas conseguiu diminuir em torno de 30% o número de amostras em campo necessárias e com isso estimar os valores necessários para a análise geoestatística a serem usados para avaliar a variabilidade espacial na área de plantio e a partir dessa análise uma possível aplicação regionalizada de calcário. Segundo Molin et al., (2015) o número de amostras é imprescindível para a qualidade do mapa final e da sua capacidade de representar adequadamente a realidade, deste modo um fator limitante para a máxima eficiência econômica é a quantidade de pontos amostrais necessários para se conseguir realizar a aplicação de corretivo a taxa variada, sendo assim a principal vantagem deste modelo é reduzir de 88 para 62 amostras necessárias, uma vez que por meio de 62 amostras o modelo consegue estimar a necessidade de calagem nos 26 pontos restantes.

Quando comparados aos dados experimentais dos 26 pontos amostrais, os dados estimados pela ANN também se mostram acurados e precisos. Quando submetidos ao Teste-t a 5% de probabilidade não apresentou diferenças significativas (Tabela 2), desta forma pode-se inferir que a ANN conseguiu estimar a necessidade de calagem com um bom grau de confiança. Aitkenhead et al. (2012), em seus estudos conseguiram prever com um bom grau de confiança a composição química e física do solo utilizando informações de observações de campo em uma rede neural.

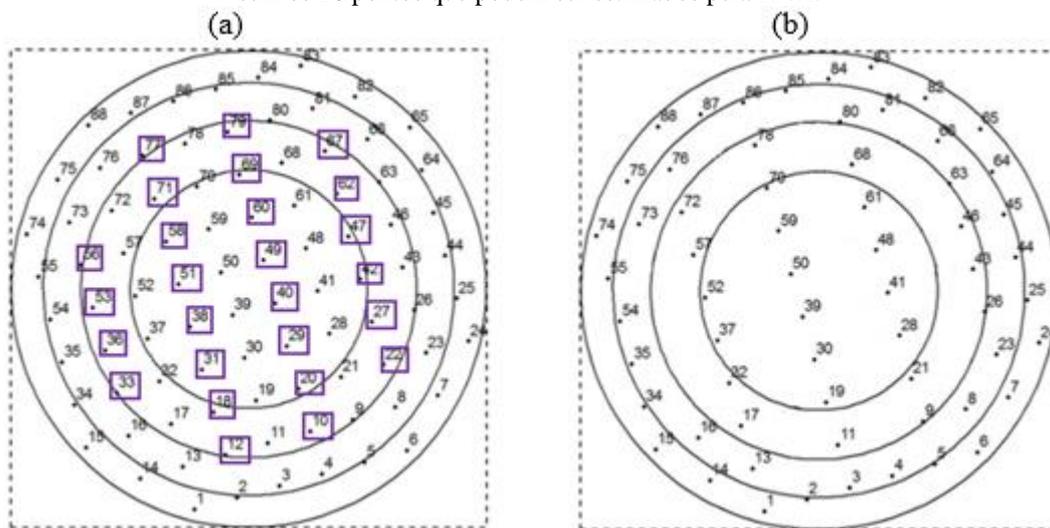
TABELA 2. Teste-t para os dados observados de necessidade de calagem em relação aos dados de necessidade de calagem estimados pela ANN1.

	Observado	Estimado (ANN)
Média	1,34	1,40
Variância	0,71	0,72
Observações	26	26
Gl	50	
p-value	0,41	

Observado: necessidade de calagem observada pelo método de saturação por bases; Estimado (ANN): necessidade de calagem estimada pela ANN1; Gl: graus de liberdade; p-value: teste de t-student significativo a 5% de probabilidade.

Desta forma a técnica de vizinhança que foi adaptada se mostrou bastante satisfatória, conseguindo diminuir consideravelmente o número de amostras a serem demandadas (Figura 7), tornando este modelo mais viável que o método tradicional, uma vez que reduzindo o número de amostras a serem coletadas, se reduz também a mão-de-obra, o tempo gasto e o custo das análises de solo, viabilizando ainda um possível mapeamento da necessidade de calagem a taxa variada. Entretanto faz-se necessário o aprimoramento da técnica para diminuir ainda mais o número de pontos a serem coletados.

FIGURA 7. (a) Esboço da grade amostral destacando os pontos que podem ser estimados pela Rede Neural Artificial (RNA) na área experimental. (b) Esboço da nova grade amostral da área experimental sem os 26 pontos que podem ser estimados pela RNA.



#### 4 CONCLUSÕES

A técnica de vizinhança dos pontos adaptada para inserção de dados na rede, conseguiu diminuir o número de amostras necessárias em aproximadamente 30%, para possíveis estudos visando a aplicação regionalizada de calcário.

A arquitetura que melhor estimou a necessidade de calagem para a redução de amostras necessárias para aplicação regionalizada foi de 6 neurônios na camada oculta, na repetição 714, com a correlação ( $r$ ) acima de 89%.

O trabalho não encerra um estudo mais aprofundado, já que o maior objetivo foi a verificação da proposta da inserção da RNA. Os dados obtidos se mostraram promissores o que incentiva mais estudos com maior espaço amostral e o aprimoramento da técnica de vizinhança.

Diante do exposto, é possível verificar que as Redes Neurais Artificiais se mostraram viáveis para a predição da necessidade de calagem, bem como sua contribuição na redução de amostras.

## REFERÊNCIAS

AITKENHEAD MJ, COULL MC, TOWERS W, HUDSON G, BLACK HIJ (2012) Predicting soil chemical composition and other soil parameters from field observations using a neural network. *Computers and Electronics in Agriculture*, Oxford, v.82, s/n, p.108–116. DOI: 10.1016/j.compag.2011.12.013.

ALVES GR, TEIXEIRA IR, MELO FR, SOUZA RTG, SILVA AG (2018) Estimating soybean yields with artificial neural networks. *Acta Scientiarum. Agronomy*, Maringá, e. 35250, v.40, n.1, mar. DOI: 10.4025/actasciagron.v40i1.35250.

BITTAR RD, ALVES SMDF, MELO FRDE (2018) Estimation of physical and chemical soil properties by artificial neural networks. *Revista Caatinga*, v31, n3, p704–712. DOI: 10.1590 / 1983-21252018v31n320rc.

BRAGA AP, CARVALHO APLF, LUDERMIR TB (2012) *Redes Neurais Artificiais: Teorias e Aplicações*. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A. 226p.

CAMARGO AP, SENTELHAS PC (1997) Avaliação do desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil. *Revista Brasileira de Agrometeorologia*, v.5, n.1, p. 89-97.

FERRAZ GAS, SILVA FM, COSTA PAN, SILVA AC, CARVALHO FM (2012) Agricultura de precisão no estudo de atributos químicos do solo e da produtividade de lavoura cafeeira. *Revista Coffee Science*, Lavras, v.7, n.1, p.59-67.

FREITAS EFM (2018) Rede neural artificial na predição de atributos físicos e químicos do solo, 2018. 103p. Dissertação (Mestrado em Engenharia Agrícola), Universidade Estadual de Goiás. Anápolis.

GALINDO FS, SILVA JC, GERLACH GAX, FERREIRA MMR, COLOMBO AS, TEIXEIRA FILHO MCM (2017) Comparação do calcário líquido com o calcário em pó na fertilidade do solo e nutrição do feijoeiro. *Cultura Agronômica, Ilha Solteira*, v.26, n.3, p.265-284.

GONZALEZ RC, WOODS RE (2010) *Processamento Digital de Imagens*, 3.ed. Pearson.

HOLLAND JE, WHITE PJ, GLENDINING MJ, GOULDING KWT, MCGRATH SP (2019) Yield responses of arable crops to liming – An evaluation of relationships between yields and soil pH from a long-term liming experiment. *European Journal of Agronomy*, v. 105, p. 176-188. DOI: 10.1016/j.eja.2019.02.016.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (2018) Banco de Dados de Informações Ambientais. Disponível em: <BDIA - Banco de Dados de Informações Ambientais (ibge.gov.br)>. Acessado em: 10 agosto de 2021.

KUNHIKRISHNAN A, THANGARAJAN R, BOLAN NS, XU Y, MANDAL S, GLEESON DB, SESHADRI B, BARTON B, TANG C, LUO J, DALALR DING W, KIRKHAM MB, NAIDU R (2016) Chapter One - Functional Relationships of Soil

Acidification, Liming, and Greenhouse Gas Flux. *Advances in Agronomy*, v. 139, p. 1-71. DOI: 10.1016/bs.agron.2016.05.001.

LEAL AJF, MIGUEL EP, BAILO FHR, NEVES DDEC, LEAL UAS (2015) Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo. *Bragantia*, v74, p436 – 444. DOI: 10.1590/1678-4499.0140.alve

LOPES AS, GUILHERME LRG (2016) A career perspective on soil management in the Cerrado Region of Brazil. In: SPARKS, D. (ed.). *Advances in Agronomy* 137. Cambridge: Academic Press, p. 1-72.

MELO RM, VIEIRA MC, CARNEVALI TO, GONÇALVES WV, TORALES EP, TOLOUEI SEL, SANTOS CC (2019) Calagem e textura do substrato afetam o desenvolvimento de *Campomanesia adamantium* (Cambess.) O. Berg. *Revista de Ciências Agrárias*, v.42, n.1, p.101-110. DOI: 10.19084/RCA18283.

MENDES A, SANTOS GR, EMILIANO PC, KALEITA A, FERREIRA MP (2019) Estimação do tamanho amostral na geoestatística usando um modelo de variograma gaussiano na presença de *outliers*. *Energia na Agricultura*, v.34, n.3, p.429-440. DOI: 10.17224/EnergAgric.2019v34n3p429-440.

MENESES KC, APARECIDO LEO, MENESES KC, FARIAS MF (2020) Estimando potencial de evapotranspiração no Estado do Maranhão usando redes neurais artificiais. *Revista Brasileira de Meteorologia*, v.35, n.4, p.675-682. DOI:10.1590/0102-77863540072.

MOLIN JP, AMARAL LR, COLAÇO AF (2015) *Agricultura de precisão*. 1. ed. São Paulo: Oficina de textos. 238p.

OLIVEIRA AJ, SILVA GF, SILVA GR, SANTOS AAC, CALDEIRA DAS, VILARINHO MKC, BARELLI MAA, OLIVEIRA TC (2020) Potencialidades da utilização de drones na agricultura de precisão. *Brazilian Journal of Development*, v.6, n.9, p.64140-64149. DOI: 10.34117/bjdv6n9-010.

PRIMAVESI O, PRIMAVESI AC, CÔRREA LA (2004) Calagem em pastagem de *Brachiaria decumbens* recuperada com adubação nitrogenada em cobertura. São Carlos: Embrapa 32p.

ROCHA IJF (2019) Estimativa do índice de área foliar para cultura irrigada por meio de pivô central utilizando imagens de sensoriamento remoto e redes neurais artificiais. 107P. Dissertação (Mestrado em Engenharia Agrícola), Universidade Estadual de Goiás. Anápolis, 2019.

ROCHA, IJF, ALVES, GR, BARBOSA, CEB, ALVES, SMF, RAMOS FR (2022) Redes neurais artificiais aplicadas na estimativa do índice de área foliar utilizando imagens de sensoriamento remoto. *Brazilian Journal of Development*, Curitiba, v.8, n.5, p.36460-36478. DOI: 10.34117/bjdv8n5-247.

SANO EE, ROSE R, SCARAMUZZA CAM, ADAMI M, BOLFE EL, COUTINHA AC, ESQUERDO JCDM, MAURANO LEP, NARVAES IS, OLIVEIRA FILHO FJB,

SILVA EB, VICTORIA DC, FERREIRA LG, BRITO JLS, BAYMA AP, OLIVEIRA GH, BAYMA-SILVA G (2019) Land use dynamics in the Brazilian Cerrado in the period from 2002 to 2013. *Pesquisa Agropecuária Brasileira*, v54, p.00138. DOI: 10.1590/S1678-3921.

SCHNEIDER PR (1998) *Análise de regressão aplicada à Engenharia Florestal*. Santa Maria: ed. UFSM/CEPEF. 236p.

SILVA MO, VELOSO CL, NASCIMENTO DL, OLIVEIRA J, PEREIRA DF, COSTA KDS (2020) Indicadores químicos e físicos de qualidade do solo. *Brazilian Journal of Development*, Curitiba, v. 6, n. 7, p. 47838-47855, jul. DOI: 10.34117/bjdv6n7-431.

SILVA J, GUEDES IMR, LIMA CEP (2012) Adubação e nutrição. IN: CLEMENTE, F.M.V.T.; BOITEUX, L (Eds.). *Produção de Tomate para processamento industrial*. Brasília DF: Embrapa. cap. 5, p. 105-127.

SOARES CS, ROBAINA AD, PEITER MX, RUSSI JL (2015) Predição da produtividade da cultura do milho utilizando rede neural artificial. *Ciência Rural*, Santa Maria, v.45, n.11, p.1987-1993. DOI:10.1590/0103-8478cr20141524.

SOUZA ZM, SOUZA GS, JÚNIOR JM, PEREIRA GT (2014) Número de amostras na análise geoestatística e na krigagem de mapas de atributos do solo. *Ciência Rural*, Santa Maria, v. 44, n. 2, p. 261-268. DOI: 10.1590/S0103-84782014000200011.

SRINIVASAN D, LIEW AC, CHANG CS (1994) A neural network short-term load forecaster. *Electric Power Systems Research*, Holanda, v.28, n.3, p.227-234. DOI: 10.1016/0378-7796(94)90037-X.

VIEIRA C, WEBER O (2017) Saturação por Bases no Crescimento e na Nutrição de Mudas de Ipê-Amarelo. *Floresta e Ambiente*, v.24, s/n. DOI: 10.1590/2179-8087.001916.

WILLMONT CJ (1981) On the validation of models. *Physical Geography*, Palm Beach, v.2, n. 2, p.184-194.

ZARIFNESHAT S, ROHANI A, GHASSEMZADEH HR, SADEGHI M, AHMADI E, ZARIFNESHAT M (2012) Predictions of apple bruise volume using artificial neural network. *Computers and Electronics in Agriculture*, Mashhad, v.82, n.6, p.75-86. DOI: 10.1016/j.compag.2011.12.015.