

# Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo

Aguinaldo José Freitas Leal (1\*); Eder Pereira Miguel (2); Fabio Henrique Rojo Baio (1); Danilo de Carvalho Neves (1); Ulcilea Alves Severino Leal (1)

(1) Universidade Federal de Mato Grosso do Sul (UFMS), Campus de Chapadão do Sul, Rodovia MS 306, 79560-000 Chapadão do Sul (MS), Brasil.

(2) Universidade de Brasília (UnB), Engenharia Florestal, 70910-900 Brasília (DF), Brasil.

(\*) Autor correspondente: [aguinaldo.leal@ufms.br](mailto:aguinaldo.leal@ufms.br)

Recebido: 6/abr./2015; Aceito: 19/jun./2015

## Resumo

O entendimento dos fatores que influenciam a produtividade é essencial para o sucesso produtivo e para adoção de manejo diferenciado em sítios específicos. Na busca de alternativas para prever a produtividade de grãos de milho a partir de atributos do solo, uma alternativa consiste no uso de redes neurais artificiais (RNAs). Diante disso, o presente estudo teve por objetivo avaliar a eficácia de adoção de atributos do solo por interface da análise de regressão, e das RNAs no estabelecimento de sítios de manejo diferenciado e predição da produtividade de grãos de milho, “segunda safra”, em solos de cerrado. Os dados foram obtidos em uma área de 41,76 ha, cultivada em 2010 e 2011. Apesar de demandar maior tempo de construção e processamento em relação à regressão linear, a adoção de RNAs permite melhor predição da produtividade de grãos. Em consonância ao estabelecimento de sítios específicos de manejo diferenciado do solo, a partir dos atributos teor de argila, capacidade de troca de cátions, matéria orgânica do solo e saturação de bases.

**Palavras-chave:** *Zea mays*, modelagem de biosistemas, zonas de manejo, agricultura de precisão.

## Artificial neural networks for corn yield prediction and definition of site-specific crop management through soil properties

### Abstract

The understanding of the factors influencing yield is essential for the crop success and adoption of site-specific management. The use of artificial neural networks (ANN) is an alternative of corn yield prediction from soil properties attributes. Thus, this study aimed to evaluate the effectiveness of adoption of soil properties by interface of the regression analysis, and ANNs in the establishment of site-specific management zones and prediction of corn yield, second crop in Cerrado's soil. Data were collected in an area of 41.76 ha cropped in 2010 and 2011. The adoption of ANNs allows better corn yield prediction despite of higher demand of construction time and processing when compared to linear regression. In consonance to the soil site-specific establishment from clay content, exchange cation capacity, organic matter and base soil saturation.

**Key words:** *Zea mays*, bioscience modeling, management zone, precision agriculture.

## 1. INTRODUÇÃO

O milho, apesar de apresentar baixa tolerância a déficit hídrico, é a cultura preferencial em “segunda safra” (“safrinha”) em condições de cerrado. Isso ocorre em função do maior valor agregado de seus grãos, em comparação ao sorgo. Entretanto, o sucesso dessa atividade é altamente dependente das condições de precipitação, capacidade de armazenamento de água e nutrientes no solo. Esses fatores são influenciados por atributos do solo, como o teor de

argila (Chen et al., 2014a) e teor de matéria orgânica (MO), que por sua vez se correlacionam com a capacidade de troca de cátions (Gruba & Mulder, 2015).

O entendimento dos fatores que influenciam a produtividade é essencial para o sucesso produtivo, principalmente em sistemas de produção modernos e em grande escala, como o cultivo de milho em “segunda safra” em Cerrado de altitude. Esse sistema carece da

adoção de manejo diferenciado, também conhecido como técnicas de agricultura de precisão, para otimizar a utilização de insumos, incrementando produtividade e lucro. Nesse sentido, a agricultura de precisão é definida como um conjunto de técnicas que permitem o gerenciamento localizado, e seu sucesso é dependente de três elementos: informação, tecnologia e gerenciamento (Srinivasan, 2006).

Com a crescente adoção de técnicas de manejo diferenciado do solo, como a adubação em taxa variável, possibilitadas pela adoção de máquinas específicas na agricultura de precisão, se faz necessário estabelecer sítios específicos de manejo (zonas de manejo diferenciado). Segundo Rodrigues et al. (2013), a gestão eficaz de sítios específicos de manejo requer uma compreensão do solo e fatores ambientais que influenciam a variabilidade da produção agrícola. Além de ser necessário avaliar as técnicas utilizadas para definir essas relações.

Na busca de entender como os atributos do solo afetam a produção, bem como a possibilidade de utilização destes na predição da produtividade de grãos, visando a definição de sítios específicos de manejo, planejamento dos investimentos, predição de produtividade e lucro, uma alternativa consiste no uso de redes neurais artificiais (RNAs). Essas redes são sistemas computacionais paralelos, constituídos por unidades de processamento simples, também denominadas neurônios artificiais, conectadas entre si de maneira específica para desempenhar determinada tarefa (Binoti et al., 2013).

As RNAs têm apresentado elevado desempenho devido a fatores como: estrutura robusta e paralelamente distribuída (camadas); eficiência no aprendizado e generalização, que as tornam capazes de resolver problemas complexos; são tolerantes a dados discrepantes ou “outliers”; podem modelar distintas variáveis e suas relações não lineares; e possibilitam a modelagem com variáveis categóricas, além das numéricas (Haykin, 2001).

Atualmente, trabalhos são desenvolvidos visando a sua adequação e parametrização para diversas situações, como as funções *tapper* ou afilamento da árvore (Leite et al., 2011), modelagem diamétrica (Binoti et al., 2013), modelo de crescimento para povoamento de *Eucalyptus* (Castro et al., 2013), estudos de características do solo (Soares et al., 2014) e, recentemente, para previsão de produtividade de grãos de trigo (Silva et al., 2014). Porém, em relação à estimativa da produtividade de grãos de milho baseado em parâmetros químicos e físicos do solo, visando estabelecer sítios específicos de manejo, o comportamento desta ferramenta não é conhecido.

Nesse contexto, há a hipótese de se estabelecer modelos biomatemáticos capazes de prever a produtividade do milho, cultivado em “segunda safra”, por meio de atributos do solo, utilizando modelos biomatemáticos de RNAs ou regressão múltipla. Além disso, esses modelos e atributos do solo podem ser utilizados para definição de sítios específicos de manejo. Diante disso, o objetivo deste estudo foi avaliar a eficácia do teor de matéria orgânica, capacidade de troca catiônica,

saturação de bases e teor de argila no estabelecimento de sítios de manejo diferenciado e na predição da produtividade de grãos de milho, por meio de interface da análise de regressão e das redes neurais artificiais (RNAs).

## 2. MATERIAL E MÉTODO

Os dados de produtividade dos grãos de milho se referem ao milho semeado em “segunda safra” (“safrinha”), nos anos de 2010 e 2011. As análises químicas do solo e produtividade de grãos se referem a uma área de 41,76 ha, localizado no município de Chapadão do Céu (GO), sudoeste de Goiás, na latitude 18,341 S e longitude 52,623 O. Essa área se encontra a 850 m de altitude, tendo o seu solo classificado como Latossolo Vermelho distrófico, segundo sistema brasileiro de classificação do solo (Santos et al., 2013). O mesmo foi cultivado em sistema plantio direto, sequência soja (verão) e milho “segunda safra”. As adubações de semeadura e cobertura foram baseadas em análise de solo da área experimental e recomendações para a cultura, quando cultivada em região do cerrado (Sousa & Lobato, 2004). A colheita, nos dois anos, foi realizada no mês de junho. Sendo a produtividade determinada através de colhedora John Deere, Modelo STS 9770. Equipada com sistema de mapeamento de colheita Harvest Doc™, o qual gera mapas de produtividade (kg ha<sup>-1</sup>) através da posição determinada pelo GPS e informações fornecidas pelos sensores instalados na colhedora, como o sensor de fluxo de grãos e o sensor de umidade (John Deere Co). O talhão foi subdividido em unidades amostrais básicas (parcelas) de 40 m × 40 m (1600 m<sup>2</sup>), nas quais foram determinadas as produtividades médias de grãos de milho, nos dois anos de cultivo (variável resposta), juntamente com as variáveis preditoras, representadas pelos atributos do solo: matéria orgânica (MO), Capacidade de troca catiônica (CTC), saturação de bases (V %); e teor de argila (TA). Obtidos através de análise de solo da camada de 0-0,2 m, segundo metodologia descrita em Silva (2009).

Para avaliar o grau de relacionamento entre as variáveis envolvidas no processo de modelagem, realizou-se análise de correlação de Pearson. Sequencialmente, os modelos pelas diferentes categorias de ajustes (redes neurais e regressão múltipla) foram ajustados. Para estimar a produtividade de milho, via regressão linear múltipla, adotou-se o método dos mínimos quadrados ordinários:

$$Y = \beta_0 + \beta_1 \times MO + \beta_2 \times CTC + \beta_3 \times V(\%) + \beta_4 \times TA. \quad (1)$$

Sendo Y = produtividade média de grãos de milho (kg ha<sup>-1</sup>), nos dois anos agrícolas avaliados; MO = teor de matéria orgânica no solo (mg dm<sup>-3</sup>); CTC = capacidade de trocas catiônicas (mmol<sub>c</sub> dm<sup>-3</sup>); V(%) = saturação de bases; TA = teor de argila (mg dm<sup>-3</sup>);  $\beta_i$  = estimadores dos parâmetros a serem ajustados, com  $i=0, 1, 2, 3$  e  $4$ .

No ajuste por redes neurais foram consideradas as mesmas variáveis. Todavia, as RNAs usam a inteligência artificial para resolver tais problemas, as quais são formadas por elementos de processamento simples, os neurônios artificiais, sendo estes ativados por uma função (função de ativação), obtendo uma única resposta. Neurônio artificial é a unidade de processamento de informação de uma RNA, constituído de “n” entradas  $x_1, x_2, \dots, x_n$  (dendritos) e uma saída  $y$  (axônio). As entradas são associadas a pesos  $w_1, w_2, \dots, w_n$  que representam as sinapses. Estas podem ser negativas ou positivas. Atualmente, um modelo básico de neurônio artificial pode ser representado matematicamente como:

$$Y_k = \varphi(V_k) \quad (2)$$

Em que:  $Y_k$  = saída do neurônio artificial;  $\varphi$  = função de ativação;  $V_k$  = resultado do combinador linear, ou seja:

$$V_k = \sum_{o=1}^m x_m \cdot w_m \quad (3)$$

Em que:  $V_k$  = combinador linear;  $x_m$  é o número de entradas; e  $w_m$  é o peso para cada entrada de  $m$ .

Para o ajuste das RNAs, as variáveis numéricas foram normalizadas linearmente no intervalo de 0 a 1. A camada de entrada foi constituída de quatro (4) neurônios, sendo um neurônio para cada variável numérica preditora (atributos do solo), em função da variável resposta/saída (produtividade de grãos de milho  $\text{kg ha}^{-1}$ ). As redes foram constituídas de apenas uma camada oculta, em que o número de neurônios nessa camada foi igual a quatro, logo, a arquitetura da rede foi 4-4-1, e como função de ativação, utilizou-se a sigmoide.

A função de ativação sigmoide é a mais habitual no desenvolvimento de redes neurais artificiais (Haykin, 2001), e matematicamente é dada por:

$$\varphi(v) = \frac{1}{1 + \exp(-\beta u)} \quad (4)$$

Em que:  $\varphi$  = função de ativação sigmoide;  $\beta$  = estimativa do parâmetro que determina inclinação da função sigmoide;  $u$  = potencial de ativação da função.

No presente estudo, o algoritmo de treinamentos utilizado foi o *resilient propagation*, como escolha mais hábil e recomendada para RNAs do tipo *Multilayer Perceptron*. No *resilient propagation* os pesos são fundamentados em informações dos dados presentes, para isso é introduzido para cada peso o valor individual de atualização. Inicialmente, os pesos de todas as redes foram gerados aleatoriamente. Sequencialmente, esse valor individual de atualização evoluiu durante o processo de aprendizado baseado na função do erro. As estimativas da produtividade de grãos foram simuladas com as combinações possíveis das variáveis de entrada, totalizando quatro combinações para a variável resposta.

Salienta-se que o aprendizado das redes foi do tipo supervisionado, com isso foram dados para a rede dois conjuntos de valores: o conjunto de valores de entrada e o conjunto de valores de saída. Desta forma, o treinamento consistiu em um problema de otimização dos parâmetros da rede (seus pesos sinápticos), para que pudessem responder às entradas conforme esperado, até que o erro entre os padrões de saída gerados pela rede alcançasse o valor mínimo desejado.

Foram treinadas cem redes do tipo *perceptrons* de múltiplas camadas, comumente conhecidas como MLP (*Multilayer Perceptron*). Como critério de parada do algoritmo de treinamento utilizou-se o número total de ciclos igual a 3.000 ou erro quadrático médio inferior a 1%, conforme sugeridos por Chen et al. (2014b). Para ajuste das redes neurais artificiais utilizou-se o programa *NeuroForest Star* (Binoti et al., 2013), e para o ajuste via regressão, como as correlações foram realizadas pelo *IBM SPSS Statistics 20*, quanto à confecção dos gráficos, utilizou-se o *Microsoft Office Excel*.

Na seleção da RNA e do modelo de regressão foram adotados os critérios tradicionais utilizados na verificação da qualidade de ajuste Rencher & Schaalje (2008): coeficiente de determinação ( $R^2$ ), erro padrão da estimativa em porcentagem ( $Syx\%$ ) e análise gráfica dos resíduos. Também foi utilizado o Critério de Informação de Akaike (AIC), que é uma medida da qualidade relativa do ajuste (Akaike, 1974), definido por:

$$AIC = -2 \ln(L_p) + 2p \quad (5)$$

Em que  $L_p$  é o valor que torna máxima a função de máxima verossimilhança do modelo estimado, e  $p$  é o número de parâmetros do modelo.

Por fim, utilizou-se o Critério de Informação Bayesiano (BIC); sua determinação, segundo Schwarz (1978), é definida pela expressão:

$$BIC = -2 \ln(L_p) + Kp \text{ Para } K = \ln(n) \quad (6)$$

onde  $K = \ln(n)$  e  $n$  o número de observações.

Após definidos os modelos com suas variáveis, do conjunto das 261 unidades amostrais básicas, 196 (75%) foram utilizadas para o ajuste e 65 (25%) foram separadas de forma aleatória para o teste de validação. Os critérios escolhidos para a validação dos ajustes, bem como a comparação entre a técnica de regressão e a de redes neurais artificiais, foi o Teste de Qui quadrado, o erro padrão da estimativa em porcentagem ( $Syx\%$ ), a diferença agregada em porcentagem ( $Da\%$ ) e o erro médio absoluto ( $Ei$ ).

### 3. RESULTADOS E DISCUSSÃO

Os valores médios, mínimos e máximos da produtividade de grãos de milho, MO, CTC, V(%) e teor argila, assim como suas medidas de dispersão, são apresentados na

tabela 1. De acordo com Gomes (2000), a variabilidade de qualquer atributo pode ser classificada segundo a magnitude do seu coeficiente de variação (CV). Se o CV for inferior a 10%, considera-se de baixa variabilidade, de 10% a 20%, média, de 20% a 30%, alta e maior que 30%, muito alta. Portanto, o teor de matéria orgânica no solo, assim como a CTC apresentaram baixa variabilidade (6,79% e 7,22% respectivamente). Contudo, Montanari et al. (2013) obtiveram valores superiores, entre 14% e 21%, na camada superficial e subsuperficial, respectivamente. Entretanto, no presente estudo a área é explorada em sistema plantio direto, sistema que pode ter contribuído para o fornecimento uniforme de material orgânico ao solo, e diminuído, conseqüentemente, a variação no teor de matéria orgânica e também na CTC.

Apesar do solo em questão ser intensamente utilizado na produção agrícola, a saturação de bases encontra-se com valor médio (V% de 39,04%) abaixo do recomendado para cultura do milho, em região de cerrado, que é 50% em cultivos de sequeiro (Sousa & Lobato, 2004). Incrementos nesse atributo são frequentemente relacionadas a variações na produtividade das culturas exigentes como o algodoeiro, cultivado no Cerrado (Freddi et al., 2014). Contudo a produtividade média de grãos (PG), nos dois anos de cultivo, foi de 5.571 kg ha<sup>-1</sup>. Produtividade essa superior à média nacional em cultivo de “segunda safra” (5.255 kg ha<sup>-1</sup>) e muito similar à média obtida na região Centro-Oeste em 2014 (5.516 kg ha<sup>-1</sup>) levantada pela CONAB (2014). Esse fato pode estar relacionado a uma possível menor exigência ou menor sensibilidade do milho cultivado em SPD (Sistema de Plantio Direto) à correção do solo. Hipótese essa apontada por Leal et al. (2013), ao verificar que modos de aplicação de calcário, em sistema de plantio direto, não incrementaram a produtividade e os teores médios de Ca e Mg em folhas de milho em relação ao tratamento controle, sem aplicação de calcário.

A produtividade de grãos apresentou correlação significativa com os quatro atributos do solo analisados (Tabela 1). Justificando, dessa forma, a utilização destes como variáveis preditoras da produtividade. Dalchiavon et al. (2011) também relataram efeito positivo dos teores de matéria orgânica na produtividade do feijoeiro, em região de cerrado em Latossolo

Vermelho distrófico. É provável que em solo cultivado em sistema de semeadura direta, como a área em questão, as correlações da MO com os demais atributos químicos do solo se manifestem mais intensamente, influenciando a disponibilidade e a absorção de nutrientes, a produtividade das culturas (Nicolodi et al., 2008), além de incrementar a capacidade de troca de cátions (CTC). Esse atributo, por sua vez, foi o que mais interferiu na produtividade (Tabela 1).

As variáveis em estudo possuem associação positiva linear entre si, apresentada na matriz de correlação (Figura 1). Rodrigues et al. (2012) também confirmaram a dependência da produtividade do milho em relação aos atributos do solo.

Os modelos ajustados (regressão e RNA) apresentaram bons desempenhos na estimativa da produtividade de milho em segundo cultivo, com valores ligeiramente baixos no erro padrão da estimativa (Syx), e consideráveis para o coeficiente de determinação (R<sup>2</sup>), conforme explicitados na tabela 2. Entretanto, os parâmetros estatísticos de avaliação da precisão do ajuste da equação via regressão linear múltipla foram inferiores aos resultados advindos da RNA (Tabela 2). Porém, mesmo que todos os estimadores de ajuste sejam bons indicadores para a seleção de modelos, a análise gráfica de resíduo é quesito fundamental na escolha de uma equação, pois erros de tendência podem ocorrer em determinada amplitude de classe da variável resposta, sem ser detectadas pelas estatísticas que medem a exatidão. Na figura 2, verifica-se os valores reais para a produtividade de milho em cada unidade amostral básica nas 196 parcelas usadas no ajuste, bem como o comportamento do modelo de regressão e redes nas suas estimativas. Além da distribuição residual pelas duas categorias de ajuste.

Tanto o AIC quanto o BIC aumentam conforme a soma dos quadrados dos erros aumenta. Dessa forma, quanto menor o valor do AIC e BIC de um modelo, menor sua discrepância geral estimada e, por consequência, menor sua distância relativa para os valores reais.

O modelo de redes neurais com quatro (4) neurônios artificiais, quando comparado com o modelo de regressão, conseguiu estimar de forma mais acurada a produtividade de grãos de milho (Figura 2). Apresentando valores de superestimativas máximas inferiores a 30% e subestimativas mínimas de 25%.

**Tabela 1.** Média, valor máximo e mínimo, desvio padrão e coeficiente de variação para os atributos produtividade de grãos de milho (PG), Matéria Orgânica (MO), Capacidade de Troca Catiônica (CTC), Saturação de bases (V%) e Teor argila (TA) e correlação de *Spearman*, entre variáveis respostas e preditoras

	PG (kg ha <sup>-1</sup> )	MO (mg dm <sup>-3</sup> )	CTC (cmol <sub>c</sub> dm <sup>-3</sup> )	V (%)	TA (g kg <sup>-1</sup> )
Média	5543,24	24,78	6,60	39,04	276,03
Máximo	8038,62	26,80	7,33	44,10	375,00
Mínimo	2056,79	19,66	5,56	25,28	196,97
Desvio Padrão	1200,88	1,68	0,48	5,13	55,86
CV (%)	21,66	6,79	7,22	13,14	20,24
Correlação entre produtividade (variável resposta) e preditoras					
	M.O	CTC	V(%)	TA	
Produtividade	0,39*	0,53*	0,45*	0,44*	

\*Significativo a 5%. MO - matéria orgânica; CTC - Capacidade de troca catiônica; V(%) - Saturação de bases do solo; TA - Teor de argila.

Também houve uma distribuição ligeiramente homogênea ao longo da linha de regressão, demonstrando adaptabilidade da rede na estimativa da variável resposta (produtividade de grãos), bem como dos atributos do solo adotados. Todavia, o modelo advindo do processo de regressão apresentou claras

tendências em superestimar nas áreas menos produtivas e subestimar as áreas mais produtivas, gerando assim erros percentuais variando de -115% a +35%.

Com base nas estatísticas que indicam exatidão, apresentadas na tabela 2, e o comportamento das equações que estimam

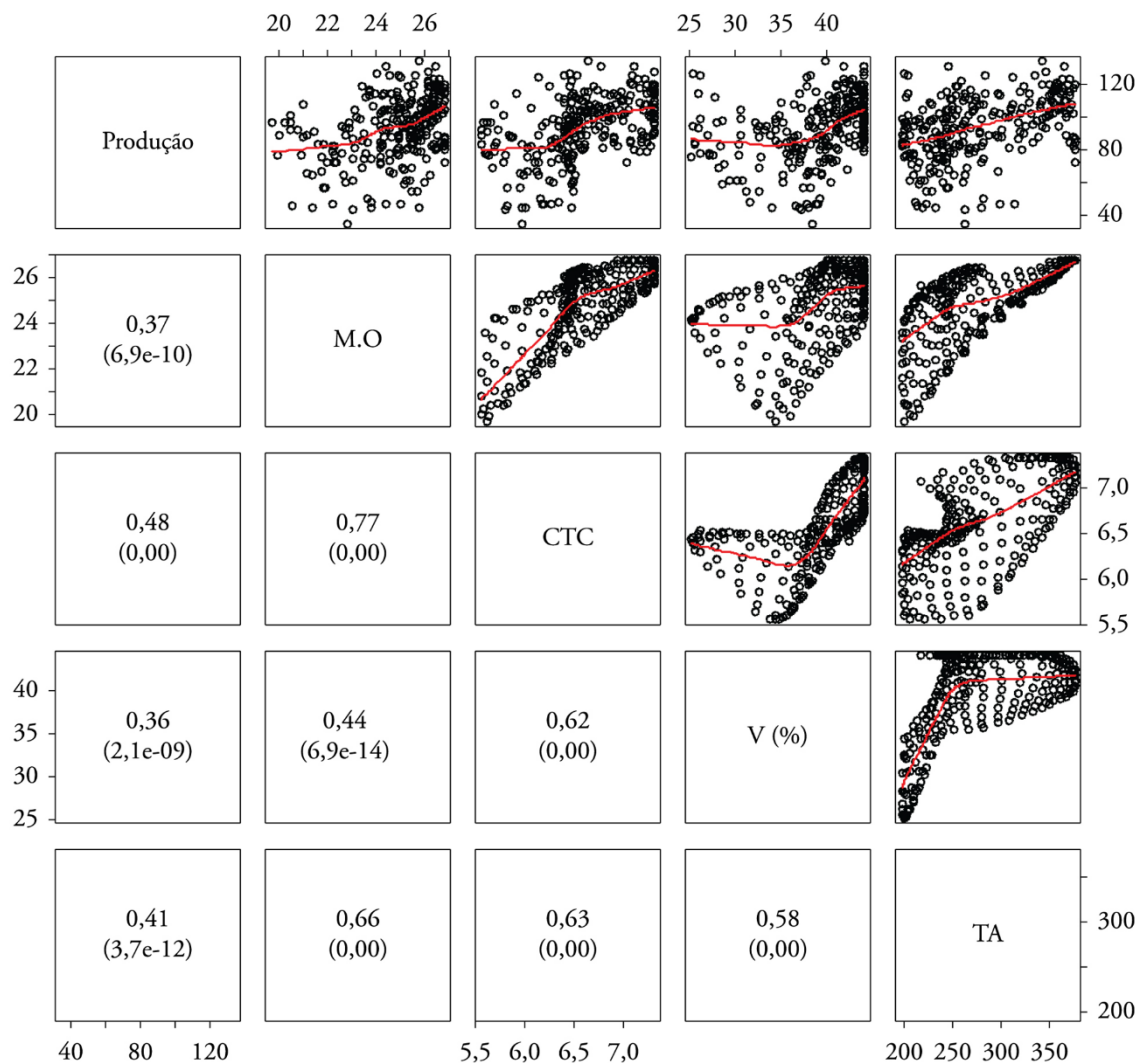


Figura 1. Correlação de Pearson e Matriz de Scatterplots entre os atributos do solo matéria orgânica (MO), Capacidade de troca catiônica (CTC), Saturação de Bases (V%), Teor de argila (TA) e a Produtividade de grãos de milho “segunda safra”.

Tabela 2. Estimativa dos parâmetros para os modelos de regressão linear e variáveis de entrada no procedimento e especificações das redes neurais artificiais (número de neurônios e número de camadas) e parâmetros estatísticos dos modelos de ajuste

Modelo de Regressão (coeficientes ajustados)											
Variável	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	Estatísticas					
						R <sup>2</sup>	Syx (%)	AIC	BIC		
Produt.	-800,23	-100,88	1050,93	16,47	4,61	0,51	17,90	1535	1571		
Redes Neurais											
Rede	Neurônios por camada				Estatísticas						
	Variáveis				Ent.	Ocultas	Saída	R <sup>2</sup>	Syx (%)	AIC	BIC
R. Prod.	MO	CTC	V(%)	TG	4	4	1	0,83	11,58	1130	1163

Em que = R<sup>2</sup>: coeficiente de determinação; Syx (%): erro padrão da estimativa percentual; AIC: critério de Akaike; BIC: critério bayesiano; R. Prod.: rede ajustada para a estimativa da produtividade de grãos de milho.

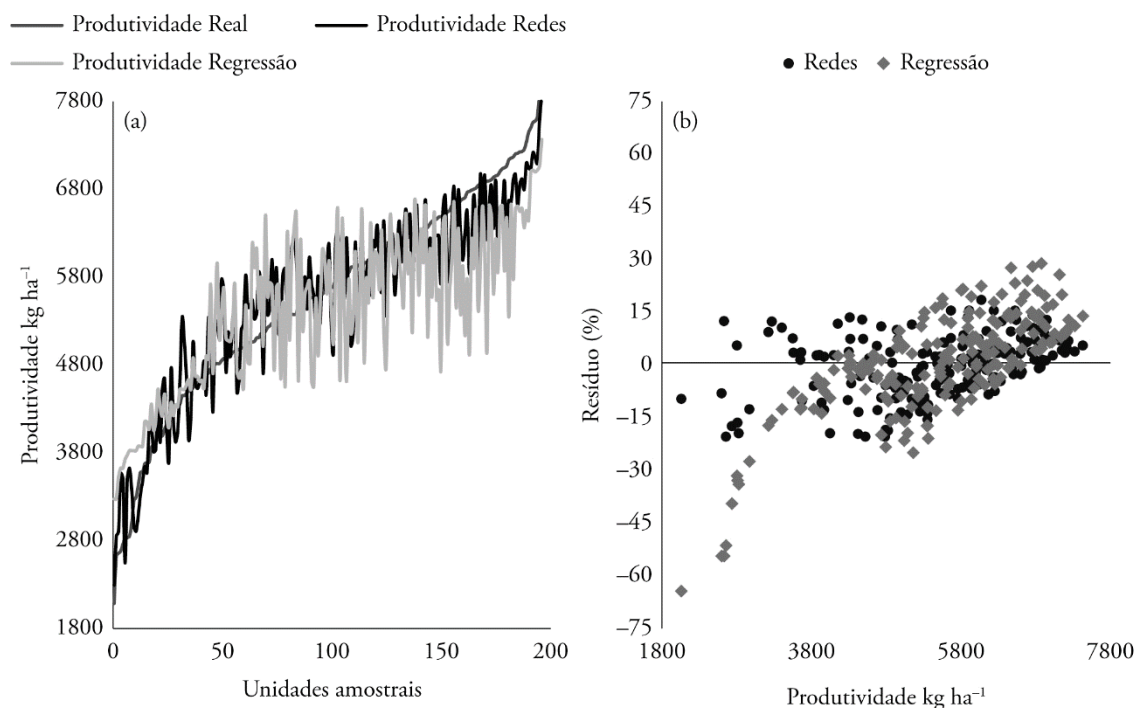
a produtividade, em consonância com os erros residuais (Figura 2), a RNA consegue prever de forma aceitável a produtividade, utilizando os atributos do solo MO, CTC, V(%) e teor de argila. Enquanto, essas mesmas variáveis, quando submetidas às técnicas de regressão perdem suas capacidades preditivas, retornam a estatísticas inferiores à RNA.

A comparação entre a técnica de regressão e a RNA pelo Teste de Qui quadrado ( $\chi^2$ ); o erro padrão da estimativa em forma percentual (Syx%); a diferença agregada (Da); e o erro médio na forma absoluta (Ei) estão apresentados na tabela 3. Os valores calculados de  $\chi^2$ , confrontados com o valor tabelado, foram não significativos a 95% de probabilidade para a RNA. Sendo aceita a hipótese de que os dois modelos são válidos e confiáveis para estimativa de produtividade de grãos de milho cultivado em segundo cultivo, em região de cerrado de altitude por meio dos atributos do solo MO, CTC, V(%) e teor de argila. Entretanto, para uma análise

mais criteriosa, na tabela 3, são apresentados outros critérios estatísticos, referentes ao comportamento dos modelos frente à amostra de validação.

As equações de regressão para as estimativas de produtividade apresentaram leve aumento no erro padrão da estimativa (de 17,9% para 23,25%); as RNA também apresentaram esse comportamento (de 11,58% para 14,93%). Esse ligeiro aumento Syx% é perfeitamente aceitável, haja vista que o número amostral reduziu-se em 75% (de 196 para 65 parcelas).

A diferença agregada (Da) é um parâmetro estatístico usado como índice de ajuste de modelos. Esta corresponde à diferença entre o somatório dos valores observados e o somatório dos valores estimados, servindo como critério indicador de sub ou superestimativas, e, no presente trabalho, para melhor visualização, a mesma foi obtida na forma percentual. Tanto a equação advinda de regressão como redes apresentaram valores positivos para a estimativa da produtividade de grãos



**Figura 2.** Comportamento dos modelos de regressão/redes neurais artificiais na estimativa da produtividade de grãos de milho “segunda safra” a partir de atributos do solo, comparados com os valores reais (a) e gráfico da distribuição residual (b).

**Tabela 3.** Atributos médios, mínimos e máximos, reais e estimados pelas duas categorias de ajuste para produtividade de grãos de milho e estatísticas para a validação

Variáveis	Estatísticas						Teste Qui quadrado	
	Média	Mínimo	Máximo	Syx (%)	Da(%)	Ei	$\chi^2$ cal.	$\chi^2$ tab. (95%; 64 Gl)
Produção	kg ha <sup>-1</sup>							
Real	5656	2626	7895					46,60
Redes	5593	3060	7600	14,93	11,05	123	18,55	
Regressão	5524	3450	7500	23,25	21,75	285	39,45	

Em que: Syx(%), erro padrão da estimativa em porcentagem; Da(%), diferença agregada em porcentagem; Ei(ABS), erro médio;  $\chi^2$  cal, qui quadrado calculado;  $\chi^2$  tab, qui quadrado tabelado; Gl, grau de liberdade da amostra; 95%, probabilidade.

de milho (21,75% e 11,05% respectivamente), caracterizando subestimativas médias nas variáveis estimadas (Tabela 3).

Os resíduos médios gerados pelas duas categorias de ajuste próximos de zero são desejáveis, pois demonstram a capacidade dos modelos em estimar as variáveis de interesse com acurácia. Por regressão, o valor médio do resíduo foi de 285,35 kg ha<sup>-1</sup>. Enquanto na RNA, este foi de 123,31 kg ha<sup>-1</sup>. A análise gráfica mais detalhada dos resíduos foi então empregada ao longo da amplitude da variável produtividade, utilizada na validação, conforme figura 3.

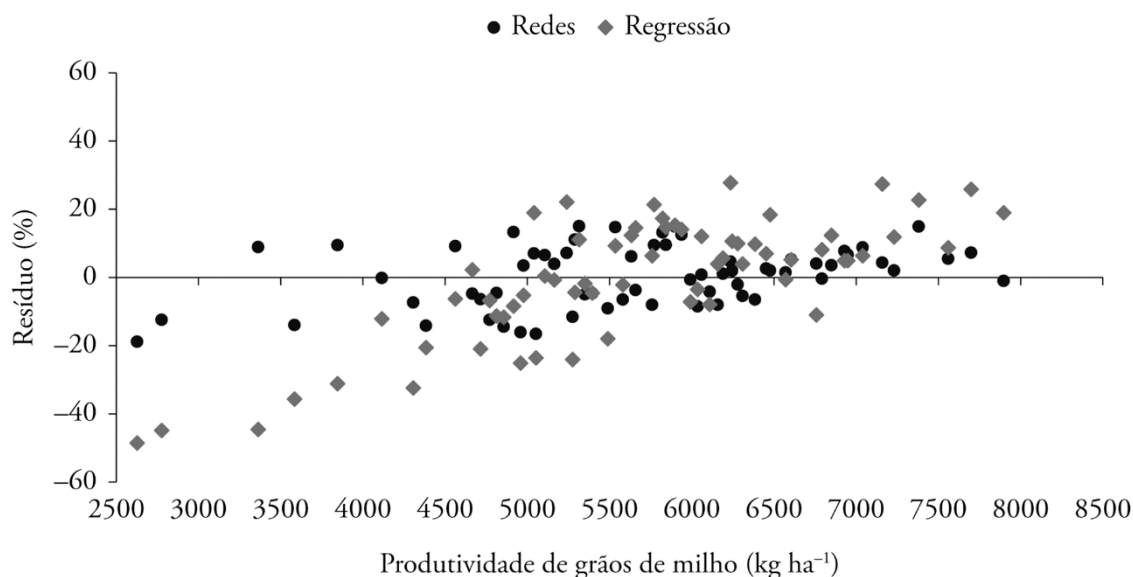
Assim como no ajuste, a equação advinda de redes neurais apresentou-se mais eficaz na estimativa da produtividade. Gerando resíduos melhor distribuídos e de forma mais compacta, não apresentando nenhum ponto crítico de tendência no decorrer de toda a linha de regressão. Por vez, a equação de regressão melhorou sua distribuição residual quando comparada com os resíduos do ajuste, no entanto, resultados estatísticos inferiores a RNA se mantêm.

Levando-se em conta a grande heterogeneidade dos solos, pode-se considerar como satisfatórias e válidas as estimativas geradas pelas duas categorias de ajuste. Todavia, ressalvas são feitas no uso do modelo de regressão para predizerem as áreas de menor produtividade, o qual apresentou uma ligeira tendência em superestimá-las, característica essa não apresentada pela rede neural. A utilização de redes neurais permite a predição da produtividade de milho, “segunda safra” cultivado em cerrado, por meio do uso de atributos do solo, fato comprovado pelo teste de  $\chi^2$ .

Embora, ambas as técnicas de estimativas não se diferiram entre si pelo teste de  $\chi^2$ , vale ressaltar que a RNA, tanto no ajuste como na validação, apresentou melhores estimativas, portanto, corroborando as afirmações de Egrioglu et al. (2014), ao mencionarem que as RNAs

apresentam vantagens sob as técnicas convencionais, devido sua capacidade de generalização, paralelismo e a possibilidade de aprendizagem. Com isso, as mesmas conseguem extrair padrões de determinada base de dados e reapplicá-los em outras com precisão, logo recomenda-se a sua utilização. A utilização da MO, CTC, (V%) e teor de argila foi eficaz na estimativa da produtividade, principalmente quando associadas à RNA. Esse resultado é de grande valia, pois impacta de forma positiva, no planejamento de semeadura, de custo, colheita e venda do produto. A produtividade de grãos pode ser influenciada por fatores como adversidades climáticas, genótipo cultivado, além de ocorrência de pragas e doenças. Entretanto, quando esses fatores não diretamente relacionados com a qualidade do solo são minimizados, a produtividade das culturas pode ser considerada um indicador da qualidade do solo (Lopes et al., 2013). Nas presentes condições, com variação pequena ou inexistente desses demais fatores, os atributos químicos do solo podem representar a qualidade deste e explicar as variações na produtividade.

Ressalta-se que os resultados apresentados possuem validade para a região na qual foi conduzida a pesquisa, considerando a variação das classes de solo, e de suas estruturas físicoquímicas. Nesse sentido, a CTC é o fator que possui maior impacto na predição da produtividade de grãos de milho “segunda safra” e na definição de sítios específicos de manejo. Entretanto, esse atributo é dependente dos teores de matéria orgânica do solo (Gruba & Mulder, 2015), que por sua vez se relacionam, em Latossolos com o teor de argila (Birani et al., 2015). Fato confirmado no presente estudo. Portanto, a CTC é característica de cada tipo de solo e depende, principalmente, dos teores de argila e de matéria orgânica. Capacidade de trocas de cátions superiores possibilitam maior retenção de nutrientes catiônicos



**Figura 3.** Distribuição de resíduos médios para as duas categorias de equações validadas, redes neurais artificiais (RNAs) e regressão múltipla linear.

(K, Ca e Mg), propiciando maior produtividade de grãos de milho, principalmente em segundo cultivo. Além disso, maior retenção de água com incremento de matéria orgânica no solo foi confirmada por Ulyett et al. (2014), ao avaliar o impacto da aplicação de Biochar no solo.

Diante do exposto, como a CTC do solo e os teores de material orgânico são fatores determinantes para a produtividade de grãos de milho em “segunda safra”. Nos sítios específicos que possuem menor potencial de produtividade, deve-se adotar manejo diferenciado para incrementar principalmente a CTC e MO do solo e, conseqüente, a produtividade de grãos de milho. Esse manejo é possível com a aplicação de Biochar (Ulyett et al., 2014; Inal et al., 2015) ou resíduos orgânicos da produção de aves (Santos et al., 2014; Inal et al., 2015), ambos aplicados em taxa variável.

Entretanto, a primeira opção carece de maiores estudos em campo, além de apresentar baixa disponibilidade, representando uma alternativa futura. Enquanto, a segunda opção apresenta-se viável, por possuir alta disponibilidade em regiões com grande produção de aves, como a em estudo. Além de estudos regionais, com comprovados benefícios de sua adoção no cultivo de milho (Guareschi et al., 2013; Silva et al., 2013).

Apesar dos resultados promissores da utilização de sistema de redes neurais para predição de produtividade de milho “segunda safra”, qualificando essa técnica para futuras análises de dados de atributos do solo para definição de unidades de manejo diferenciado do solo, ressalta-se que novos trabalhos devem ser realizados com quantidade maior de variáveis preditoras. Por fim, sugere-se também testar distintas configurações de redes neurais, de forma a conseguir maior correlação entre os dados de produtividade e atributos do solo, obtendo assim maiores acurácias nas estimativas das variáveis de interesse nas lavouras.

#### 4. CONCLUSÃO

A análise dos atributos do solo matéria orgânica, capacidade de troca de cátions, saturação de bases e teor de argila por meio de adoção de redes neurais artificiais, permite a predição da produtividade do milho, cultivado em “segunda safra” em região de cerrado, em consonância ao estabelecimento de sítios específicos de manejo diferenciado do solo.

A técnica de regressão múltipla apresenta inferioridade nas estatísticas que medem precisão do modelo e tende a superestimar as menores classes de produtividade e subestimar as maiores.

#### AGRADECIMENTOS

Os autores agradecem à FUNDECT - Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, pela concessão de financiamento que cobriu parte dos custos da presente pesquisa.

#### REFERÊNCIAS

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723. <http://dx.doi.org/10.1109/TAC.1974.1100705>.
- Binoti, D. H. B., Binoti, M. L. M. S., & Leite, H. G. (2013). BR Patent No 13410-5. Brasil: Instituto Nacional da Propriedade Industrial.
- Birani, S. M., Fernandes, A. R., de Souza Braz, A. M. S., Pedroso, A. J. S., & Alleoni, L. R. F. (2015). Available contents of potentially toxic elements in soils from the Eastern Amazon. *Chemie der Erde - Geochemistry*, 75, 143-151. <http://dx.doi.org/10.1016/j.chemer.2015.01.001>.
- Castro, R. V. O., Soares, C. P. B., Leite, H. G., Souza, A. L., Nogueira, G. S., & Martins, F. B. (2013). Individual growth model for Eucalyptus Stands in Brazil using artificial neural network. *ISRN Forestry*, 2013, 1-12. <http://dx.doi.org/10.1155/2013/196832>.
- Chen, C., Ren, T., Hu, K., Li, B., & Wang, Y. (2014a). Estimation of soil clay content using Hygroscopic water content at an arbitrary humidity. *Soil Science Society of America Journal*, 78, 119-124. <http://dx.doi.org/10.2136/sssaj2013.06.0247>.
- Chen, W. C., Tseng, L. Y., & Wu, C. S. (2014b). A unified evolutionary training scheme for single and ensemble of feedforward neural network. *Neurocomputing*, 143, 347-361. <http://dx.doi.org/10.1016/j.neucom.2014.05.057>.
- Companhia Nacional de Abastecimento – CONAB (2014). Acompanhamento da Safra Brasileira: Grãos, 1, 1-127. Recuperado em 3 de abril de 2013, de [www.conab.gov.br/OlalaCMS/uploads/arquivos/14\\_09\\_10\\_14\\_35\\_09\\_boletim\\_graos\\_setembro\\_2014.pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/14_09_10_14_35_09_boletim_graos_setembro_2014.pdf)
- Dalchiavon, F. C., Carvalho, M. P., Freddi, O. S., Andreotti, M., & Montanari, R. (2011). Variabilidade espacial da produtividade do feijoeiro correlacionada com atributos químicos de um Latossolo Vermelho Distroférrico sob sistema de semeadura direta. *Bragantia*, 70, 908-916. <http://dx.doi.org/10.1590/S0006-87052011000400025>.
- Egrioglu, E., Yolcu, U., Aladag, C. A., & Bas, E. (2014). Recurrent multiplicative neuron model artificial neural network for non-linear time series forecasting. *Procedia: Social and Behavioral Sciences*, 109, 1094-1100. <http://dx.doi.org/10.1016/j.sbspro.2013.12.593>.
- Freddi, O. S., Stoquero, A., & Peres, F. S. C. (2014). Saturação do solo por bases na produtividade e qualidade de fibra do algodoeiro. *Nativa. Sinop*, 2, 6-12. <http://dx.doi.org/10.14583/2318-7670.v02n01a02>.
- Gomes, F. P. (2000). Curso de estatística experimental (14 ed.). Piracicaba: Degaspari. 477 p.
- Gruba, P., & Mulder, J. (2015). Tree species affect cation exchange capacity (CEC) and cation binding properties of organic matter in acid forest soils. *The Science of the Total Environment*, 511, 655-662. <http://dx.doi.org/10.1016/j.scitotenv.2015.01.013>. PMID:25596350.
- Guareschi, R. F., Silva, A., Silva, H. R., Jr., Perin, A., & Gazolla, P. R. (2013). Adubação orgânica na produção de biomassa de milho em Latossolo de cerrado. *Global Science and Technology*, 6, 66-733. <http://dx.doi.org/10.14688/1984-3801.v06n02a08>.
- Haykin, S. (2001). Redes neurais: princípios e prática. Porto Alegre: Bookman. 900 p.



- Inal, A., Gunes, A., Sahin, O., Taskin, M. B., & Kaya, E. C. (2015). Impacts of biochar and processed poultry manure, applied to a calcareous soil, on the growth of bean and maize. *Soil Use and Management*, 31, 106-113. <http://dx.doi.org/10.1111/sum.12162>.
- Leal, A. J. F., Lazarini, E., Rodrigues, L. R., & Marcandalli, L. H. (2013). Adubação nitrogenada para milho com o uso de plantas de cobertura e modos de aplicação de calcário. *Revista Brasileira de Ciencia do Solo*, 37, 491-501. <http://dx.doi.org/10.1590/S0100-06832013000200020>.
- Leite, H. G., Silva, M. L. M., Binoti, D. H. B., Fardin, L., & Takizawa, F. H. (2011). Estimation of inside-bark diameter and heartwood diameter for *Tectona grandis* Linn. trees using artificial neural networks. *European Journal of Forest Research*, 130, 263-269. <http://dx.doi.org/10.1007/s10342-010-0427-7>.
- Lopes, A. A. C., Sousa, D. M. G., Chaer, G. M., Reis, F. B., Jr., Goedert, W., & Mendes, I. C. (2013). Interpretation of microbial soil indicators as a function of crop yield and organic carbon. *Soil Science Society of America Journal*, 77, 461-472. <http://dx.doi.org/10.2136/sssaj2012.0191>.
- Montanari, R., Carvalho, M.P., Silva, C.A., Jr., Corrêa, A.R., Dalchiavon, F.C., & González, A.P. (2013). Relations between the yield of bean (*Phaseolus vulgaris* L.) and chemical attributes of an Acrustox under no-tillage. *Journal of soil science and plant nutrition*, 13, 367-379. <http://dx.doi.org/10.4067/S0718-95162013005000030>.
- Nicolodi, M., Anghinoni, I., & Gianello, C. (2008). Indicadores da acidez do solo para recomendação de calagem no sistema plantio direto. *Revista Brasileira de Ciencia do Solo*, 32, 237-247. <http://dx.doi.org/10.1590/S0100-06832008000100023>.
- Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics*. New Jersey: John Wiley & Sons. 663 p.
- Rodrigues, M. S., Corá, J. E., & Fernandes, C. (2012). Spatial relationships between soil attributes and corn yield in no-tillage system. *Revista Brasileira de Ciencia do Solo*, 36, 599-609. <http://dx.doi.org/10.1590/S0100-06832012000200029>.
- Rodrigues, M. S., Corá, J. E., Castrignanò, A., Mueller, T. G., & Rienzi, E. (2013). A spatial and temporal prediction model of corn grain yield as a function of soil attributes. *Agronomy Journal*, 105, 1878-1887. <http://dx.doi.org/10.2134/agronj2012.0456>.
- Santos, H. G., Jacomine, P. K. T., Anjos, L. H. C., Oliveira, V. A., Lumbreiras, J. F., Coelho, M. R., Almeida, J. A., Cunha, T. J. F., & Oliveira, J. B. (2013). *Sistema brasileiro de classificação de solos* (3 ed.). Brasília, DF: Embrapa. 353 p.
- Santos, L. B., Castagnara, D. D., Bulegon, L. G., Zoz, T., Oliveira, P. S. R., Gonçalves, A. C., Jr., & Neres, M. A. (2014). Substituição da adubação nitrogenada mineral pela cama de frango na sucessão aveia/milho. *Bioscience Journal*, 30, 272-281. Recuperado em 4 de abril de 2015, de [www.seer.ufu.br/index.php/biosciencejournal/article/view/17979/14538](http://www.seer.ufu.br/index.php/biosciencejournal/article/view/17979/14538)
- Schwarz, G. (1978). Estimating the dimensional of a model. *Annals Statistics*, Hayward, 6, 461-464. <http://dx.doi.org/10.1214/aos/1176344136>.
- Silva, A. A. V., Silva, I. A. F., Teixeira, M. C. M., Fo., Buzetti, S., & Teixeira, M. C. M. (2014). Estimativa da produtividade de trigo em função da adubação nitrogenada utilizando modelagem neuro fuzzy. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 18, 180-187. <http://dx.doi.org/10.1590/S1415-43662014000200008>.
- Silva, F. C. (2009). *Manual de análises químicas de solos, plantas e fertilizantes* (3 ed.). Brasília, DF: Embrapa. 627 p.
- Silva, T. R., Menezes, J. F. S., Simon, G. A., & Assis, R. L. (2013). Desenvolvimento inicial do milho em um Latossolo Vermelho distrófico com aplicação de cama de poedeira. *Global Science and Technology*, 6, 1-7. <http://dx.doi.org/10.14688/1984-3801.v06n03a01>.
- Soares, F. C., Robaina, A. D., Peiter, M. X., Russi, J. L., & Vivan, G. A. (2014). Redes neurais artificiais na estimativa da retenção de água do solo. *Ciência Rural*, 44, 293-300. <http://dx.doi.org/10.1590/S0103-84782014000200016>.
- Sousa, D. M. G., & Lobato, E. (2004). Adubação com nitrogênio. In D.M.G. Sousa, & E. Lobato (Eds.), *Cerrado: correção do solo e adubação* (2 ed., p. 129-145). Brasília: Embrapa Informação Tecnológica.
- Srinivasan, A. (2006). *Handbook of precision agriculture: principles and applications*. New York: CRC. 683 p.
- Ulyett, J., Sakrabani, R., Kibblewhite, M., & Hann, M. (2014). Impact of biochar addition on water retention, nitrification and carbon dioxide evolution from two sandy loam soils. *European Journal of Soil Science*, 65, 96-104. <http://dx.doi.org/10.1111/ejss.12081>.