

REDES NEURAS ARTIFICIAIS PARA ESTIMAÇÃO DA PRODUTIVIDADE DA SOJA

GABRIELA KAROLINE MICHELON¹, PAULO LOPES DE MENEZES², CLAUDIO LEONES BAZZI², JEAN CARLOS GRANDO³, FERNANDO ROBERTO ROCKENBACH⁴.

¹ Graduada em Ciência da Computação, Discente do Programa de Mestrado em Tecnologias Computacionais para o Agronegócio, UTFPR, Medianeira – PR, Fone: (0XX45) 3240-8000, gabrielamichelon@alunos.utfpr.edu.br.

² Doutor, Docente do Programa de Mestrado em Tecnologia Computacionais para o Agronegócio, UTFPR, Medianeira – PR.

³ Discente do curso de Engenharia Agrônômica, UDC, Foz do Iguaçu - PR.

⁴ Graduado em Engenharia Civil, Discente do Programa de Mestrado em Tecnologias Computacionais para o Agronegócio, UTFPR, Medianeira – PR.

Apresentado no
XLVI Congresso Brasileiro de Engenharia Agrícola - CONBEA 2017
30 de julho a 03 de agosto de 2017 - Maceió - AL, Brasil

RESUMO: O gerenciamento de propriedades agrícolas por meio da agricultura de precisão tem se tornado fundamental para aumentar o potencial produtivo de áreas agrícolas. Além disso, a tecnologia da informação, por meio das ferramentas computacionais, auxilia e melhora a prática agrícola, tanto no manejo do solo como da planta. Este trabalho objetivou a aplicação de Redes Neurais Artificiais para obter um modelo computacional de predição da produtividade da soja. Com a predição da produtividade, se torna possível a correção localizada nos pontos amostrais preditos com baixas produtividades, viabilizando melhores condições à planta com aumento do seu potencial produtivo. Em contrapartida, em locais onde há produtividade considerada alta, pode-se administrar a aplicação, reduzindo-se assim os custos com insumos e possíveis impactos ambientais causados pelo excesso de fertilizantes. Para a criação do modelo de predição, utilizou-se dados dos macronutrientes da folha da soja (Nitrogênio, Potássio, Fósforo, Cálcio e Magnésio), buscando-se obter as condições de fertilidade do solo. Como resultado, obteve-se um modelo de rede neural artificial capaz de prever, aproximadamente, 74% dos dados reais com apenas um neurônio em uma única camada oculta, utilizando para isto, os nutrientes da folha da soja no estágio V3.

PALAVRAS-CHAVE: Inteligência Artificial, macronutrientes, agricultura de precisão.

ARTIFICIAL NEURAL NETWORKS TO ESTIMATION OF SOYBEAN PRODUCTIVITY

ABSTRACT: The management of agricultural properties through precision agriculture has become fundamental to increase the potential of agricultural areas. Besides that, information technology, through computational tools, helps and improves agricultural practice, both in soil and plant management. This work aimed the application of Artificial Neural Networks to obtain a computational model of prediction of soybean yield. With the prediction of productivity, it becomes possible the localized correction at the predicted sample points with low productivity, allowing better conditions to the plant enabling an increase in its productive potential. On the other hand, in places where there is high productivity, the application can be administered, reducing costs with inputs and possible environmental impacts caused by excess of fertilizers. For the creation of the prediction model, was used data from the macronutrients of the soybean leaf (Nitrogen, Potassium, Phosphorus, Calcium and

Magnesium), seeking to obtain soil fertility conditions. As result, it was obtained an artificial neural network model, capable of predicting, approximately, 74% of the actual data, with only one neuron in a single hidden layer, using for this the nutrients of soybean leaf in the V3 stadium.

KEYWORDS: Artificial Intelligence, macronutrients, precision agriculture.

INTRODUÇÃO

Estratégias de gerenciamento de propriedades agrícolas que utilizam a tecnologia da informação para aumentar a produção e minimizar impactos ambientais são conhecidas mundialmente como técnicas de Agricultura de Precisão (AP). As técnicas de AP estão se tornando cada vez mais importantes devido ao aumento da população e conseqüentemente da demanda por alimentos, o que faz com que seja requerido o aumento da eficiência produtiva de áreas produtoras para atender a esta demanda (MONDO et al., 2012).

Os avanços da AP têm contribuído para obter ferramentas de predição da produtividade, as quais melhoram a performance de produção de uma área produtiva. Entre os muitos fatores que podem influenciar na produtividade e servir como indicadores na predição da produtividade, estão os nutrientes do solo que são absorvidos pela planta e influenciam diretamente no seu desenvolvimento (VITTI; TREVISAN, 2000).

Entre as principais culturas produzidas no Brasil, a soja (*Glycine max*) é uma das oleaginosas mais importantes da produção agrícola do mundo, devido a sua alta qualidade nutricional, referente a alta concentração de proteínas e aminoácidos e por possuir um maior rendimento em relação a outras culturas (GANDHI, 2009). Para obter uma boa produtividade, a soja necessita de solos com disponibilidade de nutrientes como Fósforo (P), Potássio (K), Cálcio (Ca), Magnésio (Mg), Enxofre (S), Boro (B), Cloro (C), Cobre (Cu), Ferro (Fe), Manganês (Mn), Molibdênio (Mo), Cobalto (Co), Zinco (Zn) e Nitrogênio (N), porém os mais absorvidos por esta cultura são os chamados macronutrientes: N, P, K, Ca e Mg, os quais exercem maior influência no seu desenvolvimento (VITTI; TREVISAN, 2000).

Este trabalho buscou realizar a predição da produtividade por meio dos nutrientes presentes na folha da soja (absorvidos do solo), sendo eles os macronutrientes N, P, K, Ca e Mg, em três estádios do desenvolvimento da planta. Para tanto, utilizou-se a técnica computacional de Redes Neurais Artificiais (RNAs) para obter um modelo de predição da produtividade. A escolha pelas RNAs é devido a sua capacidade de resolver problemas não lineares (KHAIRUNNIZA-BEJO et al., 2014).

A Rede Neural Artificial é baseada no funcionamento do cérebro humano, pois possui neurônios artificiais interconectados por meio de sinapses, os quais são capazes de adquirir conhecimento por meio de um processo de aprendizagem (HAYKIN, 2001).

Portanto, este trabalho objetivou obter um modelo de predição da produtividade da soja utilizando a técnica computacional de Redes Neurais Artificiais *Multilayer Perceptron Feedforward* com o algoritmo de treinamento *backpropagation* utilizando como entrada para a rede os nutrientes: N, P, K, Ca e Mg e como saída a produtividade da soja.

Um modelo capaz de prever a produtividade da soja somente com os nutrientes da folha, facilitará a aplicabilidade da agricultura de precisão, pois sabendo-se os locais que possuem menor probabilidade de ter alta produtividade, poderão ser realizadas correções dos nutrientes do solo de forma localizada antes da colheita, quando em estágios iniciais do desenvolvimento da planta, ou futuras adubações, caso a cultura esteja muito avançada, nos locais necessários para aumentar a produtividade na próxima safra.

MATERIAL E MÉTODOS

Os nutrientes analisados neste trabalho foram coletados em 3 estádios do desenvolvimento da soja em duas áreas experimentais (Talhão 1, com 10 hectares (ha) e Talhão 2, com 23,8 ha), com solos classificados como LATOSSOLO VERMELHO Distroférico típico (EMBRAPA, 2006), ambas cultivadas sob sistema de plantio direto por pelo menos 10 anos com sucessão de cultura soja/milho para fins comerciais.

As áreas estão localizadas no município de Serranópolis do Iguaçu/PR, sob coordenadas geográficas centrais 25°24'28" S e 54°00'17" O com elevação média de 355 m. Foram determinados por meio do software Pathfinder, duas grades amostrais irregulares, contendo 42 pontos amostrais no talhão 1 e 73 no talhão 2, com densidade de aproximadamente 3 pontos por ha.

Para realizar a coleta das folhas, realizou-se de forma manual no ano de 2012, em 3 estádios da safra da soja 2012/2013, V3, V6 e R1, onde o estádio V3 é o estádio vegetativo da planta com 3 folhas, o V6 é o estádio da planta com 6 folhas e o estádio reprodutivo da planta já com flores, chamado de R1 (WOLLENHAUPT et al., 1997). Os pontos foram localizados por meio de um receptor GPS (GPS Trimble Pathfinder Pro XT). Para cada ponto amostral em um raio de 3 metros, coletou-se 12 folhas da soja, localizada no meio da planta (entre o dossel e o solo) (WOLLENHAUPT et al., 1997).

As folhas foram acondicionadas em sacolas de papel e posteriormente secas em estufa a 105°. Após a secagem, as folhas foram moídas por um moinho de facas modelo SL 31 da marca SOLAB. Ao final de todo este processo as folhas foram encaminhadas para análise laboratorial dos macronutrientes Ca, Mg, P, K, N, no Laboratório de Solos da Universidade Tecnológica Federal do Paraná (UTFPR) - Câmpus Pato Branco em convênio com o Instituto Agrônomo do Paraná (IAPAR), os quais seguem a metodologia de Tedesco et al. (1995) para a determinação de macronutrientes em tecido vegetal.

A produtividade da soja 2012/2013 foi obtida por meio de coleta manual, no final do mês de janeiro, coletando-se para cada ponto amostral, duas linhas de plantio, por um percurso de 1 m, representando uma porção de cerca de 0,9 m². Após a coleta, foi realizada a debulha manual e o procedimento de secagem em estufa a 105° para posteriormente ser calculada a produtividade em toneladas por hectare (ton ha⁻¹), considerando uma umidade de grãos de 12%.

O software utilizado para implementar a RNA foi o Stuttgart Neural Network Simulator (SNNS). Os dados foram formatados (após sua normalização) de acordo com o Software utilizado. O SNNS permite simular a eficiência e flexibilidade da rede neural criada para posteriormente aplicar aos dados novos a rede treinada que obteve o melhor desempenho.

Outro software utilizado foi o Java Neural Network Simulator (JavaNNS), apenas para visualizar a rede treinada no SNNS, pois ele não possui formas de implementar algoritmos que possam automatizar a criação, testes e validação de RNAs.

A função de ativação utilizada para as RNAs foi a sigmoide logística e a função de aprendizagem adotado foi o *backpropagation* padrão, o qual possui duas fases, a fase para frente (*forward*) e para trás (*backward*). Na fase *forward*, cada neurônio da primeira camada intermediária, recebe os valores já ponderados pelas conexões de entrada e, a partir deste, produz os valores de saída que serão utilizados como entrada para a sua camada seguinte, assim sucessivamente até que os neurônios da camada de saída produzam cada um o seu valor de saída. Ao fim deste processo, é analisado o erro, que é a diferença entre as saídas produzidas e os respectivos valores desejados (FACELI et al., 2011).

O erro obtido na fase *forward* é utilizado na fase *backward*, na qual pelo método do gradiente descendente é ajustado os pesos da rede, desde a camada de saída até a primeira camada intermediária. Desta forma, o algoritmo *Backpropagation* faz com que um erro de uma

determinada camada intermediária seja estimado com a soma dos erros dos neurônios da camada seguinte (MITCHELL, 1997; SILVA et al., 2010).

A arquitetura das redes neurais criadas foi a *multilayer Perceptron* (MLP), as quais possuem uma ou mais camadas intermediárias, sendo alimentadas (direção em que ocorre o fluxo dos dados da rede) sempre em uma direção, da camada de entrada para a camada de saída (SILVA, et al. 2010). Segundo Braga *et al.* (2007), MLPs constituem aproximadores globais e são recomendadas para resolução de problemas mais complexos.

Para aplicar a técnica de Inteligência Artificial foi necessário realizar previamente a normalização dos dados e a modelagem de acordo com o formato requerido por cada ferramenta a ser utilizada. Portanto, normalizou-se os dados entre 0 e 1, de acordo com a Equação 1 (SU et al., 1997) antes de aplicá-los às ferramentas computacionais.

$$f(x) = y = (y_{\max} - y_{\min}) * \frac{x - x_{\min}}{x_{\max} - x_{\min}} + y_{\min} \quad (1)$$

Em que, y é o valor normalizado correspondente ao valor original x , e x_{\min} e x_{\max} são os valores de mínimo e de máximo entre todos os valores da base de dados e y_{\min} corresponde a 0 e y_{\max} corresponde a 1.

Os dados utilizados como entrada no programa computacional de treinamento da RNA, foram dos nutrientes N, P, K, Ca e Mg e o valor de saída foi o da produtividade da soja. O conjunto de treinamento representou uma porção de 60% do conjunto total, enquanto o conjunto de validação representou 30% dos dados e o conjunto de teste, os outros 10% restantes como sugerido por Silva *et al.* (2010).

O conjunto de validação foi utilizado na RNA logo após o treinamento, para verificar se o modelo aprendeu a obter a saída (produtividade da soja) por meio das entradas. O conjunto de teste, só foi utilizado após escolhida a rede com menor erro para avaliar se as saídas por ela estimadas são significativamente semelhantes ou não aos dados observados.

A técnica utilizada para determinar o momento de parar o treinamento das RNA foi a do treinamento com parada antecipada (HAYKIN, 2001), a qual permite comparar a cada ciclo de treinamento o valor estimado pela RNA com o valor real da produtividade da soja. A partir disto, obtém-se um erro, e caso o erro for abaixo do aceitável, a RNA treinada foi satisfatória e encerram-se os treinamentos, caso contrário, deve ser continuado o treinamento até que se obtenha o erro esperado, ou até que o erro de validação comece a aumentar. O erro analisado nas RNAs foi o *Mean squared error* (MSE).

Para avaliar o modelo baseado nas técnicas de inteligência artificial resultante deste trabalho utilizou-se: teste de normalidade dos dados para confirmar se os testes estatísticos paramétricos eram aplicáveis neste estudo; teste de hipóteses H_0 e H_1 para testar se os dados reais e os simulados poderiam ser considerados provenientes de uma mesma população (o que é esperado); coeficiente de determinação R^2 para testar o quanto o modelo simulado se ajustou aos dados reais e o coeficiente de correlação de Pearson R_p para quantificar a correlação dos dados reais e simulados.

RESULTADOS E DISCUSSÃO

Realizou-se os testes de normalidade dos dados dos nutrientes e da produtividade pelos testes de Anderson - Darling, Kolmogorov - Smirnov, Shapiro - Wilk, Ryan - Joiner. Para os atributos, nitrogênio, potássio e cálcio, não foi possível verificar normalidade ao nível de 5% ($\alpha = 0,05$) de significância. Portanto, transformou-se os dados pela transformação logarítmica, raiz quadrada dos dados e a divisão de $1/x$ chamada de transformação do quociente. Porém, como resultado, os testes de normalidade apresentaram p-valor mais distantes do nível de 5% de significância do que p-valores obtidos antes da transformação dos dados. Portanto,

rejeitou-se a hipótese de que os dados dos nutrientes seguem uma distribuição normal. A não normalidade dos dados pode ser atribuída a deficiência ou excesso de nutrientes ou de elementos químicos nocivos à planta. Deste modo, considerando que a normalidade dos dados não pode ser assumida, os testes de avaliação das redes neurais a serem apresentados foram baseados em métodos estatísticos não paramétricos.

Criou-se Redes Neurais Artificiais do tipo Perceptron Multicamadas utilizando o algoritmo backpropagation para cada uma das três coletas de folhas realizadas (estádios V3, V6 e R1), sendo utilizadas topologias de redes com até três camadas intermediárias, variando a quantidade de 1 a 500 neurônios em cada camada, totalizando aproximadamente 12000 redes neurais artificiais treinadas. A primeira camada da rede neural, recebeu como entradas os nutrientes N, P, K, Ca e Mg em porcentagem e a última camada, a camada de saída, o valor da produtividade da soja em toneladas por hectare. Como trata-se de unidades métricas distintas, os dados foram padronizados em um intervalo entre 0 e 1 antes de serem utilizados na RNA.

Os resultados apresentados na Tabela 1, são de redes que tiveram o menor MSE em cada coleta, sendo ainda apresentada suas respectivas topologias utilizadas, a quantidade de épocas até o momento de parada do treinamento da RNA (quando o erro de validação começa a subir), os valores de MSE, os coeficientes de determinação (R^2) e os coeficientes de correlação de Pearson (Rp).

Apesar de terem sido realizados testes com RNA de até três camadas intermediárias, os melhores desempenhos foram obtidos com uma camada intermediária e com uma pequena quantidade de neurônios, variando de somente um neurônio na RNA da primeira coleta, 6 neurônios na camada intermediária da RNA da segunda coleta e 5 neurônios na camada intermediária da RNA da terceira coleta. Como mostrado por Cybenko (1989), a partir de extensões do Teorema de Kolmogorov, são necessárias no máximo duas camadas intermediárias, com um número suficiente de unidades (neurônios) por camada, para se produzir quaisquer mapeamentos. Também foi mostrado que apenas uma camada intermediária é suficiente para aproximar qualquer função contínua, deste modo, os testes realizados foram considerados suficientes para verificar a capacidade de aprendizado das redes neurais criadas.

TABELA 1. Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (pontos amostrais aleatórios para conjunto de treinamento, validação e teste em cada coleta)

Coleta	Topologia	Épocas	MSE	R^2	R
1	1x5_1x1_1x1	67	0,0387	0,7407	-0,8606
2	1x5_1x6_1x1	15	0,0242	0,0085	0,0922
3	1x5_1x1_1x1	27	0,0153	5E-05	0,0071

O número que antecede a X representa o número de camadas e o valor posterior a X o número de neurônios em cada camada, onde o primeiro trio de caracteres descreve a camada de entrada, o segundo trio de caracteres a camada intermediária e o último a camada de saída da rede neural artificial.

O coeficiente de determinação é utilizado para avaliar a qualidade do ajuste do modelo, pois indica quanto o modelo foi capaz de explicar os dados coletados, sendo que, quanto mais próximo de 1 este coeficiente, o modelo pode ser considerado melhor. A primeira coleta dos nutrientes da folha foi a única que teve um bom desempenho, pois a melhor RNA obtida com os dados da primeira coleta teve um modelo que conseguiu explicar 74% da variação dos dados reais. Já nas demais coletas tiveram modelos muito ruins, que não conseguiram explicar nem 1% da variação dos dados originais.

O coeficiente de correlação linear de Pearson foi utilizado para medir a correlação linear entre os dados reais e os simulados pela RNA. De acordo com a classificação de Callegari e Sídia (2003), a RNA obtida com os dados da primeira coleta tem uma forte correlação linear, apesar de ter sido negativa, o que significa que o modelo de RNA é muito bom, pois apresentou correlação de -0,86. Para a segunda e terceira coleta a correlação linear é classificada como fraca.

Estes resultados podem ter sido influenciados pela menor quantidade de nutrientes restante na folha no estágio V6 e R1, pois segundo Mundstock e Thomas (2005) e Ritchie (1998), no estágio V6 e nos estádios reprodutivos, se inicia um período de rápido e constante acúmulo diário das taxas de matéria seca e de nutrientes pela planta, que se inicia nas partes vegetativas (folhas, hastes, pecíolos e raízes), deslocando-se gradualmente para as vagens e sementes em formação, enquanto as partes vegetativas finalizam o seu desenvolvimento.

Segundo Rosolem (2002) quanto mais velha menor a absorção de nutrientes da folha e os nutrientes são transmitidos para as flores e vagens da planta.

Nas Figuras 1, 2 e 3 pode ser visualizado os gráficos de dispersão para os dados reais (eixo x) e os dados simulados pela melhor RNA (eixo y) separadas por coleta.

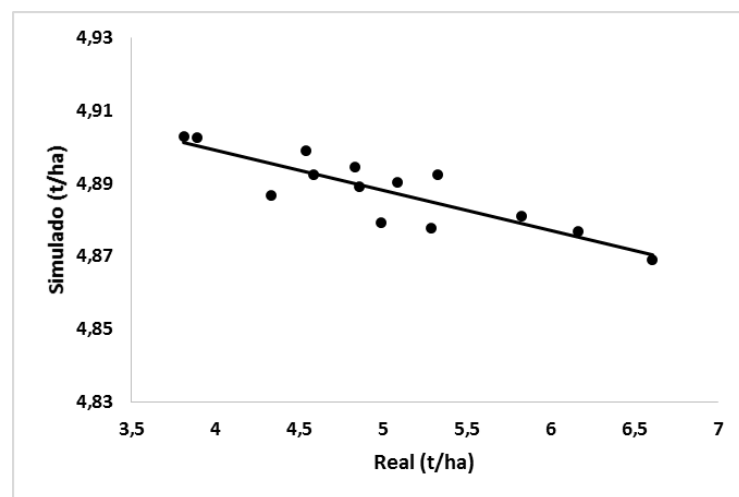


FIGURA 1. Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da primeira coleta de nutrientes da folha.

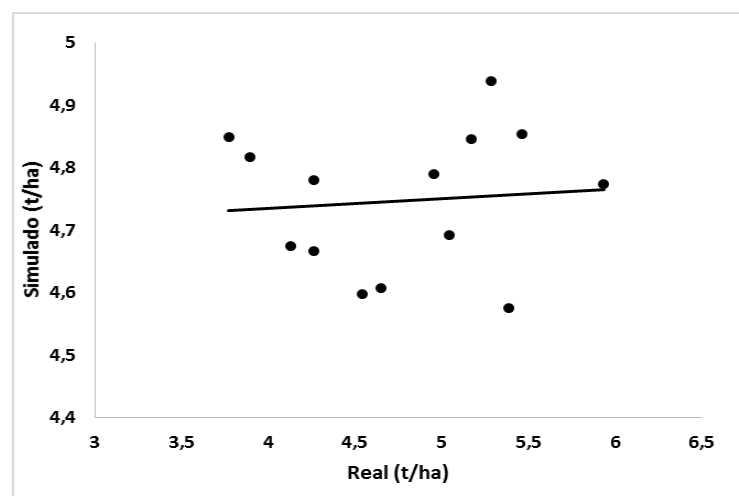


FIGURA 2. Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da segunda coleta de nutrientes da folha.

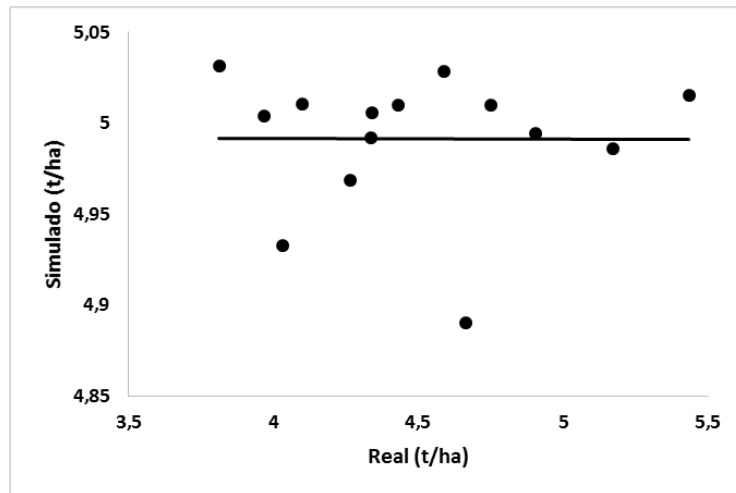


FIGURA 3. Resultado dos valores estimados da produtividade da soja com a RNA obtida com os dados da terceira coleta de nutrientes da folha.

Na Figura 4, está ilustrado como ficou a topologia do melhor modelo de rede neural artificial obtida para coleta 1 dos nutrientes da soja e os pesos sinápticos de cada conexão entre os neurônios.

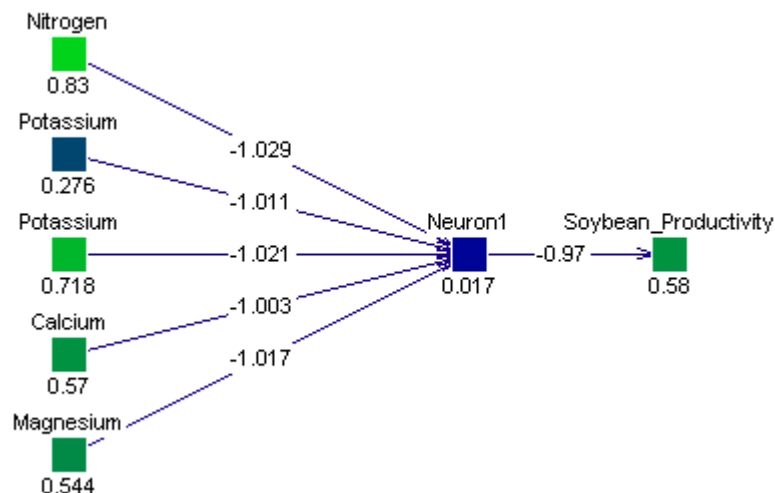


FIGURA 4. Ilustração do melhor modelo de RNA com os sinais de entrada, os pesos sinápticos, o limiar de ativação, potencial de ativação e função de ativação, utilizando os dados da primeira coleta da folha da soja.

Foram realizados modelos de RNAs para as áreas A e B separadas, para verificar se o desempenho dos modelos estava sendo prejudicado pelas diferenças entre as áreas, desde diferenças nutricionais até as diferenças de relevo e plantio. A Tabela 2 possui os resultados obtidos pelas melhores RNAs de cada coleta dos nutrientes da soja somente para a área A.

Para a área B, os resultados não foram tão bons quanto na área A. As topologias obtiveram maiores quantidades de neurônios nas camadas intermediárias e os coeficientes de determinação não explicaram mais que 36% dos dados reais.

Para fazer um comparativo dos dados reais e os simulados pelas RNAs e verificar o quanto as amostras são homogêneas entre os dois conjuntos de dados, realizou-se o teste não paramétrico de Kruskal-Wallis (PONTES E CORRENTES, 2001), devido aos dados brutos não apresentarem normalidade (Tabela 3).

TABELA 2. Topologia das melhores redes para cada uma das coletas dos nutrientes da folha da soja para a área A e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta)

Coleta	Topologia	Épocas	MSE	R ²	R _p
1	1x5_1x3_1x1	18806	0,0186	0,3195	-0,565
2	1x5_1x1_1x1	59	0,0200	0,6147	0,7840
3	1x5_1x1_1x1	57	0,0199	0,5220	0,7225

TABELA 3. Topologia das melhores redes para cada uma das coletas dos nutrientes da folha na área A e seus respectivos erros quadráticos médios, quantidade de épocas, coeficiente de determinação e coeficiente de correlação de Pearson (mesmos pontos amostrais para conjunto de treinamento, validação e teste em cada coleta)

Kruskal-Wallis		
	H	P-valor
C1	0	1
C2	0,019	0,8903
C3	10,0513	0,0015

Na Tabela 3, observa-se que as RNAs da coleta 1 e 2 dos nutrientes da soja obteve p-valor $> 0,05$, o que significa que o teste de Kruskal-Wallis aceita a hipótese de homogeneidade entre os dados reais e simulados ao nível de 5% de significância. Já o modelo de RNA da terceira coleta dos nutrientes da folha da soja, rejeita a hipótese de que os dados reais e simulados vêm de populações semelhantes.

O valor crítico é 3,565 (valor analisado na tabela do qui-quadrado utilizada para testes de Kruskal-Wallis ao nível de 5% de confiança), o que implica que se o valor de H for menor que o valor crítico, os dados reais e simulados não possuem diferenças significativas, como ocorre na coleta 1 e 2, porém na coleta 3, além de os dados não serem semelhantes, o valor crítico comprova que há diferenças significativas entre os grupos analisados (grupo dos dados reais e outro dos dados simulados), pois o valor de H é maior que o valor crítico.

Na Figura 5 visualiza-se a comparação dos erros quadráticos médios de treinamento e validação da RNA treinada com os nutrientes da primeira coleta da folha da soja.

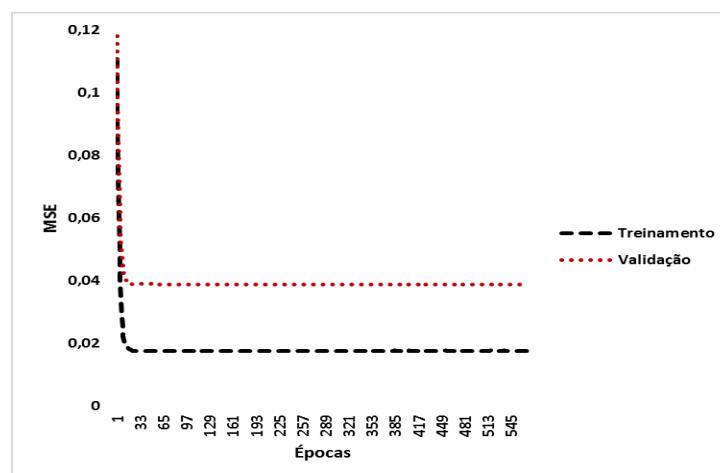


FIGURA 5. Gráfico de comparação dos erros quadráticos médios de treinamento e validação até o término da execução da RNA treinada com os nutrientes da primeira coleta da folha da soja.

Como o melhor modelo de RNA foi encontrado para a primeira coleta dos nutrientes da folha da soja, utilizou-se a Figura 5 para mostrar os erros quadráticos médios de treinamento e validação da rede até o término de sua execução, onde o erro de validação passou a aumentar. Pode se perceber que o erro de treinamento começou em 0,1 e abaixou para 0,02, ao longo das épocas de treinamento, sendo sempre menor que o erro de validação da rede.

CONCLUSÕES

Pela análise do coeficiente de determinação da predição da produtividade da soja, a primeira coleta é a única época na qual é viável uma RNA predizer a produtividade da soja, pois como a primeira coleta deve ser realizada em um estágio bem inicial do desenvolvimento da planta, isto permite ao produtor reverter possíveis produtividades baixas a tempo da colheita.

A análise dos resultados obtidos com a separação das áreas A e B, foram melhores do que os resultados das áreas juntas, pois em todas as coletas o coeficiente de determinação apresentou 31 a 61% da variância dos dados reais. As topologias dos melhores modelos foram com apenas uma camada intermediária e com muito poucos neurônios (apenas um na segunda e na terceira coleta). A correlação dos dados preditos e reais pelo modelo de rede neural da segunda e terceira coleta (com melhores coeficientes de determinação) é forte e correlação moderada negativa para a primeira coleta dos nutrientes da folha da soja. O erro variou de 0,01 a 0,02.

Para a área B, observou-se pelo coeficiente de determinação que as coletas 1 e 3 tiveram modelos que explicaram de 3 a 4% dos dados reais e mesmo a coleta 2 que obteve um coeficiente de determinação maior, o modelo explicou apenas 36% dos dados. Os erros quadráticos médios foram um pouco maiores também que os erros da área A e as correlações variaram de fraca a moderada correlação dos dados reais e preditos.

Portanto, este trabalho mostrou que é possível obter um modelo de predição da produtividade da soja com os nutrientes da folha, pois uma rede neural artificial do tipo Perceptron Multicamadas utilizando o algoritmo backpropagation conseguiu explicar 74% da variação dos dados reais, utilizando para isto, 5 neurônios na camada de entrada, um neurônio em uma única camada intermediária e um neurônio na camada de saída.

A correlação dos dados reais e preditos pelo melhor modelo da técnica de inteligência artificial obtido ficou classificada como forte negativa ($R_p = -0,86$), e o critério de parada antecipada baseado na validação teve o ponto considerado como mínimo local na 67ª época.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), à Universidade Tecnológica Federal do Paraná (UTFPR) e ao Ministério de Agricultura, Pecuária e Abastecimento (MAPA) pelo apoio financeiro e incentivo à pesquisa.

REFERÊNCIAS

BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. 2. ed. Rio de Janeiro: LTC, 2007. 226 p.
CALLEGARI, J.; SIDIA, M. **Bioestatística: princípios e aplicações**. Porto Alegre: Artemed, 2003. 253 p.

- CYBENKO, G. **Approximations by superpositions of sigmoidal functions**. 2. ed. Mathematics of Control, Signals and Systems, 1989. p. 303-314.
- EMBRAPA. **Sistema brasileiro de classificação de solo**. Rio de Janeiro: CNPSO, 2006. 353 p.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. de. **Inteligência Artificial: Uma abordagem de aprendizado de máquina**. 1. ed. Rio de Janeiro: LTC, 2011. 378 p.
- GANDHI, A. P. Quality of soybean and its food products. **Internacional Food Resource Journal**, v. 16, p. 11-19, 2009.
- HAYKIN, S. **Redes neurais: princípios e prática**. 2. ed. Porto Alegre: Bookman, 2001. 898 p.
- KHAIRUNNIZA-BEJO, S., MUSTAFFHA, S., ISMAIL, W. I. W. Application of Artificial Neural Network in Predicting Crop Yield: A Review. **Journal of Food Science and Engineering**, v. 4, p. 1-9, 2014.
- MITCHELL, T. M. **Machine Learning**. 1. ed. New York: McGraw-Hill, Inc., 1997. 432 p.
- MONDO, V. H. V., JUNIOR, F. G. G., PINTO, T. L. F., MARCHI, J. L. DE MOTOMIYA, A. V. DE A., MOLIN, J. P., CICERO, S. M. Spatial variability of soil fertility and its relationship with seed physiological potential in a soybean production area. **Revista Brasileira de Sementes**, v. 34, p. 193–201, 2012.
- MUNDSTOCK, C. M., THOMAS, A. L. **Soja: fatores que afetam o crescimento e o rendimento de grãos**. Porto Alegre: Evangraf, 2005. 31 p.
- PONTES, A. C. F.; CORRENTE, J. E. Comparações múltiplas não-paramétricas para o delineamento com um fator de classificação simples. **Revista de Matemática e Estatística**, v. 19, p. 179–197, 2001.
- RITCHIE, S.W. **Como a planta de soja se desenvolve**. Piracicaba: POTAFOS, 1998. 23 p.
- ROSOLEM, C. A. Recomendação e aplicação de nutrientes via foliar. Lavras: UFLA/FAEPE, 2002. 98 p.
- SILVA, I. N. D., SPATTI, D. H., FLAUZINO, R. A. **Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático**. 1. ed. São Paulo: Artliber, 2010. 399 p.
- SU, C. T., TONG, L. I., LEOU, C. M. Combination of time series and neural network for reliability forecasting modeling. **Journal of the Chinese Institute of Industrial Engineers**, Beijing, v. 14, n. 4, p. 419-429, 1997.
- TEDESCO, M. J., GIANELLO, C., BISSANI, C. A., BOHNEN, H., VOLKWEISS, S. J. **Análise de solo, plantas e outros materiais**. 2. ed. Porto Alegre, 1995. 174 p.
- VITTI, G. C., TREVISAN, W. **Manejo de macro e micronutrientes para alta produtividade da soja**. Piracicaba, 2000. 16 p.
- WOLLENHAUPT, N. C., MULLA, D. J., CROWFORD, G. C. Soil sampling and interpolation techniques for mapping spatial variability of soil properties. **The Site-Specific Management for Agricultural Systems**, p. 19-53, 1997.