

UTILIZAÇÃO DO ALGORÍTIMO FLORESTA ALEATÓRIA NA ESTIMATIVA DA EVAPORAÇÃO DO TANQUE CLASSE “A” EM AMBIENTE PROTEGIDO

Pietros André Balbino dos Santos¹, Thiago Henrique Burgarelli Teixeira², Gustavo Ferreira Jardim³, Luiz Gonsaga de Carvalho⁴, Felipe Scherz⁵, Victor Buono da Silva Baptista⁶

¹ Engenheiro Agrícola, Doutorando em Recursos hídricos em Sistemas Agrícolas, UFLA-MG, pietros.balbino@gmail.com

² Engenheiro Agrícola, Doutorando em Recursos hídricos em Sistemas Agrícolas, UFLA-MG, teixeira.agricola@gmail.com

³ Graduando em Engenharia Agrícola, UFLA-MG, gustavo.jardim@estudante.ufla.br

⁴ Engenheiro Agrícola, Professor Titular, DEA/UFLA-MG, lgonsaga@ufla.br

⁵ Engenheiro Agrônomo, Prof. Adjunto I, DEA/UFLA-MG.

⁶ Engenheiro Agrícola, Prof. Adjunto I, DEA/UFLA-MG, victor.buonosb@ufla.br

Apresentado no

XLIX Congresso Brasileiro de Engenharia Agrícola - CONBEA 2020

23 a 25 de novembro de 2020 - Congresso On-line

RESUMO: A estimativa da evaporação do tanque Classe A pode auxiliar produtores no manejo adequado da irrigação. Assim, objetivou-se com o presente estudo avaliar o algoritmo Floresta Aleatória (FA) na estimativa da evaporação do tanque Classe A em ambiente protegido por meio de diferentes combinações de variáveis meteorológicas. A FA foi implementada pelo WEKA e foram utilizados duas diferentes configurações. Os dados de entrada foram temperatura máxima (Tmax), mínima (Tmin) e média (Tmed) do dia, além da umidade relativa máxima (URmax), mínima (URmin) e média (URmed) do ar utilizados na estimativa da evaporação. Os dados de entrada foram arranjados em 3 configurações distintas (C1, C2 e C3). O algoritmo FA foi capaz de estimar a evaporação de forma confiável em todas as situações avaliadas, entretanto o arranjo C1, com todas variáveis meteorológica, combinado com a configuração FA1 (*Break ties randomly when several attributes look equally good* não ativado associado à 100 interações) resultou nos melhores índices estatísticos.

PALAVRAS-CHAVE: Aprendizagem de Máquina, Ambiente Protegido, Dados Meteorológicas.

USE OF ALGORITHM RANDOM FOREST IN THE ESTIMATION OF CLASS "A" PAN EVAPORATION UNDER GREENHOUSE

ABSTRACT: Estimating of the Class “A” pan evaporation can assist producers in the proper management of irrigation. Thus, the objective of this study was to use the Random Forest (RF) algorithm to estimate the evaporation of the Class “A” pan in a protected environment by means of different combinations of meteorological variables. The RF was implemented by WEKA and two different configurations were used. The input data were maximum (Tmax), minimum (Tmin) and average (Tave) daily temperature, in addition to the maximum (URmax), minimum (URmin) and average (URave) relative humidity used to estimate evaporation. The input data were arranged in 3 different configurations (C1, C2 and C3). The RF algorithm was able to reliably estimate evaporation in all situations evaluated, however the C1 arrangement, with all meteorological variables, combined with the RF1 configuration (no activation of Break ties

randomly when several attributes look equally good associated with 100 interactions) resulted in the best statistical indices.

KEYWORDS: Greenhouse, Machine Learning, Weather Data.

INTRODUÇÃO: O tanque Classe “A” é utilizado para obter a evaporação, e a intensidade dessa evaporação está relacionada com a condição meteorológica do ambiente. O tanque Classe “A” foi desenvolvido pelo Serviço Meteorológico Norte-Americano (U.S.W.B.) e é uma das técnicas utilizadas na condução da irrigação (PEREIRA et al., 1997). Entretanto, a aquisição deste equipamento requer um investimento inicial relativamente alto, o que inviabiliza sua utilização em determinadas situações. A utilização de métodos de estimativa é necessária visando uma forma alternativa de obter a evaporação. No entanto, a eficácia do método influencia na qualidade da estimativa. A Floresta Aleatória (FA) é uma opção de método de modelagem de dados, e é um método estatístico não paramétrico (Breiman, 2001). Além disso é uma técnica de classificação e regressão cuja função principal é fazer crescer árvores de decisão em conjunto de forma que a correlação entre as árvores permaneça a mais baixa possível. Dessa forma, esta técnica tem sido utilizada na estimativa em diferentes áreas de estudo. Diante do exposto, o presente trabalho teve como objetivo analisar o desempenho do algoritmo FA na estimativa da evaporação do Tanque Classe “A” em ambiente protegido utilizando diferentes arranjos de dados meteorológicos como varável de entrada.

MATERIAL E MÉTODOS: Os dados meteorológicos e evaporação foram obtidos em casa de vegetação coberta com filme de polietileno de baixa densidade (PEBD), leitoso, impermeável e transparente, localizada na Universidade Federal de Lavras, Minas Gerais, Brasil. No centro da estufa foi instalado a estação meteorológica Vantage Pro2 da Agrosystem, para aquisição dos dados de temperatura e umidade relativa do ar. Foi instalado o evaporímetro Tanque Classe “A” sobre um estrado de madeira para aquisição dos dados de evaporação. As leituras foram realizadas diariamente nos meses de agosto, setembro e outubro de 2018. Os dados obtidos foram utilizados para a construção da floresta aleatória (FA). A FA foi implementada pelo Waikato Environment for Knowledge Analysis (WEKA) versão 3.8.2 © 1999-2017 desenvolvido pela University of Waikato, Hamilton, Nova Zelândia (1993), tendo como base Breiman, 2001 e disponível gratuitamente. A FA tem etapas que podem ser resumidas conforme mostrado na Figura 1.

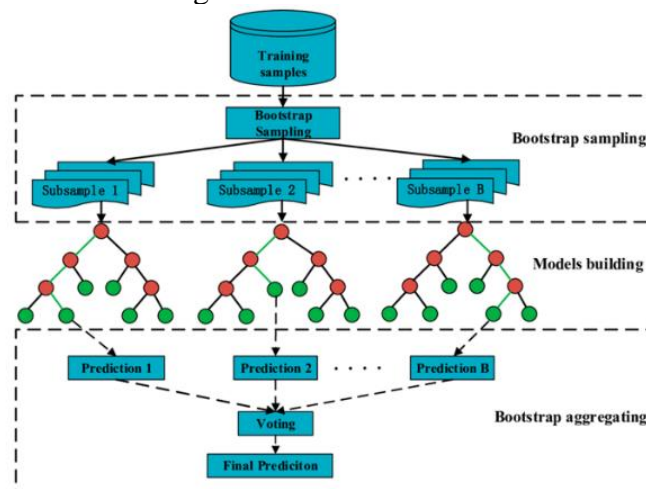


Figura 1. Esquema de processamento da Random Forest (Fonte: Wang et al., 2019).

Foram utilizadas duas configurações distintas na implementação da FA (Tabela 1).

Tabela 1. Configuração da implementação da Floresta aleatória pelo WEKA.

	Tmean	
	FA1	FA2
Break ties randomly when several attributes look equally good		x
Size of each bag, as percentage of the training set size	100	100
Number of iterations	100	500
Minimum number of instances per leaf	1	1
Maximum depth of the tree	unl.	unl.

Os dados de entrada foram temperatura máxima (Tmax), mínima (Tmin) e média (Tmed) do ar, além da umidade relativa máxima (URmax), mínima (URmin) e média (URmed) do ar. O dado de saída (estimado) foi a evaporação do tanque Classe “A”. Desses dados, 75% foram utilizados no treinamento da FA. Os dados de entrada foram combinados de diferentes formas, buscando a maior capacidade preditiva e o menor erro (Tabela 2).

Tabela 2. Arranjo das variáveis meteorológicas utilizadas na estimativa

Combinação	Tmax	Tmin	Tmed	URmax	URmin	URmed
C1	X	X	X	X	X	X
C2	X	X	X			
C3				X	X	X

Cada modelo gerou indicadores estatísticos de performance que foram: coeficiente de correlação de Pearson (r), que indica o grau de dispersão do dado em relação à média; erro absoluto médio (EAM) e raiz quadrada média do erro (RQME) que indica o quão próximos os valores previstos estavam dos observados, representando a precisão de cada modelo.

RESULTADOS E DISCUSSÃO: Na tabela 3 estão dispostos os resultados dos índices estatísticos de cada configuração de rede avaliada e cada combinação de dados de entrada. Observa-se que todas as configurações com cada combinação apresentaram valores de r acima de 0,75, indicando uma boa capacidade preditiva do algoritmo Floresta Aleatória nesse tipo de análise. Entretanto, a configuração FA1 com a combinação C1, que utiliza todos os dados de entrada, apresentaram os melhores resultados com r, EAM e RQME igual a 0,84, 0,74 e 0,90, respectivamente. Esse resultado indica que, apesar dos arranjos de entrada C2 e C3 resultarem em uma boa estimativa, a combinação C1 é a melhor na estimativa da evaporação do tanque Classe A durante esse período de análise.

Tabela 3. Índices estatísticos de cada configuração da floresta aleatória e arranjo avaliada

	FA1			FA2		
	r	EAM	RQME	r	EAM	RQME
C1	0,84	0,74	0,90	0,82	0,78	0,92
C2	0,79	0,91	1,13	0,75	0,94	1,17
C3	0,78	0,82	1,01	0,75	0,89	1,08

De acordo com os resultados observados, a não ativação do *Break ties randomly when several attributes look equally good* associado à 100 interações aumentou a capacidade de predição, no entanto, de forma geral, o algoritmo FA foi capaz de estimar com boa acurácia. Essa capacidade

preditiva da FA pode ser explicada pelas vantagens de não fazer suposições distributivas sobre os preditores, medindo a importância de cada variável e sendo menos sensível a ruído ou overfitting (Armitage; Ober, 2010).

Na figura 2 estão plotados a importância de cada atributo de entrada na estimativa da evaporação do Tanque Classe “A” do arranjo C1, que apresentou os melhores índices estatísticos. Observa-se que a URmed e URmin representa os maiores pesos na estimativa da evaporação na configura FA1. Tmax e Tmin apresentaram os menores pesos. Na configuração FA2 a URmin e URmed apresentaram os maiores pesos na estimativa. Esse resultado mostra a influência da UR na evaporação, apesar da temperatura indicar a energia disponível no sistema, os níveis de evaporação estão atrelados à variação da UR.

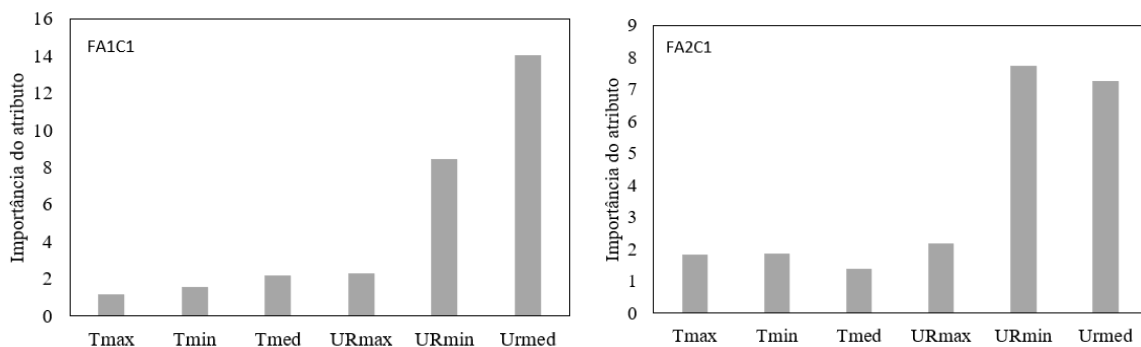


Figura 2. Importância dos atributos de entrada na estimativa

CONCLUSÕES: O presente estudo mostrou que a FA é capaz de estimar a evaporação do tanque Classe A com acurácia, e o arranjo com todas as variáveis disponível com a configuração apresentada na FA1 resultaram nos melhores índices estatísticos. Portanto essa combinação é mais indicada na estimativa da evaporação em ambiente protegido.

AGRADECIMENTOS: À Universidade Federal de Lavras – UFLA, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES pelas bolsas concedidas e apoio financeiro à pesquisa.

REFERÊNCIAS:

ARMITAGE, D. W.; OBER, H. K. A comparison of supervised learning techniques in the classification of bat echolocation calls. **Ecological Informatics**, 5(6), 465–473, 2010.

BREIMAN, L. Random forests. **Machine Learning**, 45(1), 5–32, 2001.

PEREIRA, A.R.; VILLA NOVA, N.A.; SEDIYAMA, G.C. *Evapo(transpi)ração*. Piracicaba: FEALQ, 1997. 183 p.

WANG, H. et al. Intelligent identification of maceral components of coal based on image segmentation and classification. **Applied Sciences (Switzerland)**, 9(16), 1–15, 2019.