



Deep Learning para a seleção de variáveis independentes na estimativa da temperatura mínima do ar mensal

Frank F. Capuchinho¹ (PG)*, Francisco R. de Melo¹ (PQ), Kedinna D. de Sousa² (PG), Gabriella A. M. Campos¹ (PG), Sabrina D. de Oliveira¹ (PG), Luan L. Leão¹ (IC), Jefferson S. Silva¹ (PG)

*frankfreirec@gmail.com

¹ Universidade Estadual de Goiás, Campus Central - Sede: Anápolis – CET. Br 153 Quadra Área Km 99 Zona Rural, Anápolis - GO, Brasil. E-mail: frankfreirec@gmail.com.

² Universidade Federal de Goiás-UFG, Campus Samambaia. Rodovia Goiânia / Nova Veneza, Km 0, CEP 74001-970, Goiânia, GO, Brasil.

Resumo: Existem diversos fatores geográficos e climáticos que afetam a T_{min}, sendo importante identificar aqueles que possibilitam melhor precisão das estimativas. O objetivo deste trabalho foi o de utilizar o Deep Learning para selecionar as variáveis independentes que proporcionam melhores estimativas da T_{min} mensal, para a região Centro-Oeste do Brasil. Foram utilizados dados de estações meteorológicas localizadas na região Centro-Oeste do Brasil. As variáveis utilizadas para as estimativas da T_{min}, foram: altitude (Alt), latitude (Lat), longitude (Long), umidade relativa do ar (UR) e velocidade do vento (VV). Foram estudadas cinco combinações de cenários (C1 a C5) dos parâmetros explicativos. As Redes Neurais Artificiais (RNAs) foram desenvolvidas, para cada mês do ano e para cada cenário, totalizando 60 RNAs. Para avaliar a precisão das estimativas foi utilizado o coeficiente de correlação de Pearson (r). Somente no C5 para os meses de maio (28%) e setembro (38%) é que foram obtidos (r) classificados como fracos (r < 0,4). Exceto nestas circunstâncias (C5 em maio e setembro), as diferentes combinações de variáveis de entrada e a aplicação de Deep Learning foram capazes de estimar a T_{min} mensal, para a região Centro-Oeste do Brasil com precisão satisfatória.

Palavras-chave: Inteligência Artificial. Agrometeorologia. Séries Temporais. Big Data.

Introdução

O desenvolvimento fisiológico de animais e vegetais está relacionado diretamente com temperatura do ar. Isso porque algumas espécies são mais suscetíveis que outras, especificamente, aos níveis mínimos de temperatura do ar. Nesse contexto, determinar a temperatura mínima do ar é importante pois possibilita que os produtores rurais e os órgãos governamentais possam planejar as atividades agropecuárias de maneira otimizada (REIS et al., 2017).





Existem diversos fatores geográficos e climáticos que afetam a T_{min}, como por exemplo a altitude, latitude, longitude, umidade relativa do ar e velocidade do vento. Dessa maneira, torna-se importante identificar os fatores e elementos de forma a possibilitar uma melhor precisão nas estimativas da T_{min}.

Dentre os métodos que podem ser aplicados para identificar a relevância das variáveis explicativas na estimativa da T_{min}, tem-se as Redes Neurais Artificiais (RNA's) Deep Learning. As RNAs são sistemas computacionais constituídos por unidades de processamento simples (neurônios artificiais), conectadas entre si de maneira específica para desempenhar determinada tarefa (BINOTI et al., 2013; HAYKIN, 1999).

Já a abordagem Deep Learning apresenta diversas camadas intermediárias e 'N' neurônios, com alterações lineares e não lineares. Usualmente a abordagem Deep Learning tem sido aplicada para a resolução de problemas que envolvam um grande volume de dados (*Big Data*), como é o caso do estudo de séries temporais climáticas (FERLIN e REZENDE, 2019). Portanto, o objetivo deste trabalho foi o de utilizar o Deep Learning para selecionar as variáveis independentes que proporcionam melhores estimativas da T_{min} mensal, para a região Centro-Oeste do Brasil.

Material e Métodos

O trabalho foi desenvolvido no Laboratório de Engenharia Agrícola da Universidade Estadual de Goiás, Campus Central – CET, sede Anápolis/GO. A região estudada foi o Centro-Oeste do Brasil, nos estados de Goiás, Mato Grosso, Mato Grosso do Sul e Distrito Federal de 81 estações meteorológicas localizadas nestes estados.

Foram utilizados dados geográficos (altitude, latitude e longitude) e dados climáticos diários da temperatura mínima do ar (T_{min}), velocidade do vento (VV) e umidade relativa do ar (UR) de estações automáticas. Os dados foram obtidos na plataforma online do Instituto Nacional de Meteorologia (2020), referentes ao período com disponibilidade de dados entre 2000 e 2020.





Para preencher as falhas de dados observados nas estações meteorológicas automáticas e completar a série estipulada 30 anos (1990-2020), foi utilizado banco de dados desenvolvido por Xavier et al. (2015) para o Brasil no *software* R versão 3.6.1. Em cada município, foram determinadas as médias mensais da T_{min} (°C), UR (%), e VV (m.s⁻¹), de modo a completar os 12 meses do ano, por um período de 30 anos. Foi utilizada uma amostragem de 60% para a etapa de treinamento da rede e 20% para a etapa de validação e 20% para o teste (COUTINHO et al., 2018).

Foram desenvolvidas Redes Neurais Artificiais Deep Learning (RNAs) no *software* R, com configuração de três camadas intermediárias (17-7-3) e função de transferência *rectified linear unit* (ReLU). Para as variáveis de entrada foram analisados 05 diferentes cenários (C1 a C5), sendo eles: C1 (Alt, Lat, Long, UR e VV); C2 (Alt, Lat, Long e UR); C3 (Alt, Lat e Long); C4 (Alt, Lat, Long e VV) e C5 (Alt, Lat e VV). Dessa maneira, foram geradas RNAs para cada mês do ano e para cada cenário analisado, totalizando 60 RNAs.

Para a validação das redes em relação aos valores observados os estimados em cada cenário, foi utilizado o coeficiente de correlação de Pearson 'r', para as parcelas (20%) dos dados que não foram utilizadas na etapa de treino e validação e em todos os meses do ano, para a temperatura mínima do ar mínima, na escala mensal.

Resultados e Discussão

Na Figura 1 são apresentados os resultados de Coeficiente de Correlação de Pearson 'r' entre valores observados e simulados pelas Redes Neurais Artificiais (RNAs) nas estimativas das temperaturas do ar mínima (T_{min}) em diferentes cenários (C1-C5) para o Centro-Oeste do Brasil.

Os coeficientes de correlação de Pearson variaram de 94,1% (C3 no mês de abril) a 28,5% (C5 no mês de maio). Somente no C5 para os meses de maio (28%) e setembro (38%) é que foram obtidos coeficientes de correlação classificados como uma correlação fraca positiva (0,1 – 0,39) (DANCEY e REIDY, 2005).



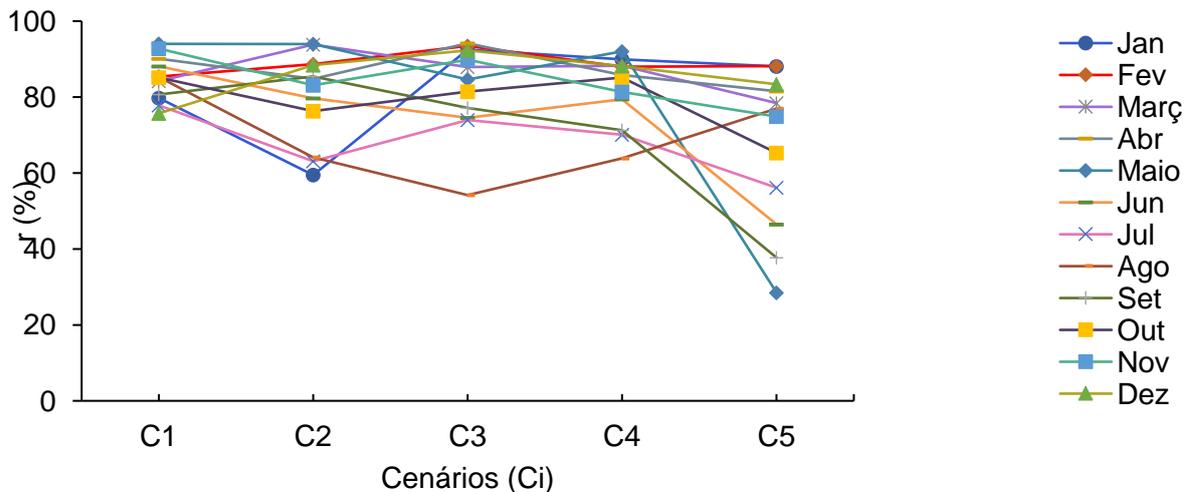


Figura 1 - Coeficiente de Correlação de Pearson (r, em %) entre valores observados e simulados pelas Redes Neurais Artificiais (RNAs) na estimativa da temperatura do ar mínima (Tmin) mensal, em diferentes cenários (C1-C5) para o Centro-Oeste do Brasil.

A correlação classificada como moderada positiva (0,4 – 0,69) (DANCEY e REIDY, 2005). Foi observado no cenário C2 em janeiro (59%), julho (63%) e agosto (64%); cenários C3 e C4 obtiveram 54 e 64%, respectivamente, ambos no mês de agosto e, no C5 junho (46%), julho (56%), e outubro (65%). Os restantes dos meses e cenários, obtiveram resultados satisfatórios de coeficiente de correlação ($r \geq 0,7$) entre valores estimados e simulados pelas RNAs, indicando que essa abordagem é capaz de estimar a Tmin mensal, com precisão satisfatória, para a região Centro-Oeste do Brasil.

Trabalhos desenvolvidos a exemplo de Prasad et al. (2017), verificaram 23 variáveis de entrada, no qual o menor e o maior valor de r (0,29 e 0,96) foi na combinação de 03 e 10 variáveis de entrada, respectivamente. Já Meenal et al. (2018) analisaram redes neurais artificiais com 07 e 05 variáveis de entrada. O menor valor ($r = 0,77$) foi obtido na rede com 05 variáveis independentes enquanto o maior valor ($r = 0,98$) foi obtido na rede com 07 neurônios na camada de entrada.

Para as RNAs que obtiveram 'r' classificados como moderados e altos, a abordagem Deep Learning foi capaz de estimar a Tmin, com correlações satisfatórias em todos os cenários estudados (5, 4 e 3 neurônios na camada de entrada).





Considerações Finais

Exceto no cenário C5 (Alt, Long e VV) nos meses de maio e setembro, as diferentes combinações de variáveis de entrada e a aplicação de Deep Learning foram capazes de estimar a T_{min} mensal, para a região Centro-Oeste do Brasil com precisão satisfatória.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) pela bolsa concedida ao autor e ao Programa de Pós-Graduação *Stricto Sensu* em Engenharia Agrícola (PPGEA) da Universidade Estadual de Goiás.

Referências

- BINOTI, M.L.M. DA S.; BINOTI, D.H.B.; LEITE, H.G. Aplicação de Redes Neurais Artificiais para estimação da altura de povoamentos equiâneos de eucalipto. **Revista Árvore**, Viçosa, v.37, n.4, p.639–645, 2013.
- COUTINHO, E.R.; SILVA, R.M.; MADEIRA, J.G.F.; COUTINHO, P.R.O.S.; BOLOY, R.A.M.; DELGADO, A.R.S. Application of artificial neural networks (ANNs) in the gap filling of meteorological time series. **Revista Brasileira de Meteorologia**, São José dos Campos, v.33, n.2, p.317-328, 2018.
- DANCEY, C.P.; REIDY, J. **Estatística sem matemática**: para psicologia usando SPSS para Windows. 3.ed. Porto Alegre: Artmed, 2006. 608p.
- FERLIN, E.P.; REZENDE, D.A. *Big Data* aplicado à cidade digital estratégica: estudo sobre o volume de dados das aplicações *smart city*. **Revista Gestão & Tecnologia**, Pedro Leopoldo, v. 19, n. 2, p. 175-194, 2019.
- HAYKIN, S. **Neural Networks and Learning Machines**. 3.ed. New Jersey: Pearson Education, 1999. 906p.
- INSTITUTO NACIONAL DE METEOROLOGIA. **Banco de dados meteorológicos para ensino e pesquisa - INMET**, 2020. Disponível em: < <https://www.inmet.gov.br/> >. Acesso em: 20 de nov. de 2020.
- MEENAL, R.; SELVAKUMAR, A.I. Assessment of SVM, empirical and ANN based solar radiation prediction models with most influencing input parameters. **Renewable Energy**, Amsterdã, v.121, p.324-343, 2018.
- PRASAD, R.; DEO, R. C.; LI, Y.; MARASENI, T. Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS AND MODWT algorithm. **Atmospheric Research**, Amsterdã, v.197, p.42-63, 2017.
- REIS, C.J.; BEIJO, L.A.; AVELAR, F.G. Temperatura mínima esperada para Piracicaba-SP via distribuições de valores extremos. **Revista Brasileira de Agricultura Irrigada**, Fortaleza, v.11, n.4, p.1639-1650, 2017.
- XAVIER, A.C.; KING, C.W.; SCANLON, B.R. Daily gridded meteorological variables in Brazil (1980-2013). **International Journal of Climatology**, Reading, v.36, p.2644-2659, 2015.

