

Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Uso de machine learning no manejo da irrigação e estimativa da produtividade
de milho em Piracicaba, SP

Bruna Marques de Queiroz

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Engenharia
de Sistemas Agrícolas

Piracicaba
2023

Bruna Marques de Queiroz
Engenheira Agrônoma

Uso de machine learning no manejo da irrigação e estimativa da produtividade de milho
em Piracicaba, SP

versão revisada de acordo com a Resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **JARBAS HONORIO DE MIRANDA**

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Engenharia
de Sistemas Agrícolas

Piracicaba
2023

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP**

Queiroz, Bruna Marques de

Uso de machine learning no manejo da irrigação e estimativa da produtividade de milho em Piracicaba, SP / Bruna Marques de Queiroz. - - versão revisada de acordo com a Resolução CoPGr 6018 de 2011. - - Piracicaba, 2023.

132 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. *Zea mays* 2. Gestão hídrica 3. Predição da produtividade 4. Rendimento agrícola 5. Predição da evapotranspiração I. Título

DEDICATÓRIA

Dedico,

Aos meus avós Romeu e Terezinha Queiroz, à minha avó Maria Aparecida Marques e ao meu padrinho Ivan Marques que hoje estão na eternidade e sempre me apoiaram. Suas lembranças me inspiram e me fazem persistir.

AGRADECIMENTOS

A Deus e a Nossa Senhora pela proteção e iluminação.

Aos meus pais, Josete e Carlos, eu devo a vida e todas as oportunidades que nela recebi e espero um dia poder-lhes retribuir todo amor e confiança.

Ao meu irmão Diogo, pelo apoio e companheirismo, principalmente nos momentos mais difíceis.

Ao meu avô José, pela força, carinho e palavras de sabedoria durante toda essa trajetória.

Aos meus familiares.

À Universidade de São Paulo, sobretudo o campus da Escola Superior de Agricultura “Luiz de Queiroz” e todos seus professores e funcionários.

Ao meu orientador Jarbas Honorio de Miranda, pelos ensinamentos, oportunidades e confiança.

À Professora Sônia Maria De Stefano Piedade, pela amizade e conselhos dados para minha vida profissional e pessoal.

Ao Professor Paulo Leonel Libardi, pelo constante apoio e ensinamentos.

Aos funcionários do Departamento de Engenharia de Biossistemas, especialmente Antônio Agostinho Gozzo (Seu Antônio), Francisco Bernardo Dias (Chiquinho), Gilmar Batista Grigolon e Paula Alessandra Bonassa Pedro, pela amizade e auxílio na execução dos experimentos.

Aos integrantes do Grupo de Práticas e Estudos em Água e Solo (GPEAS) pela amizade e conselhos, especialmente Isabela, Janderson, Lorena, Martiliana e Rafaela.

A todos os meus amigos, especialmente Caroline, Jéssica, Larissa e Mikaela por estarem comigo desde o primeiro dia de ESALQ. E ao meu grande amigo Vitor, pelo companheirismo, compreensão e apoio.

Aos professores do Departamento de Engenharia de Biossistemas, pelos ensinamentos e orientação no dia a dia.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos durante o período.

A todos aqueles que de alguma forma positiva fizeram parte do meu processo de formação.

E por fim, a todos os que contribuem para o avanço da ciência no Brasil.

EPÍGRAFE

“Estou entre aqueles que acham que a ciência tem uma grande beleza.”

Marie Curie

SUMÁRIO

RESUMO	9
ABSTRACT	10
1. INTRODUÇÃO GERAL	11
Referências	13
2. REVISÃO BIBLIOGRÁFICA	17
2.1. Aspectos da cultura do milho	17
2.2. Manejo da água no cultivo de milho	19
2.3. Modelagem na agricultura	20
2.3.1. Modelos de simulação de cultura	21
2.3.2. Algoritmos de machine learning	23
Referências	26
3. USO DE MACHINE LEARNING NA ESTIMATIVA DA EVAPOTRANSPIRAÇÃO DE REFERÊNCIA EM PIRACICABA, SP	35
Resumo	35
Abstract	35
3.1. Introdução	36
3.2. Material e Métodos	38
3.2.1. Área de estudo e coleta dos dados	38
3.2.2. Determinação da ETo pela abordagem do modelo FAO-56 Penman-Monteith....	40
3.2.3. Pré-processamento e desenvolvimento dos modelos	42
3.2.4. Algoritmos aplicados	43
3.2.5. Avaliação da performance dos modelos	46
3.3. Resultados e Discussão	47
3.3.1. Correlação entre as variáveis do banco de dados	47
3.3.2. Avaliação dos conjuntos	49
3.3.3. Avaliação dos algoritmos para experimento de campo	54
3.4. Conclusões	59
Referências	59
4. USO DE MACHINE LEARNING E MODELO DSSAT/CERES-MAIZE NA ESTIMATIVA DA PRODUTIVIDADE DE MILHO EM PIRACICABA, SP.....	63
Resumo.....	63

Abstract.....	63
4.1. Introdução.....	64
4.2. Material e Métodos.....	66
4.2.1. Descrição sobre o modelo CERES-MAIZE.....	66
4.2.2. Propriedades físicas do solo.....	67
4.2.3. Dados meteorológicos.....	67
4.2.4. Dados do genótipo.....	68
4.2.5. Outras informações simuladas.....	69
4.2.6. Criação do banco de dados, pré-processamento e desenvolvimento do modelo....	70
4.2.7. Algoritmos aplicados.....	71
4.2.8. Avaliação da performance dos modelos.....	75
4.3. Resultados e Discussão.....	75
4.3.1. Produtividade de milho simulada pelo DSSAT.....	75
4.3.2. Correlação entre as variáveis de entrada e produtividade.....	79
4.3.3. Avaliação dos algoritmos.....	83
4.4. Conclusões.....	87
Referências.....	88
5. MANEJO DA IRRIGAÇÃO E ESTIMATIVA DA PRODUTIVIDADE DE MILHO PARA A REGIÃO DE PIRACICABA, SP, UTILIZANDO MACHINE LEARNING.....	93
Resumo.....	93
Abstract.....	93
5.1. Introdução.....	94
5.2. Material e Métodos.....	96
5.2.1. Localização da área experimental.....	96
5.2.2. Preenchimento dos recipientes e instalação dos tensiômetros.....	96
5.2.3. Parâmetros da curva de retenção de água no solo.....	99
5.2.4. Delineamento experimental.....	101
5.2.5. Condução do experimento.....	102
5.2.6. Colheita.....	105
5.2.7. Avaliações biométricas e componentes de produção.....	105
5.2.8. Análises estatísticas.....	105
5.3. Resultados e Discussão.....	106
5.3.1. Condições meteorológicas.....	106
5.3.2. Determinação da ETo.....	109

5.3.3. Manejo da irrigação	111
5.3.4. Respostas biométricas no desenvolvimento vegetativo do milho.....	113
5.3.5. Respostas biométricas no desenvolvimento das espigas de milho	115
5.3.6. Avaliação dos componentes de produção do milho.....	117
5.3.7. Estimativa da produtividade com algoritmos de machine learning	119
5.4. Conclusões	121
Referências.....	122
ANEXOS	127

RESUMO

Uso de machine learning no manejo da irrigação e estimativa da produtividade de milho em Piracicaba, SP

A agricultura contemporânea vem encontrando desafios, tais como, mudanças climáticas, aumento nos custos de produção, escassez de terras agricultáveis e recursos hídricos. Nesse cenário, torna-se cada vez mais necessária a busca por ferramentas para otimizar os processos agrícolas. O milho caracteriza-se como um produto fundamental para o agronegócio brasileiro e sua produção é influenciada pela variabilidade climática, fertilidade do solo, tratamentos culturais, manejo de irrigação e fatores bióticos. Alinhado ao cultivo, o incremento de recursos computacionais e o surgimento da teoria da inteligência artificial, pesquisas utilizando técnicas do aprendizado de máquina vêm demonstrando a capacidade dessas ferramentas em otimizar o uso de recursos e auxiliar em tomadas de decisão. A gestão hídrica associada ao manejo da irrigação, com base na demanda da água pela cultura, representada pela taxa da evapotranspiração, ocupa posição de destaque no cultivo de milho, sendo que para sua obtenção, o método da FAO-56 Penman-Monteith (PM) vem sendo o mais aplicado. Como alternativa a esse método, estudos mostraram que algoritmos de *machine learning* podem ser eficientes em estimar a evapotranspiração de referência (ET_o), bem como o rendimento da produção. Dessa forma, considerando a importância das ferramentas computacionais (e.g., DSSAT, ciência de dados) para a otimização das atividades agrícolas, os principais objetivos dessa pesquisa foram: a) desenvolver modelos preditivos de *machine learning* para a estimativa da ET_o diária e do rendimento agrícola do cultivo de milho na região de Piracicaba e b) avaliar a performance desses modelos comparados a um experimento de campo. Para alcançar esses objetivos, a pesquisa foi dividida em três etapas: etapa 1) uso de uma série climática de 21 anos de dados diários (Posto Meteorológico da ESALQ/USP) para desenvolvimento de modelos preditivos de *machine learning* da ET_o (*support vector regression* (SVR), *histogram-based gradient boosting regression tree* (HGBR) e *multilayer perceptron regressor* (MLPR)); etapa 2) simulação da produtividade de milho com o modelo DSSAT/CERES-MAIZE (*Decision Support System for Agrotechnology Transfer*) para as condições de Piracicaba para uma série de 30 anos de dados climáticos e desenvolvimento de modelos preditivos do rendimento agrícola utilizando os algoritmos *extra-trees* (EXT) e *random forest* (RF) e etapa 3) Aplicação dos modelos desenvolvidos em comparação a dados coletados a partir de um experimento em campo referente à condução do ciclo de cultivo do milho. Diante dos resultados obtidos, houve uma indicativa de que os modelos aplicados apresentaram desempenho satisfatório na predição da ET_o e da produtividade do milho na região de Piracicaba, SP. O algoritmo HGBR obteve estimativas mais precisas para a ET_o em comparação ao SVR e MLPR. Enquanto que, o RF obteve melhor performance na estimativa da produtividade do milho em comparação ao EXT. Portanto, diante das condições da pesquisa, conclui-se que os algoritmos de aprendizado de máquina apresentaram-se como ferramentas eficazes na predição da ET_o e da produtividade do milho para as condições de Piracicaba, SP, inclusive quando comparados à coleta de dados em campo, passando a ser uma alternativa importante a modelos e métodos clássicos utilizados na gestão eficiente do sistema produtivo.

Palavras-chave: *Zea mays*, Gestão hídrica, Predição da produtividade, Rendimento agrícola, Predição da evapotranspiração

ABSTRACT

Use of machine learning in irrigation management and maize yield estimation in Piracicaba, SP

Contemporary agriculture is facing challenges such as climate change, rising production costs, scarcity of arable land and water resources. In this scenario, the search for tools to optimize agricultural processes becomes increasingly necessary. Maize is a fundamental product for Brazilian agribusiness, and its production is influenced by climate variability, soil fertility, cultural practices, irrigation management and biotic factors. Aligned with cultivation, the increase in computational resources and the emergence of artificial intelligence have shown the capacity of machine learning techniques to optimize resource use and aid in decision-making. Water administration associated with irrigation management, based on crop water demand represented by evapotranspiration rate, occupies a prominent position in maize cultivation, with the FAO-56 Penman-Monteith (PM) method being the most applied. As an alternative to this method, studies have shown that machine learning algorithms can be efficient in estimating reference evapotranspiration (ET_o) as well as crop yield. Thus, considering the importance of computational tools (e.g., DSSAT, data science) for optimizing agricultural activities, the main objectives of this research were to: a) develop predictive machine learning models for estimating daily ET_o estimation and maize agricultural yield in the Piracicaba region and b) evaluate the performance of these models compared to a field experiment. To achieve these objectives, the research was divided into three stages: stage 1) using a 21-year climate series of daily data (Meteorological Station of ESALQ/USP) for developing predictive machine learning models of ET_o (support vector regression (SVR), histogram-based gradient boosting regression tree (HGBR) and multilayer perceptron regressor (MLPR)); stage 2) simulating maize productivity with the DSSAT/CERES-MAIZE model (Decision Support System for Agrotechnology Transfer) for Piracicaba conditions based on a 30-year series of climate data and developing predictive models of agricultural yield using the extra-trees (EXT) and random forest (RF) algorithms and 3) applying the developed models in comparison to data collected from a field experiment regarding the maize cultivation cycle. Based on the results obtained, there was an indication that the applied models showed a satisfactory performance in predicting ET_o and maize productivity in the Piracicaba, SP, region. The HGBR algorithm provided more accurate estimates for ET_o compared to SVR and MLPR, while RF performed better in estimating maize yield compared to EXT. Therefore, under the conditions of this research, it is concluded that machine learning algorithms proved to be effective tools in predicting ET_o and maize yield for the conditions of Piracicaba, SP, becoming an important alternative to classical models and methods used in efficient management of the productive system.

Keywords: *Zea mays*, Water management, Productivity prediction, Agricultural yield, Evapotranspiration prediction

1. INTRODUÇÃO GERAL

Diante do crescente aumento da população mundial e sua demanda por alimentos, as projeções dos desafios globais para os próximos 50 anos evidenciam que os principais problemas da agricultura contemporânea serão aspectos relacionados à água, energia e alimentos, sendo que a maior dificuldade estará em aliar a produção de alimentos com a proteção ambiental, o que impulsiona a ciência e os formuladores de políticas a oferecer novas abordagens e soluções para essa problemática (AMELUNG et al., 2020).

Nesse cenário de crescente demanda por alimentos, torna-se necessária a busca por ferramentas capazes de otimizar cada vez mais os processos agrícolas, possibilitando aumento da produtividade de modo consciente (OLIVEIRA et al., 2018).

No cenário dos processos agrícolas, o milho (*Zea mays* L.) constitui-se em um dos mais importantes cereais cultivados, consumidos e estudados no mundo, sendo responsável por fornecer produtos utilizados para a alimentação humana e animal e matérias-primas para a indústria, tais como combustíveis, óleos, bebidas, polímeros, entre outros. A produção mundial de milho corresponde a mais de 1.186 milhões de toneladas, com os EUA figurando como os maiores produtores e consumidores, seguidos pela China e Brasil (GREAVES e WANG, 2017; MIRANDA, 2018; USDA, 2022).

A cadeia produtiva do milho constitui-se em uma das mais importantes do cenário agrícola brasileiro, correspondendo a 37% da produção nacional de grãos. A principal região produtora no Brasil é a Centro-Oeste, seguida pelas regiões Sul, Nordeste, Sudeste e Norte. No Brasil, foram cultivados 21,58 milhões de hectares e produzidos 113,13 milhões de toneladas de milho na safra 2021/22, com produtividade de 5.242 kg ha⁻¹, sendo as maiores áreas observadas nos estados do Mato Grosso, Paraná e Mato Grosso do Sul, os quais foram responsáveis pela produção de 41,62, 16,42 e 12,46 milhões de toneladas, respectivamente. (CONAB, 2022).

Segundo Kipkulei et al. (2022), a produção da cultura de milho é influenciada pela variabilidade climática, fertilidade do solo, tratos culturais, data de semeadura, manejo de irrigação, fatores bióticos, entre outros. Os ganhos de produtividade observados nas últimas safras de milho se referem ao incremento tecnológico desenvolvido através de cultivares mais adaptadas e eficientes técnicas de manejo obtidas através de tomadas de decisões baseadas em informações e insights precisos (DUARTE, 2023).

A aplicação da tecnologia na agricultura tornou-se um recurso indispensável no auxílio de tomadas de decisões, visto que a partir da agricultura digital tornou-se possível unir a ciência de dados e as atividades agrícolas, visando estabelecer uma relação entre o vasto número de dados disponíveis de modo a favorecer a gestão antecipada dos sistemas produtivos (KIM et al., 2019).

Nesse sentido, com o rápido avanço dos recursos computacionais e o surgimento da teoria da inteligência artificial, pesquisas utilizando técnicas do aprendizado de máquina demonstraram a capacidade dessas ferramentas em otimizar o uso de recursos (e.g., água, fertilizantes e energia) a partir de algoritmos que podem analisar os dados e fornecer recomendações precisas sobre a quantidade a ser aplicada, bem como em estimar a produtividade final e gerenciar as safras com base em dados históricos, auxiliando os agricultores a tomar decisões em relação aos momentos ideais de plantio, tratos culturais e colheita (FILIPPI, 2019; MAYA-GOPAL e CHINTALA, 2020; TIBBETTS, 2018).

A disponibilidade e a gestão hídrica ocupam posição de destaque no ciclo de cultivo do milho, uma vez que o excesso ou a falta de água causam injúrias e redução da produtividade das plantas, especialmente quando ocorrem nos estádios de florescimento e enchimento de grãos. Por isso, o conhecimento e entendimento em relação à quantidade de água a ser fornecida para a cultura tornam-se essenciais, a fim de se obter um manejo consciente, o qual permitirá que as plantas se desenvolvam de forma satisfatória (KIPKULEI et al., 2022; PALHARES, 2003; TAIZ e ZEIGER, 2013).

A reposição de água ao solo por meio da irrigação, na quantidade e no momento oportuno, apresenta-se como fator decisivo para maximizar o processo produtivo. Para tal, o manejo da irrigação baseado na necessidade de água da cultura, definida pela demanda evapotranspirométrica, é amplamente utilizado (BERNARDO et al., 2019).

Por representar o principal uso consuntivo da água agrícola, a evapotranspiração (ET) quantifica a demanda hídrica das plantas, sendo, assim, um fator crítico para uma gestão adequada das práticas de irrigação. Portanto, a estimativa precisa desse processo é fundamental no estabelecimento de um sistema sustentável, mitigando os impactos da escassez e do uso excessivo da água (FERREIRA et al., 2019; WANNIARACHCHI e SARUKKALIGE, 2022).

Diferentes métodos foram desenvolvidos para a medição da evapotranspiração de referência (ET_o) incluindo diretos, indiretos e modelos baseados em dados (*data driven models*). Dentre os métodos indiretos, o modelo FAO-56 Penman-Monteith (PM) vem sendo considerado como método padrão. Sua aplicação, porém, é dificultada por necessitar de um número expressivo de variáveis climáticas de entrada, com dados de alta qualidade e integração de constantes empíricas. O que vem de encontro à limitação das estações meteorológicas que não satisfazem esses requisitos de observação, especialmente em países em desenvolvimento (YAMAÇ e TODOROVIC, 2020).

Para contornar essa problemática, o uso da ciência de dados, tais como, *machine learning* vêm sendo aplicadas à estimativa da ET_o como demonstrado por Chen et al., 2020; Chia et al., 2020; Granata, 2019; Mohammadi e Mehdizadeh, 2020; Raza et al., 2020; e Wang et al., 2019. Essa

abordagem pode ser utilizada para determinação da ETo diária, a qual é empregada para a realização do manejo de irrigação no ciclo de cultivo do milho.

Além do manejo da irrigação, modelos capazes de simular a interação solo-planta-atmosfera e auxiliar na previsão da safra de determinada cultura também se destacam como ferramentas essenciais para otimização dos processos agrícolas. Dentre os modelos mais citados na literatura, destaca-se o DSSAT (*Decision Support System for Agrotechnology Transfer*, DSSAT version 4.6, JONES et al., 2003; HOOGENBOOM et al., 2014) como sendo um dos principais sistemas que empregam modelos de simulação de crescimento e desenvolvimento de culturas ao longo do tempo a partir de variáveis climáticas da região, estado nutricional e sanitário das plantas, características de solo e tratos culturais.

No entanto, tais modelos geralmente requerem elevada quantidade de parâmetros de entrada, o que torna mais difícil a sua calibração precisa (XI et al., 2017). Por isso, diferentes abordagens baseadas no aprendizado de máquina para estimar o crescimento e desenvolvimento de plantas vêm sendo amplamente difundidas como alternativas promissoras e ferramentas complementares para a modelagem comumente usada (BASSO e LIU, 2019; CRANE-DOESCH, 2018; MISHRA et al., 2016; MUPANGWA et al., 2020).

Com base na importância dessas ferramentas computacionais para a otimização das atividades agrícolas, a hipótese da presente pesquisa é que o uso de algoritmos de *machine learning* no manejo da irrigação e na estimativa da produtividade do milho pode contribuir na gestão estratégica e eficiente do sistema produtivo.

De modo a testar a hipótese citada, os objetivos dessa pesquisa foram: i) desenvolver modelos preditivos de *machine learning* para a estimativa da ETo diária na região de Piracicaba; ii) desenvolver modelos preditivos de *machine learning* para a estimativa do rendimento agrícola do cultivo de milho; e iii) avaliar a performance desses modelos em um experimento de campo.

Referências

- AMELUNG, W.; BOSSIO, D.; de VRIES, W. Towards a global-scale soil climate mitigation strategy. **Nat Commun**, v. 11, 5427, 2020.
- BASSO, B.; LIU, L. Chapter four - seasonal crop yield forecast: methods, applications, and accuracies. **Advances in Agronomy** 154, 201–255, 2019.
- BERNARDO, S.; MANTOVANI, E. C.; SILVA, D. D.; SOARES, A. A. **Manual de Irrigação**. 9. ed. Viçosa: UFV, 545 p., 2019.

- CHEN, Z.; ZHU, Z.; JIANG, H.; SUN, S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. **Journal of Hydrology**, v. 591, p. 125–286, 2020.
- CHIA, M. Y.; HUANG, Y. F.; KOO, C. H. Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. **Computers and Electronics in Agriculture**, v. 175, p. 1–12, 2020.
- CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos, v. 10, Safra 2022/23 – Primeiro levantamento**, Brasília, p. 1–77, 2022.
- CRANE-DROESCH, A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. **Environmental Research Letters**, v. 13, n. 11, 114003, 2018.
- DUARTE, Y. C. N. **Impacto das mudanças climáticas no Brasil e sua influência no cultivo e na produtividade do milho: uma abordagem multi-modelos**. Tese. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 250 p., 2023.
- FERREIRA, L. B.; da CUNHA, F. F.; de OLIVEIRA, R. A.; FERNANDES FILHO, E. I. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – A new approach. **Journal of Hydrology**, v. 572, p. 556–570, 2019.
- FILIPPI, P.; JONES, E. J.; WIMALATHUNGE, N. S.; SOMARATHNA, P. D.; POZZA, L. E.; UGBAJE, S. U.; BISHOP, T. F. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. **Precision Agriculture**, p. 1-15, 2019.
- GRANATA, F.; GARGANO, R.; de MARINIS, G. Artificial intelligence based approaches to evaluate actual evapotranspiration in wetlands. **Science of the Total Environment**, v. 703, p. 1–15, 2020.
- GREAVES, G. E.; WANG, Y. M. Effect of regulated deficit irrigation scheduling on water use of corn in southern Taiwan tropical environment. **Agricultural Water Management**, v. 188, p. 115–125, 2017.
- HOOGENBOOM, G.; JONES, J. W.; WILKENS, P. W.; PORTE, C. H.; BOOTE, K. J.; HUNT, L. A.; SINGH, U.; LIZASO, J. L.; WHITE, J. W.; URYASEV, O.; ROYCE, F. S.; OGOSHI, R.; GIJSMAN, A. J.; TSUJI, G. Y. **Decision Support System for Agrotechnology Transfer (DSSAT): version 4.6**. DSSAT Foundation, Prosser, Washington, 2014.
- JONES, J. W.; HOOGENBOOM, G.; PORTER, C. H.; BOOTE, K. J.; BATCHELOR, W. D.; HUNT, L. A.; WILKENS, P. W.; SINGH, U.; GIJSMAN, A. J.; RITCHIE, J. T. The DSSAT cropping system model. **European Journal of Agronomy**, v. 18, p. 235–265, 2003.
- KIM, N.; HA, K. J.; PARK, N. W.; CHO, J.; HONG, S.; LEE, Y.W. A Comparison Between Major Artificial Intelligence Models for Crop Yield Prediction: Case Study of the Midwestern United States, 2006–2015. **ISPRS International Journal of GeoInformation**, 8.5: 240, 2019.

- KIPKULEI, H. K.; BELLINGRATH-KIMURA, S.D.; LANA, M. et al. Modelling cropland expansion and its drivers in Trans Nzoia County, Kenya. **Modeling Earth Systems and Environment**, v. 8, p. 5761–5778, 2022.
- MAYA-GOPAL, P. S.; CHINTALA, B. R. Others Big Data challenges and opportunities in agriculture. **Int. J. Agric. Environ. Inf. Syst.** 11, p. 48–66, 2020.
- MIRANDA, R. A. Uma história de sucesso da civilização. Embrapa Milho e Sorgo. **A Granja**, v. 74, n. 829, p. 24-27, 2018.
- MISHRA, S.; MISHRA, D.; SANTRA, G. H. Applications of machine learning techniques in agricultural crop production: A review paper. **Indian Journal of Science and Technology**, v. 9(38), p. 1–14, 2016.
- MOHAMMADI, B.; MEHDIZADEH, S. Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. **Agricultural Water Management**, v. 237, p. 1–13, 2020.
- MUPANGWA, W.; CHIPINDU, L.; NYAGUMBO, I.; MKUHLANI, S.; SISITO, G. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. **SN Applied Sciences** 2, 952, 2020.
- OLIVEIRA, L. A.; MIRANDA, J. H.; COOKE, R. A. C. Water management for sugarcane and corn under future climate scenarios in Brazil. **Agricultural Water Management**, v. 201, p. 199–206, 2018.
- PALHARES, M. **Distribuição e população de plantas e produtividade de grãos de milho**. Dissertação. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 90 p., 2003.
- RAZA, A.; SHOAIB, M.; FAIZ, M. A.; BAIG, F.; KHAN, M. M.; ULLAH, M. K.; ZUBAIR, M. Comparative Assessment of Reference Evapotranspiration Estimation Using Conventional Method and Machine Learning Algorithms in Four Climatic Regions. **Pure and Applied Geophysics**, v. 177, n. 9, p. 4479–4508, 2020.
- TAIZ, L.; ZEIGER, E. *Fisiologia Vegetal*, 5ª Edição. **Artmed**, Porto Alegre, 954 p., 2013.
- TIBBETTS, J. H. The Frontiers of artificial intelligence: Deep learning brings speed, accuracy to the life sciences. **Bioscience**, 68(1), p. 5–10, 2018.
- USDA - United States Department of Agriculture, 2022.
- WANG, S.; LIAN, J.; PENG, Y.; HU, B.; CHEN, H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. **Agricultural Water Management**, v. 221, p. 220–230, 2019.
- WANNIARACHCHI, S.; SARUKKALIGE, R. A Review on Evapotranspiration Estimation in Agricultural Water Management: Past, Present, and Future. **Hydrology** 9, 123, 2022.

XI, M. L.; LU, D. A.; GUI, D. W.; QI, Z. M.; ZHANG, G. N. Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization. **Journal of Hydrology**, 544, p. 456–466, 2017.

YAMAÇ, S. S.; TODOROVIC, M. Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. **Agricultural Water Management**. v. 228, 105875, 2020.

2. REVISÃO BIBLIOGRÁFICA

2.1. Aspectos da cultura do milho

O milho é uma gramínea pertencente à família *Poaceae*, gênero *Zea* e espécie *Zea mays* L. A sua propagação é do tipo sexuada e depende do cruzamento entre plantas para a formação das espigas. O seu sistema radicular é característico das gramíneas, sendo do tipo fasciculado e dependendo das condições do ambiente, pode atingir mais de um metro de profundidade. Ao longo do crescimento da planta, surgem raízes adventícias que são responsáveis por proporcionar à planta maior capacidade de fixação no solo e de se manter ereta (BARROS e CALADO, 2014).

As folhas do tipo lanceoladas, dispostas ao longo do caule de forma alternada, são morfológicamente classificadas como completas, uma vez que apresentam limbo, bainha e pecíolo. É uma planta monóica, ou seja, apresenta flores unissexuais distribuídas no mesmo indivíduo, sendo que a inflorescência masculina se apresenta agrupada em uma estrutura denominada panícula, enquanto que a feminina é a espiga. A planta de milho é classificada como protândrica, pois a inflorescência masculina se desenvolve antes da feminina (BARROS e CALADO, 2014).

Os eventos fenológicos que ocorrem durante o ciclo de milho foram descritos por Ritchie et al. (1993), com uma representação por letras e números, facilitando a identificação de cada etapa (Tabela 1).

Tabela 1. Escala fenológica do milho descrita por Ritchie et al. (1993).

Fase	Estágio fenológico	Tempo decorrido
Vegetativa	VE – Emergência	0 dias
	V1 – uma folha aberta	1 semana após emergência
	V4 – quatro folhas abertas	2 semanas após emergência
	V12 – doze folhas abertas	6 semanas após emergência
	V _n – n folhas abertas	-
	VT – pendoamento	8 semanas após emergência
Reprodutiva	R1 – espigamento	9-10 semanas após emergência
	R2 – grãos em bolha	-
	R3 – grão pastoso	12 dias após polinização
	R4 – grão pastoso	24 dias após polinização
	R5 – grão dentado	48 dias após polinização
	R6 – maturidade fisiológica	55 dias após polinização

Atualmente, o milho vem sendo cultivado em quase todos os continentes, em regiões compreendidas entre 58° de latitude norte à 40° de latitude sul, sendo considerado uma das

principais espécies utilizadas no mundo. Os principais países produtores e fornecedores da cultura no mercado mundial são os Estados Unidos, China e Brasil (USDA, 2022).

Segundo Fancelli e Dourado Neto (2000), a interação de tempo e condições climáticas a que uma planta de milho é exposta determina a intensidade com que ela pode expressar seu potencial genético. Nesse contexto, a temperatura influencia diretamente no crescimento das raízes até a absorção de nutrientes e água, além de determinar a duração das fases fenológicas da planta a partir da soma térmica (CARON et al., 2017). De acordo com Bai et al. (2016), a elevação da temperatura pode afetar as taxas de fotossíntese, respiração e enchimento de grãos. Ademais, temperaturas em torno de 25 °C favorecem o florescimento e a maturação dos grãos de milho.

Os valores de temperatura basal da cultura do milho podem variar de acordo com a cultivar utilizada e as condições de cultivo, mas geralmente é utilizado um valor de 10 °C como temperatura basal inferior e 32 °C como temperatura basal superior. Logo, valores fora desse intervalo podem resultar em baixos rendimentos e na diminuição da qualidade dos grãos (ASSIS et al., 2006; ROMANO, 2005).

Sendo uma planta C4, o milho apresenta elevada eficiência em realizar atividade fotossintética e converter fotoassimilados em biomassa (LIU et al., 2012). O processo da fotossíntese está fortemente relacionado com a radiação solar, a qual também é responsável pela abertura e fechamento de estômatos de acordo com a temperatura dentro da planta e do balanço de energia. Logo, reduções na intensidade da radiação solar resulta em diminuição da capacidade da planta em produzir biomassa seca, resultando em baixos incrementos agrícolas (COSTA, 1997).

A produtividade do milho está intimamente ligada à quantidade de radiação fotossinteticamente ativa (RFA) absorvida por suas folhas e à eficiência com que converte energia luminosa em energia química através do processo da fotossíntese. A quantidade de RFA capturada está diretamente relacionada à eficiência com que as folhas da planta de milho a interceptam, sendo que fatores como o índice de área foliar, a arquitetura das folhas, a cultivar, a densidade e época de plantio influenciam essa eficiência de interceptação (DUARTE, 2023; KUNZ et al., 2007; MÜLLER e BERGAMASCHI, 2005).

A temperatura, radiação solar e a precipitação são caracterizadas como as variáveis meteorológicas que mais influenciam no crescimento e desenvolvimento das plantas de milho, sendo as principais responsáveis pelas oscilações dos valores de produtividade observados ao longo das safras, particularmente quando essas variáveis apresentam comportamentos atípicos (GUO et al., 2017; SHIM et al., 2017; YANG et al., 2017).

2.2. Manejo da água no cultivo de milho

O manejo da irrigação desempenha um papel fundamental no regime hídrico das áreas de produção agrícola, tendo como objetivo monitorar o fluxo de água extraído pelas plantas e a sua transferência para o ambiente, a fim de suprir esse volume por meio da irrigação. A adoção de práticas eficientes na gestão da irrigação pode resultar em uma redução de 30 a 70% da água aplicada, além de proporcionar um aumento de 20 a 90% da produção (SACCON, 2018).

A demanda evaporativa da atmosfera e as características da cultura determinam a necessidade hídrica de uma planta ao longo do seu ciclo de cultivo, a qual está diretamente relacionada à interação com a radiação solar, fase fenológica, distribuição espacial do dossel, profundidade do sistema radicular, índice de área foliar e arquitetura da planta (BERGAMASCHI e MATZENAUER, 2014).

O milho demanda uma elevada quantidade de água ao longo do seu ciclo, necessitando de uma lâmina de aproximadamente 600 mm para que possa expressar um adequado desenvolvimento. Essa necessidade hídrica varia em decorrência da época e local de plantio, do tipo de solo, das práticas de manejo adotadas e da cultivar utilizada (ABDRABBO et al., 2016; FANCELLI e DOURADO NETO, 2000).

Para o milho, os períodos de máxima exigência de água concentram-se na fase de emergência, florescimento e enchimento dos grãos, por isso déficits hídricos que ocorrem durante esses períodos são os que provocam maiores reduções de produtividade. De acordo com Fancelli e Dourado Neto (2000), a ocorrência de deficiência de água em uma semana durante o desenvolvimento da inflorescência masculina pode acarretar em uma redução de 50% da produção, enquanto que após a polinização, as perdas podem ser de 25%.

Em resposta ao estresse hídrico, as plantas de milho acionam seus mecanismos de defesa, os quais provocam alterações no seu desempenho fisiológico com o objetivo de permitir que a cultura consiga tolerar essa condição climática. Essas modificações incluem sintomas de maturidade tardia, redução da altura da planta, do índice de área foliar, da produção de biomassa e do rendimento de grãos (TAIZ e ZEIGER, 2009).

Nesse contexto, a disponibilidade e a gestão hídrica ocupam posição de destaque no ciclo de cultivo, uma vez que o excesso ou a falta de água podem causar injúrias e redução da produtividade das plantas. Por isso, a reposição de água ao solo por meio da irrigação, na quantidade e no momento oportuno, apresenta-se como fator decisivo para otimizar a produção de milho. (BERNARDO et al., 2019; KIPKULEI et al., 2022; TAIZ e ZEIGER, 2013).

O manejo da irrigação da cultura do milho consiste em determinar o momento e o volume apropriados para a aplicação de água, para isso diversos critérios podem ser adotados, como o

critério baseado no uso das características físico-hídricas do solo e na estimativa da evapotranspiração da cultura (ETc); baseado em sensores para monitoramento do potencial ou da umidade do solo; e o critério conjunto com o uso de sensores aliado à um método para estimativa da ETo. A escolha da abordagem adotada geralmente depende da sua viabilidade econômica e operacional (ALBUQUERQUE, 2010). Nesse sentido, o manejo da irrigação baseado na necessidade de água da cultura, definida pela demanda evapotranspirométrica, é amplamente utilizado e recomendado.

A evapotranspiração (ET) é um importante componente do ciclo da água e do balanço hídrico na agricultura, o qual consiste no principal parâmetro para determinação do consumo de água das plantas, sendo, assim, um fator crítico para uma gestão adequada das práticas de irrigação. Por isso, estimativa precisa desse processo proporciona a quantidade ideal de água a ser aplicada na cultura, contribuindo no estabelecimento de um sistema sustentável e eficaz (FERREIRA et al., 2019; JING et al., 2019; WANG et al., 2019 WANNIARACHCHI e SARUKKALIGE, 2022).

Tradicionalmente, a ET pode ser medida usando lisímetros (ALLEN et al., 2011; HOLMES, 1984), porém essa técnica é trabalhosa, demorada e custo elevado, além de ser limitada nas escalas temporal e espacial (FALAMARZI et al., 2014). Consequentemente, métodos indiretos e baseados em dados utilizando variáveis meteorológicas foram propostos, uma vez que Allen et al. (1998) enfatizaram que a ET é altamente influenciada pelos fatores climáticos.

2.3. Modelagem na agricultura

A agricultura contemporânea está sendo cada vez mais desafiada pelas mudanças climáticas, aumento nos custos de produção, escassez de terras agricultáveis e recursos hídricos. Esses fatores representam uma ameaça à sustentabilidade ambiental e econômica dos sistemas de abastecimento de alimentos atuais e futuros. Por isso, a busca por inovações científicas e tecnológicas é indispensável para assegurar o suprimento suficiente de alimentos e otimizar os processos agrícolas diante de uma população global em constante crescimento (ANDERSEN et al., 2018; JUNG et al., 2021).

Ao longo do tempo, pesquisadores têm empregado tecnologias de ponta e explorado formas de incorporá-las aos sistemas agrícolas de forma a adquirir melhor compreensão sobre as interações plantas-clima-solo, as quais envolvem complexos processos que podem ser simulados através de modelos. Os modelos dinâmicos de simulação de culturas, bem como as técnicas da ciência de dados, têm se mostrado ferramentas valiosas para a integração de diferentes elementos

dos agrossistemas, as quais permitem a investigação do funcionamento desses componentes dentro do sistema (JUNG et al., 2021).

De acordo com Reynolds (1979) e Monteith (1996), modelos são constituídos por conjuntos de equações que visam representar um determinado processo com base em parâmetros e variáveis de entrada. Tais modelos podem ser divididos em empíricos ou mecanísticos. Os modelos empíricos caracterizam-se por serem baseados em observações e dados experimentais, os quais são utilizados no estabelecimento de relações e padrões estatísticos, sem a obrigação de compreender os mecanismos subjacentes do fenômeno em estudo. A construção desse tipo de modelo é baseada em correlações estatísticas e envolve o ajuste de parâmetros aos dados observados.

Já os modelos mecanísticos são baseados em princípios físicos, químicos ou biológicos que descrevem, a partir de equações matemáticas, os processos envolvidos no fenômeno em questão e têm como objetivo explicar o porquê e como ele ocorre. Por serem baseados em princípios fundamentais, esses tipos de modelos podem apresentar maior robustez, em comparação aos empíricos, quando aplicados em condições distintas das utilizadas para desenvolvê-los (BECHTEL, 2005).

Portanto, quando esse conceito de modelo é aplicado na agricultura, permite-se a representação, de forma simplificada, dos sistemas produtivos através de equações matemáticas responsáveis por simular o crescimento, desenvolvimento e o desempenho produtivo das culturas sem a necessidade da implantação de experimentos agrícolas, o que possibilita ações de planejamento que busquem otimizar a produção, como a escolha da época de plantio e de cultivares adaptados, além do agendamento de irrigação e fertilização (BOUMAN et al., 1996; DUARTE, 2023; MARTINS, 2012).

2.3.1. Modelos de simulação de cultura

As interações planta-clima-solo são caracterizadas por envolverem processos complexos que podem ser simulados através dos modelos mecanísticos de simulação de cultura, os quais possibilitam interagir diferentes características da planta e do solo em diversas condições climáticas, fornecendo informações sobre as respostas da cultura, formas de manejo que reduzam os prejuízos devido às condições adversas de cultivo e os riscos da produção (GODOY, 2007; MARIN et al., 2011).

Os modelos de simulação das culturas são amplamente desenvolvidos e empregados com o propósito de simular o desenvolvimento fisiológico e o desempenho produtivo das plantas a partir de variáveis ambientais, coeficientes genéticos e dados do solo (MONTEITH, 1996). Esses modelos utilizam variáveis climáticas da região, estado nutricional e sanitário das plantas, características de

solo e tratos culturais como dados de entrada para estimar o rendimento da produção e fornecer orientações para o gerenciamento agrícola, como o agendamento de irrigação e adubação (MACCARTHY et al., 2017; MALIK et al., 2019; FU et al., 2020; XIAO et al., 2020).

Dentre os modelos mais citados na literatura, pode-se destacar o modelo DSSAT (*Decision Support System for Agrotechnology Transfer*, DSSAT version 4.6, JONES et al., 2003; HOOGENBOOM et al., 2014), o qual é um dos principais sistemas que empregam modelos de simulação de crescimento e desenvolvimento de culturas ao longo do tempo, sendo composto por ferramentas de suporte que auxiliam na manipulação de dados de solo, clima, condições experimentais e informações sobre o genótipo.

O modelo DSSAT possui abordagem mecanística e baseia-se nas interações dos componentes inseridos em um sistema agrícola, assim, pode-se simular interações entre as culturas e o ambiente de diversas formas e realizar análises do desempenho do cultivo para diferentes épocas de plantio, tipos de solo, população de plantas, tipos de solo manejos de irrigação e adubação com o intuito de obter práticas que envolvam menor risco à produção (JAMES e CURTFORTH, 1996; JONES et al., 1996; MAVROMATIS et al., 2002).

Esse modelo vem sendo amplamente utilizado para atender a uma variedade de aplicações, como a otimização de práticas agrônômicas (e.g., manejo racional da irrigação) para intensificar a produção (Kadiyala et al., 2015; Mubeen et al., 2016; Oliveira et al., 2018), avaliação do desempenho de várias cultivares em diferentes condições ambientais (Chisanga et al., 2021; Feleke et al., 2021), análise dos impactos das mudanças climáticas na produção agrícola (Jiang et al., 2021; Oliveira et al., 2018) e avaliação de estratégias para lidar com condições limitadas de solo e clima (Malik e Dechmi, 2019).

2.3.1.1. Modelo DSSAT/CERES-MAIZE

O DSSAT inclui o modelo CERES-MAIZE, componente de uma série de modelos de simulação dos processos fisiológicos denominada CERES (*Crop Environment Resource Synthesis*), específico para simular o rendimento da cultura de milho sob condições ambientais distintas, o qual, de acordo com Jones e Kiniry (1986), é um dos mais detalhados na predição dos estádios fenológicos e número de folhas do milho. Para isso, são considerados a sensibilidade ao fotoperíodo e os efeitos da elevada temperatura no número final de folhas.

No modelo CERES-MAIZE são consideradas as seguintes fases fenológicas para a cultura do milho: germinação, emergência, fim da fase juvenil, iniciação da floração masculina, início do enchimento de grãos e maturidade fisiológica (HOOGENBOOM et al., 2010). O desenvolvimento das

plantas é controlado pela soma térmica ou graus-dia ($^{\circ}\text{C}$ dia), dado pelos graus-dia acumulados desde a emergência até o estágio juvenil (temperatura base de 8°C) (P1), pelo coeficiente de sensibilidade ao fotoperíodo (P2), pelos graus-dia acumulados do florescimento até o ponto de maturidade fisiológica (P5) e graus-dia necessários para que ocorra a emergência de uma folha (PHINT). Já a produtividade é controlada pelos coeficientes G2 e G3, os quais representam o número potencial de grãos por planta e a taxa potencial de crescimento do grão, respectivamente (HOOGENBOOM et al., 1994, 2010; JONES et al., 2003).

Para realizar a simulação do crescimento e desenvolvimento do milho, o modelo CERES-MAIZE considera o processo do balanço de água no solo, a interceptação da radiação pelo dossel, a produção de matéria seca e a partição da biomassa. Além disso, para que a simulação ocorra são necessários no mínimo dados de entrada de clima, do solo e práticas de manejo, tais como época de semeadura, cultivar, aplicação de irrigação e fertilizantes, densidade de plantas e espaçamento entre linhas (KINIRY et al., 1997).

O modelo tem sido testado em diferentes condições, obtendo bons resultados de estimativa do rendimento agrícola. Demonstrando a utilidade da ferramenta, Borges Alves et al. (2012) testaram o DSSAT/CERES-MAIZE e obtiveram um desempenho satisfatório na determinação das janelas de semeadura do milho, assim como a produtividade média esperada.

2.3.2. Algoritmos de *machine learning*

Nos últimos anos tem sido observado um crescimento significativo no uso de técnicas provenientes da área de inteligência artificial, especialmente do aprendizado de máquina, no contexto da análise de dados e monitoramento da produção com o objetivo de auxiliar os produtores na tomada de decisões (SARKER et al., 2021).

O aprendizado de máquina consiste em um subcampo da inteligência artificial, o qual tem como foco o desenvolvimento de algoritmos e técnicas baseando-se na capacidade de aprendizado e identificação de padrões a partir dos dados disponíveis em um processo comumente denominado de treino, onde os algoritmos são expostos à repetição de testes fundamentados em definições, exemplos e feedbacks, utilizando um conjunto de métodos lógicos e probabilísticos, que permitem o ajuste de seus parâmetros, bem como uma melhora no desempenho ao longo do tempo através da sua experiência (KERSTING, 2018; van KLOMPENBURG, KASSAHUN e CATAL, 2020).

Os modelos do aprendizado de máquina podem ser divididos em modelos supervisionados e não supervisionados. No primeiro, o treinamento é realizado com base em um conjunto de dados rotulados, os quais são, posteriormente, validados para avaliar a acurácia do algoritmo na predição

da variável resposta. Enquanto que os modelos de aprendizagem não supervisionado são caracterizados por não dependerem de rótulos previamente definidos e por isso não têm como objetivo a predição da variável resposta, mas sim a exploração de dados desconhecidos capazes de identificar importantes padrões e ajustar os dados em agrupamentos (VERMA, SINGH e DIXIT, 2019).

Na agricultura, o aprendizado de máquina tem uma ampla gama de aplicações, incluindo o reconhecimento de doenças e deficiências nutricionais (Abade, Ferreira e Vidal, 2021; Sujatha et al., 2021; Yashodha e Shalini, 2020), identificação de plantas daninhas (Hasan et al., 2021; Su, 2020), predição da produtividade (Chlingaryan, Sukkarieh e Whelan, 2018; Souza, 2021; Van Klompenburg, Kassahun e Catal, 2020), programação de colheita (Thilakarathne et al., 2022) e manejo da irrigação (Chen et al., 2020; Sun e Scanlon, 2019; Virnodkar et al., 2020).

Algoritmos de diversas metodologias de predição já foram propostos como possíveis alternativas para a estimativa da ETo e do rendimento agrícola em várias condições climáticas ao redor do mundo, como demonstrado em Chen et al., 2020; Chia et al., 2020; Granata, 2019; Mohammadi e Mehdizadeh, 2020; Raza et al., 2020; e Wang et al., 2019.

Raza et al. (2020) utilizaram cinco algoritmos de *machine learning*, entre eles *support vector machine* (SVM) e *multilayer perceptron* (MLP), para estimar a ETo a partir dos dados meteorológicos de seis estações meteorológicas do Paquistão. Esses modelos, então, foram aplicados em outras estações que não possuem dados de treinamento disponíveis localizadas na China, Nova Zelândia e EUA para validar as suas eficientes performances. Os autores concluíram que os cinco algoritmos foram capazes de estimar a ETo de forma satisfatória em diferentes zonas climáticas do Paquistão, bem como em outras partes do mundo, com erro médio de predição de $0.016 \text{ mm dia}^{-1}$, sendo que o SVM apresentou a melhor performance para as condições do estudo.

Algoritmos baseados nas chamadas *support vector machine* (SVM), os *Support vector regression* (SVR), aplicação do SMV à análise de regressão, são capazes de aprender e modelar relações não-lineares entre os dados de entrada e saída em uma dimensão superior, minimizando assim o erro de treinamento observado e o erro de distribuição suficientemente para alcançar a eficiência de regressão generalizada. Esse algoritmo utiliza um conjunto de funções matemáticas que são definidas como o kernel, as quais são responsáveis por pegar os dados como entrada e transformá-los no formato necessário, ou seja, transformar dados linearmente inseparáveis em dados linearmente separáveis (YU et al., 2018).

Gradient boostings (GB) são formas de algoritmos de aprendizado de máquina baseados em modelos aditivos, caracterizados como uma metodologia *ensemble* (*ensemble method*), em que vários modelos preditivos “fracos” são criados e combinados com o objetivo de se obter um melhor desempenho como um todo. A ideia central dos métodos *boosting* é treinar modelos

sequencialmente, onde cada previsor adicionado tenta corrigir ou minimizar os erros de previsão realizados pelo algoritmo antecessor (FRIEDMAN et al., 2000; 2001; NATEKIN e KNOLL, 2013).

A abordagem GB trata o processo de treinamento como um modelo aditivo, ou seja, tenta minimizar as perdas pela implementação de mais estimadores. Também permite que funções de perda arbitrárias sejam utilizadas, onde o objetivo é minimizar a função de perda dos previsores fracos, obtendo médias de ponderação e melhorando a capacidade da técnica. O HGBR, GB baseado em histograma, é uma abordagem do aprendizado de máquina introduzida recentemente por Chen e Guestrin (2016), caracterizada por agrupar variáveis contínuas em compartimentos (bins) discretos e utilizar esses compartimentos para construir histogramas das variáveis durante o treinamento e determinar as falhas dos modelos.

Já o *Multi-layer perceptron regressor* (MLPR) é um algoritmo baseado em redes neurais artificiais, composto por camadas de entrada e saída e uma ou mais camadas ocultas com vários neurônios empilhados juntos. Os neurônios de uma camada estão ligados aos neurônios de outra pela conexão chamada “pesos”, os quais são responsáveis por transportar os resultados de uma camada para outra. Para processar a entrada de dados entre neurônios de cada camada é usada a função de ativação (HASSANPOUR KASHANI et al., 2008).

Entre os algoritmos mais utilizados pode-se citar o *Random Forest* (RF), algoritmo de aprendizado supervisionado baseado na utilização de diversos algoritmos menores (árvores de decisão), o qual usa um conjunto de métodos (*ensemble methods*) para resolver problemas de regressão e classificação. A técnica de amostragem utilizada por esse modelo é conhecida como *bootstrapping*, em que subconjuntos de observações são criados aleatoriamente a partir do conjunto de dados original (BREIMAN, 2001; GARCÍA-PEÑALVO et al., 2018; SATTARI et al., 2020).

O algoritmo de aprendizado de máquina supervisionado *Extra-Trees* (*Extremely Randomized Trees*) pode ser caracterizado como uma extensão altamente aleatório do modelo *Random Forest*, proposto por Geurts et al. (2006), o qual funciona gerando aleatoriamente diversas árvores de decisão e, então, combina os resultados de cada árvore para encontrar a resposta final. Seu principal diferencial é que esse processo é caracterizado por ser extremamente aleatório, o que contribui para a construção de modelos mais generalizáveis. Segundo Geurts et al. (2006), as diferenças em relação ao RF resultam em modelos menos suscetíveis ao *overfitting* e desempenho superior.

Referências

- ABADE, A.; FERREIRA, P.; VIDAL, F. Plant Diseases recognition on images using convolutional neural networks: A systematic review. **Computers and Electronics in Agriculture**, v. 185, p. 1–47, 2021.
- ABDRABBO, M. A. A.; SALEH, S. M.; FARAG, A. A. Water requirements for maize under climate change. **Journal of Applied Sciences Research**, v. 12, p. 19–28, 2016.
- ALBUQUERQUE, P. E. P. **Manejo de irrigação na cultura do milho**. Embrapa Milho e Sorgo, Sete Lagoas. Sistema de Produção 1, Versão Eletrônica 6ª edição, 2010.
- ALLEN, R. G.; PEREIRA, L. S.; HOWELL, T. A.; JENSEN, M. E. Evapotranspiration information reporting: I. Factors governing measurement accuracy. **Agricultural Water Management**, v. 98, n. 6, p. 899–920, 2011.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. **Crop evapotranspiration: guidelines for computing crop water requirements**. Rome: FAO, 300 p., 1998 (FAO irrigation and drainage paper no. 56).
- ANDERSEN, M. A.; ALSTON, J. M.; PARDEY, P. G.; SMITH, A. A century of U.S. productivity growth: a surge then a slowdown. **American Journal of Agricultural Economics**, v. 100, n. 4, p. 1072-1090, 2018.
- ASSIS, J. P.; DOURADO NETO, D.; NASS, L. L.; MANFRON, P. A.; BONNECARRERE, R. A. G.; MARTIN, T. N. Simulação estocástica de atributos do clima e da produtividade potencial de milho utilizando-se distribuição triangular. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, p. 539–543, 2006.
- BAI, H.; TAO, F.; XIAO, D.; LIU, H.; ZHANG, H. Attribution of yield change for rice-wheat rotation system in China to climate change, cultivars and agronomic management in the past three decades. **Climatic Change** 135, p. 539–553, 2016.
- BARROS, J. F. C.; CALADO, J. G. A Cultura do Milho. **Évora: Universidade de Évora**, Escola de ciências e tecnologia, Departamento de Fitotecnia. 52 p., 2014.
- BECHTEL, W. Explanation: A mechanist alternative. **Studies in History and Philosophy of Biological and Biomedical Sciences**, v. 36, n. 2, p. 421–441, 2005.
- BERGAMASCHI, H.; MATZENAUER, R. O milho e o clima. **Emater/RS-Ascar**, Porto Alegre, 84 p., 2014.
- BERNARDO, S.; MANTOVANI, E. C.; SILVA, D. D.; SOARES, A. A. **Manual de Irrigação**. 9. ed. Viçosa: UFV, 545 p., 2019.
- BORGES ALVES, M. E.; PAIXÃO, J. S.; ANDRADE, C. L. T.; AMARAL, T. A.; STEIDLE NETO, A. J.; SILVA, D. F.; FERREIRA RODRIGUES, C. C. Épocas de semeadura de milho em Minas Gerais: utilização do modelo DSSAT para análise do risco climático. In: CONGRESSO NACIONAL DE MILHO E SORGO, 29, 2012, Águas de Lindóia. **Anais... Águas de Lindóia: ABMS**, p. 3591-3597, 2012.

- BOUMAN, B. A. A.; van KEULEN, H.; van LAAR, H. H.; RABBINGE, R. The “School of de Wit” crop growth simulation models: a pedigree and historical overview. **Agricultural Systems**, v. 52, p. 171–198, 1996.
- BREIMAN, L. Random forests. **Machine Learning**, v. 45, p. 5–32, 2001.
- CARON, B. O.; OLIVEIRA, D. M.; ELLI, E. F.; ELOY, E.; SCHWERZ, F.; SOUZA, V. Q. Elementos meteorológicos sobre características morfológicas e produtivas do milho em diferentes épocas de semeadura. **Científica**, Jaboticabal, v. 45, p. 105–114, 2017.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In **Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, p. 785-794, 2016.
- CHEN, Z.; ZHU, Z.; JIANG, H.; SUN, S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. **Journal of Hydrology**, v. 591, p. 125–286, 2020.
- CHIA, M. Y.; HUANG, Y. F.; KOO, C. H. Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. **Computers and Electronics in Agriculture**, v. 175, p. 1–12, 2020.
- CHLINGARYAN, A.; SUKKARIEH, S.; WHELAN, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. **Computers and Electronics in Agriculture**, v. 151, p. 61–69, 2018.
- CHISANGA, C. B.; PHIRI, E.; CHINENE, V. R. N. Evaluating APSIM-and-DSSAT-CERES-Maize models under rainfed conditions using zambian rainfed maize cultivars. **Nitrogen**, v. 2(4), p. 392–414, 2021.
- COSTA, L. C. Modelagem e simulação em agrometeorologia. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA 10, Piracicaba, 1997. **Anais...** Piracicaba: SBA, p. 3–6, 1997.
- DUARTE, Y. C. N. **Impacto das mudanças climáticas no Brasil e sua influência no cultivo e na produtividade do milho: uma abordagem multi-modelos**. Tese. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 250 p., 2023.
- FALAMARZI, Y.; PALIZDAN, N.; HUANG, Y. F.; LEE, T. S. 2014. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). **Agricultural Water Management**, v. 140, p. 26–36, 2014.
- FANCELLI, A. L.; DOURADO NETO, D. Produção de milho. **Livraria e Editora Agropecuária**, Guaíba, 360 p., 2000.
- FELEKE, H. G.; SAVAGE, M. J.; TESFAYE, K. Calibration and validation of APSIM–Maize, DSSAT CERES–Maize and AquaCrop models for Ethiopian tropical environments. **South African Journal of Plant and Soil**, v. 38(1), p. 36–51, 2021.

- FERREIRA, L. B.; da CUNHA, F. F.; de OLIVEIRA, R. A.; FERNANDES FILHO, E. I. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – A new approach. **Journal of Hydrology**, v. 572, p. 556–570, 2019.
- FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). **Annals of Statistics**, v. 28, p. 337–407, 2000.
- FRIEDMAN, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 1189–1232.
- FU, C.; WANG, J.; GONG, S.; ZHANG, Y.; WANG, C.; MO, Y. Optimization of irrigation and fertilization of drip-irrigated corn in the chernozem area of north-east China based on the CERES-Maize model. **Irrigation and Drainage**, v. 69, p. 714–731, 2020.
- GARCÍA-PEÑALVO, F.; CRUZ-BENITO, J.; MARTÍN-GONZÁLEZ, M.; VÁZQUEZ-INGELMO, A.; SÁNCHEZ-PRIETO, J. C.; THERÓN, R. Proposing a machine learning approach to analyze and predict employment and its factors. **International Journal of Interactive Multimedia and Artificial Intelligence**, v. 5, p. 39–45, 2018.
- GEURTS, P.; ERNST, D.; WEHENKEL, L. Extremely randomized trees. **Machine Learning**, v. 63, p. 3–42, 2006.
- GODOY, A. P. **Modelagem de processos de acumulação de biomassa e de açúcar da cana-de-açúcar via sistemas nebulosos**. Dissertação. Faculdade de Engenharia Elétrica e Computação, Universidade Estadual de Campinas, Campinas, 254 p., 2007.
- GRANATA, F.; GARGANO, R.; de MARINIS, G. Artificial intelligence based approaches to evaluate actual evapotranspiration in wetlands. **Science of the Total Environment**, v. 703, p. 1–15, 2020.
- GUO, E.; ZHANG, J.; WANG, Y.; ALU, S.; WANG, R.; LI, D.; HA, S. Assessing non-linear variation of temperature and precipitation for different growth periods of maize and their impacts on phenology in the Midwest of Jim province, China. **Theoretical and Applied Climatology**, v. 126, p. 1–15, 2017.
- HASAN, A. S. M. M.; SOHEL, F.; DIEPEVEEN, D.; LAGA, H.; JONES, M. G. K. A survey of deep learning techniques for weed detection from images. **Computers and Electronics in Agriculture**, v. 184, 106067, 2021.
- HASSANPOUR KASHANI, M., MONTASERI, M.; LOTFOLLAHI YAGHIN, M. A. Flood estimation at ungauged sites using a new hybrid model. **Journal of Applied Sciences**, v. 8, p. 1744–1749, 2008.
- HOLMES, J. W. Measuring evapotranspiration by hydrological methods. **Agricultural Water Management**, v. 8, n. 1–3, p. 29–40, 1984.

- HOOGENBOOM, G.; JONES, J. W.; WILKENS, P. W.; BATCHELOR, W. D.; BOWEN, W. T.; HUNT, L. A.; PICKERING, N. B.; SINGH, U.; GODWING, D. C.; BAER, B.; BOOTE, K. J.; RITCHIE, J. T.; WHITE, J. W. Crop models. In: TSUJI, G. Y.; UEHARA, G.; BALAS, S. (Ed.). **DSSAT v3, Decision Support System for Agrotechnology Transfer**. Honolulu: University of Hawaii, v. 2, cap. 2, p. 95–244, 1994.
- HOOGENBOOM, G.; JONES, J. W.; PORTER, C. H.; WILKENS, P. W.; BOOTE, K. J.; HUNT, L. A.; TSUJI, G. Y. **Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.5**. University of Hawaii, Honolulu, Hawaii, 2010.
- HOOGENBOOM, G.; JONES, J. W.; WILKENS, P. W.; PORTE, C. H.; BOOTE, K. J.; HUNT, L. A.; SINGH, U.; LIZASO, J. L.; WHITE, J. W.; URYASEV, O.; ROYCE, F. S.; OGOSHI, R.; GIJSMAN, A. J.; TSUJI, G. Y. **Decision Support System for Agrotechnology Transfer (DSSAT): version 4.6**. DSSAT Foundation, Prosser, Washington, 2014.
- JAMES, Y. W.; CURTFORTH, H. W. Crop growth models for decision support systems. **Canadian Journal of Plant Science**, v. 76, p. 9–19, 1996.
- JIANG, R.; HE, W.; HE, L.; YANG, J. Y.; QIAN, B.; ZHOU, W.; HE, P. Modelling adaptation strategies to reduce adverse impacts of climate change on maize cropping system in Northeast China. **Scientific Reports**, v. 11, 810, 2021.
- JING, W.; YASEEN, Z. M.; SHAHID, S.; SAGGI, M. K.; TAO, H.; KISI, O.; SALIH, S. Q.; AL-ANSARI, N.; CHAU, K. W. Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions. **Engineering Applications of Computational Fluid Mechanics**, v. 13, p. 811–823, 2019.
- JONES, C. A.; KINIRY, J. R. **Ceres-Maize: a simulation model of maize growth and development**. College Station: Texas A & M University Press, 194 p., 1986.
- JONES, J.; TSUJI, G.; HOOGENBOOM, G.; HUNT, L.; THORNTON, P.; WILKENS, P.; IMAMURA, D.; BOWEN, W.; SINGH, U. Decision support system for agrotechnology transfer DSSAT. In: TSUJI, G.; HOOGENBOOM, G.; THORNTON, P. (Ed.). **International benchmark sites network for agrotechnology transfer: a system approach to research and decision makins**. Dordrecht: Kluwer Academic Press, v. 3, 1996.
- JONES, J. W.; HOOGENBOOM, G.; PORTER, C. H.; BOOTE, K. J.; BATCHELOR, W. D.; HUNT, L. A.; WILKENS, P. W.; SINGH, U.; GIJSMAN, A. J.; RITCHIE, J. T. The DSSAT cropping system model. **European Journal of Agronomy**, v. 18, p. 235–265, 2003.
- JUNG, J.; MAEDA, M.; Anjin CHANG, A.; BHANDARI, M.; ASHAPURE, A.; LANDIVAR-BOWLES, J. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. **Current Opinion in Biotechnology**, v. 70, p. 15–22, 2021.

- KADIYALA, M. D. M.; JONES, J.W.; MYLAVARAPU, R. S.; LI, Y. C.; REDDY, M. D. Identifying irrigation and nitrogen best management practices for aerobic rice–maize cropping system for semi-arid tropics using CERES-rice and maize models. **Agricultural Water Management**, v. 149, p. 23–32, 2015.
- KERSTING, K. Machine learning and artificial intelligence: Two fellow travelers on the quest for intelligent behavior in machines. **Sec. Machine Learning and Artificial Intelligence**, v.1, 2018.
- KINIRY, J. R.; WILLIAMS, J. R.; VANDERLIP, R. L.; ATWOOD, J. D.; REICOSKY, D. C.; MULLIKEN, J.; COX, W. J.; MASCANI JR., H. J.; HOLLINGER, S. E.; WIEBOLD, W. J. Evaluation of two maize models for nine U. S. locations. **Agronomy Journal**, v. 89, p. 421–426, 1997.
- KIPKULEI, H. K.; BELLINGRATH-KIMURA, S.D.; LANA, M. et al. Modelling cropland expansion and its drivers in Trans Nzoia County, Kenya. **Modeling Earth Systems and Environment**, v. 8, p. 5761–5778, 2022.
- KUNZ, J. H.; BERGONCI, J. L.; BERGAMASCHI, H.; DALMAGO, G. A.; HECLER, B. M. M.; COMIRAN, F. Uso da radiação solar pelo milho sob diferentes preparos de solo, espaçamento e disponibilidade hídrica, 2007.
- LIU, Z. A.; YANG, X.; HUBBARD, K. G.; LIN, X. Maize potential yields and yield gaps in the changing climate of northeast China. **Global Change Biology**, v. 18, p. 3441–3454, 2012.
- MACCARTHY, D. S.; ADIKU, S. G.; FREDUAH, B. S.; GBEFO, F. Using CERES-Maize and ENSO as decision support tools to evaluate climate-sensitive farm management practices for maize production in the northern regions of Ghana. **Frontiers in Plant Science** 8:31, 2017.
- MALIK, W.; DECHMI, F. DSSAT modelling for best irrigation management practices assessment under Mediterranean conditions. **Agricultural Water Management**, v. 216, p. 27–43, 2019.
- MALIK, W.; ISLA, R.; DECHMI, F. DSSAT-CERES-maize modelling to improve irrigation and nitrogen management practices under Mediterranean conditions. **Agricultural Water Management**, v. 213, p. 298–308, 2019.
- MARIN, F. R.; JONES, J. W.; ROYCE, F.; SUGUITANI, C.; DONZELI, J. L.; PALLONE FILHO, W. J.; NASSIF, D. S. P. Parameterization and evaluation of predictions of DSSAT/CANEGRO for sugarcane Brazilian production systems. **Agronomy Journal**, v. 103, p. 100–110, 2011.
- MARTINS, M. A. **Estimativa de produtividade das culturas do milho e do sorgo a partir de modelos agrometeorológicos em algumas localidades da região nordeste do Brasil**. Tese. Instituto Nacional de Pesquisas Espaciais – INPE, São José dos Campos, 2012.
- MAVROMATIS, T.; BOOTE, K. J.; JONES, J. W.; WILKERSON, G. G.; HOOGENBOOM, G. Repeatability of model genetic coefficients derived from soybean performance trials across different states. **Crop Science**, v. 42, p. 76–89, 2002.

- MOHAMMADI, B.; MEHDIZADEH, S. Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. **Agricultural Water Management**, v. 237, p. 1–13, 2020.
- MONTEITH, J. The quest balance in crop modeling. **Agronomy Journal**, v. 88, p. 695–697, 1996.
- MUBEEN, M.; AHMAD, A.; WAJID, A.; KHALIQ, T.; HAMMAD, H. M.; SULTANA, S. R.; AHMAD, S.; FAHAD, S.; NASIM, W. Application of CSM-CERES-Maize model in optimizing irrigated conditions. **Outlook on Agriculture**, v. 45(3), p. 173–184, 2016.
- MÜLLER, A. G.; BERGAMASCHI, H. Eficiências de interceptação, absorção e de uso da radiação fotossinteticamente ativa pelo milho (*Zea mays* L.) em diferentes disponibilidades hídricas. **Revista Brasileira de Agrometeorologia**, v. 13, p. 27–33, 2005.
- NATEKIN, A.; KNOLL, A. Gradient boosting machines, a tutorial. **Frontiers in Neurorobotics**, v. 7, 21, 2013.
- OLIVEIRA, L. A.; MIRANDA, J. H.; COOKE, R. A. C. Water management for sugarcane and corn under future climate scenarios in Brazil. **Agricultural Water Management**, v. 201, p. 199–206, 2018.
- RAZA, A.; SHOAIB, M.; FAIZ, M. A.; BAIG, F.; KHAN, M. M.; ULLAH, M. K.; ZUBAIR, M. Comparative assessment of reference evapotranspiration estimation using conventional method and machine learning algorithms in four climatic regions. **Pure and Applied Geophysics**, v. 177, n. 9, p. 4479–4508, 2020.
- REYNOLDS, J. F. Some misconceptions of mathematical modeling. **Plant Physiology**, v. 10, p. 41–44, 1979.
- RITCHIE, S. W.; HANWAY, J. J.; BENSON, G. O. How a corn plant develops. **Special Bulletin**, n. 48, 1993.
- ROMANO, M. R. **Desempenho fisiológico da cultura de milho com plantas de arquitetura contrastante: parâmetros para modelos de crescimento**. Tese. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 100 p., 2005.
- SACCON, P. Water for agriculture, irrigation management. **Applied Soil Ecology**, v. 123, p. 793–796, 2018.
- SARKER, I. H.; FURHAD, M. H.; NOWROZY, R. AI-driven cybersecurity: an overview, security intelligence modeling and research directions. **SN Computer Science**, v. 2, 173, 2021.
- SATTARI, M. T.; APAYDIN, H.; SHAMSHIRBAND, S. Performance evaluation of deep learning-based gated recurrent units (GRUs) and tree-based models for estimating ETo by using limited meteorological variables. **Mathematics**, v. 8, 972, 2020.
- SHIM, D.; LEE, K. J.; LEE, B. W. Response of phenology – and yield-related traits of maize to elevated temperature in a temperate region. **The Crop Journal**, v. 5, p. 305–316, 2017.

- SOUZA, P. V. D. **Rede neural artificial para predição da produtividade da cultura do milho**. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) - Universidade Tecnológica Federal do Paraná, Santa Helena, 2021.
- SU, W. -H. Advanced machine learning in point spectroscopy, RGB- and hyperspectral-imaging for automatic discriminations of crops and weeds: A review. **Smart Cities**, v. 3, p. 767–792, 2020.
- SUN, A.Y.; SCANLON, B.R. How can Big Data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. **Environmental Research Letters**, v. 14, n. 7, 073001, 2019.
- SUJATHA, R.; CHATTERJEE, J. M.; JHANJHI, N. Z.; BROHI, S. N. Performance of deep learning vs machine learning in plant leaf disease detection. **Microprocessors and Microsystems**, v. 80, 103615, 2021.
- TAIZ, L.; ZEIGER, E. *Fisiologia Vegetal*, 4ª Edição. **Artmed**, Porto Alegre, 848 p., 2009.
- TAIZ, L.; ZEIGER, E. *Fisiologia Vegetal*, 5ª Edição. **Artmed**, Porto Alegre, 954 p., 2013.
- THILAKARATHNE, N. N.; BAKAR, M.; ABAS, P. E.; YASSIN, H. A cloud enabled crop recommendation platform for machine learning-driven precision farming. **Sensors**, v. 22, 6299, 2022.
- USDA - United States Department of Agriculture, 2022.
- van KLOMPENBURG, T.; KASSAHUN, A.; CATAL, C. Crop yield prediction using machine learning: A systematic literature review. **Computers Electronics Agriculture**, v. 177, 105709, 2020.
- VERMA, K. K.; SINGH, B. M.; DIXIT, A. A review of supervised and unsupervised machine learning techniques for suspicious behavior recognition in intelligent surveillance system. **International Journal of Information Technology**, v. 14, p. 397–410, 2019.
- VIRNODKAR, S. S.; PACHGHARE, V. K.; PATIL, V. C.; JHA, S. K. Remote sensing and machine learning for crop water stress determination in various crops: A critical review. **Precision Agriculture**, v. 21, n. 5, p. 1121–1155, 2020.
- WANG, S.; LIAN, J.; PENG, Y.; HU, B.; CHEN, H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. **Agricultural Water Management**, v. 221, p. 220–230, 2019.
- WANNIARACHCHI, S.; SARUKKALIGE, R. A Review on Evapotranspiration Estimation in Agricultural Water Management: Past, Present, and Future. **Hydrology** 9, 123, 2022.
- XIAO, D.; LI LIU, D.; WANG, B.; FENG, P.; WATERS, C. Designing high-yielding maize ideotypes to adapt changing climate in the North China Plain. **Agricultural Systems**, v. 181, 102805, 2020.
- YANG, H.; GRASSINI, P.; CASSMAN, K. G.; AIKEN, R. M.; COYNE, P. I. Improvements to the Hybrid-Maize model for simulating maize yields in harsh rainfed environments. **Field Crops Research**, v. 204, p. 180–190, 2017.

- YASHODHA, G.; SHALINI, D. An integrated approach for predicting and broadcasting tea leaf disease at early stage using IoT with machine learning—A review. **Materials Today: Proceedings**, v. 37, p. 484–488, 2020.
- YU, X.; ZHANG, X.; QIN, H. A data-driven model based on Fourier transform and support vector regression for monthly reservoir inflow forecasting. **Journal of Hydro-environment Research**, v. 18, p. 12–24, 2018.

3. USO DE MACHINE LEARNING NA ESTIMATIVA DA EVAPOTRANSPIRAÇÃO DE REFERÊNCIA EM PIRACICABA, SP

Resumo

A evapotranspiração é composta basicamente por dois processos: evaporação da superfície do solo e transpiração das plantas. A sua estimativa com adequada precisão é fundamental para a compreensão do balanço hídrico e dos processos hidrológicos, sendo um fator relevante para se alcançar uma gestão eficiente e sustentável do uso da água. A limitação da aplicação do modelo padrão (FAO-56 Penman-Monteith), recomendado para estimativa da evapotranspiração de referência (ET_o), ocorre quando é necessário a aplicação de dados meteorológicos detalhados, os quais apresentam uma certa dificuldade em estarem disponíveis, fato que é observado principalmente em países emergentes, como o Brasil. Diante disso, pesquisas vêm demonstrando que a utilização de técnicas de ciência de dados, tais como *machine learning* (aprendizado de máquina) pode se tornar em uma alternativa ao modelo padrão FAO 56 ao ser aplicado diante de uma redução de variáveis climáticas. Nesse contexto, o objetivo da presente pesquisa foi avaliar algoritmos de aprendizado supervisionado visando a estimativa da ET_o para a região de Piracicaba, SP, com diferentes combinações de dados. Para tal, foram obtidos dados climáticos do posto meteorológico da ESALQ/USP, em uma série de 21 anos de dados diários e assim calculados os valores da ET_o pelo método FAO-56 e analisadas as correlações desses dados com a variável-alvo (ET_o). Após a escolha das variáveis de entrada, três algoritmos de aprendizado de máquina – *support vector regression* (SVR), *histogram-based gradient boosting regression tree* (HGBR) e *multilayer perceptron regressor* (MLPR) foram treinados e testados para a predição da ET_o. Pelos resultados obtidos, pode-se notar que as variáveis meteorológicas radiações solares global e líquida, temperatura média e umidade relativa média apresentaram elevadas correlações com a ET_o para a região de Piracicaba. Em relação aos algoritmos, os altos e baixos valores de R² e erro quadrado médio (RMSE), respectivamente, indicaram que todos os modelos aplicados apresentaram desempenho satisfatório na predição da ET_o, sendo que o HGBR forneceu estimativas mais precisas em comparação ao SVR e MLPR. Além disso, foi observado que o tamanho da série histórica não influenciou no desempenho dos modelos. Dessa forma, pode-se concluir que os algoritmos de aprendizado de máquina adotados podem se tornar ferramentas com aplicabilidade satisfatória na predição da ET_o, apresentando-se como uma alternativa para a quantificação desse processo, diante de uma quantidade reduzida de variáveis climáticas.

Palavras-chave: Evapotranspiração, Algoritmos, *Support vector*, *Histogram-based gradient boosting*, *Multilayer perceptron*

Abstract

Evapotranspiration is basically composed of two processes: soil surface evaporation and plant transpiration. Its accurate estimation is fundamental for understanding the water balance and hydrological processes, being a relevant factor for achieving efficient and sustainable water use management. The limitation of the standard model (FAO-56 Penman-Monteith), recommended for estimating reference evapotranspiration (ET_o), occurs when detailed meteorological data are required, which are often not readily available, particularly in emerging countries like Brazil. Therefore, research has shown that the use of data science techniques such as machine learning can become an alternative to the standard FAO-56 model when faced with a reduction in climatic variables. In this context, the objective of this research was to evaluate supervised machine learning algorithms in estimating ET_o in the Piracicaba, SP, region using different combinations of

input data. Climatic daily data from the ESALQ/USP meteorological station for 21 years were obtained, thus, ETo values were calculated using the FAO-56 method and it was analyzed the correlations of these data with the target variable (ETo). After selecting the input variables, three machine learning algorithms – support vector regression (SVR), histogram-based gradient boosting regression (HGBR), and multilayer perceptron regressor (MLPR) – were trained and tested for ETo prediction. Results showed that global and net solar radiation, mean temperature and mean relative humidity had high correlations with ETo for the Piracicaba region. Regarding the algorithms, the high and low valued of R^2 and mean squared error (RMSE), respectively, indicated that all models performed satisfactorily in predicting ETo, with HGBR providing more precise estimates than SVR and MLPR. Furthermore, it was observed that the size of historical time series had no influence on models performance. Therefore, it can be concluded that the adopted machine learning algorithms can become tools with satisfactory applicability in the prediction of ETo, presenting itself as an alternative for the quantification of this process when in a situation with reduced availability of climatic variables.

Keywords: Evapotranspiration, Algorithms, Support vector, Histogram-based gradient boosting, Multilayer perceptron

3.1. Introdução

A evapotranspiração (ET) é um importante componente do ciclo da água e do balanço hídrico na agricultura, já que mais de 60% da precipitação global é devolvida à atmosfera por meio desse processo, o qual representa o efeito combinado da evaporação da água da superfície do solo e da transpiração das plantas via liberação de moléculas de água pelo estômato (JING et al., 2019; WANG et al., 2019).

Tanto a chuva quanto a ET caracterizam o clima de uma região e são utilizadas como ferramentas de apoio em tomadas de decisão sobre o gerenciamento da água na agricultura. Por representar o principal uso consuntivo da água agrícola, a ET quantifica a demanda hídrica das plantas, sendo, assim, um fator crítico para uma gestão adequada das práticas de irrigação. Portanto, a estimativa precisa do consumo hídrico em áreas agrícolas torna-se fundamental no estabelecimento de um sistema sustentável e eficaz, mitigando os impactos da escassez e do uso excessivo da água (FERREIRA et al., 2019; WANNIARACHCHI e SARUKKALIGE, 2022).

Krishna (2018) destacou que a estimativa precisa da ET é essencial, pois o entendimento e a quantificação dos processos que regem a ET esclarecem as incertezas em torno do comportamento do ciclo hidrológico em meio às mudanças climáticas. Segundo Djaman et al. (2018), estimativas com precisão da ET são importantes para o monitoramento da seca, previsões meteorológicas, programação da irrigação, validações de modelos hidrológicos e previsão de incêndios florestais.

Diferentes métodos foram desenvolvidos para a medição da ET incluindo diretos, indiretos e modelos baseados em dados (*data driven models*). Tradicionalmente, a ET pode ser medida usando lisímetros (ALLEN et al., 2011; HOLMES, 1984), porém essa técnica é trabalhosa, demorada e de

custo elevado, além de ser limitada nas escalas temporal e espacial (FALAMARZI et al., 2014). Consequentemente, métodos indiretos e baseados em dados utilizando variáveis meteorológicas foram propostos, uma vez que Allen et al. (1998) enfatizaram que a ET é altamente influenciada pelos fatores climáticos.

A abordagem do modelo FAO-56 Penman-Monteith (PM) é amplamente recomendada e usada para a estimativa da evapotranspiração de referência (ET_o). Esse método baseia-se nas leis do balanço de energia e aerodinâmica, incorporando a equação de Penman-Monteith modificada por Allen et al. (1998) e fornece taxas da ET para uma cultura “ideal”, definida como uma grama verde hipotética com altura uniforme de 8 a 15 cm, com crescimento ativo e sem deficiência hídrica, resistência da superfície ao transporte de vapor d’água igual a 70 m s^{-1} e um albedo (poder refletor) de 0,23. No entanto, a utilização do modelo de PM requer quantidade e qualidade significativas dos dados, necessitando de variáveis climáticas como densidade do fluxo de calor, constante psicrométrica, inclinação da curva de pressão de vapor, pressão de saturação e pressão do vapor real, por isso a sua aplicação é limitada quando dados meteorológicos detalhados não estão disponíveis, fazendo-se necessário o desenvolvimento de modelos mais simples que utilizem uma quantidade reduzida de dados com adequada precisão.

Com o rápido avanço dos recursos computacionais e o surgimento da teoria da inteligência artificial, pesquisas utilizando técnicas do aprendizado de máquina foram realizadas para o desenvolvimento de métodos práticos e de baixo custo que estimem eficientemente a ET_o, como demonstrado em Chen et al., 2020; Chia et al., 2020; Granata, 2019; Mohammadi e Mehdizadeh, 2020; Raza et al., 2020; e Wang et al., 2019.

A eficiência de algoritmos de aprendizado de máquina em modelar a ET, caracterizada por ser um sistema não-linear, dinâmico e complexo, pode ser atribuída a capacidade desses modelos em encontrar relações potenciais dentro do conjunto de treinamento utilizado, o que se reflete diretamente no momento de teste, logo são capazes de lidar com problemas complexos e não-lineares (HU et al., 2021).

Raza et al. (2020) utilizaram cinco algoritmos de *machine learning*, entre eles *support vector machine* (SVM) e *multilayer perceptron* (MLP), para estimar a ET_o a partir dos dados meteorológicos de seis estações meteorológicas do Paquistão. Esses modelos, então, foram aplicados em outras estações que não possuem dados de treinamento disponíveis localizadas na China, Nova Zelândia e EUA para validar as suas eficientes performances. Os autores concluíram que os cinco algoritmos foram capazes de estimar a ET_o de forma satisfatória em diferentes zonas climáticas do Paquistão, bem como em outras partes do mundo, com erro médio de predição de $0.016 \text{ mm dia}^{-1}$, sendo que o SVM apresentou a melhor performance para as condições do estudo.

Nesse contexto, a hipótese da presente pesquisa é que, o uso de algoritmos de aprendizado de máquina pode ser uma alternativa ao método padrão de determinação da evapotranspiração de referência, utilizando-se uma quantidade reduzida de variáveis climáticas.

De modo a testar a hipótese citada, os objetivos dessa pesquisa foram: i) compreender as correlações entre as variáveis meteorológicas e a ETo na região de Piracicaba; ii) avaliar os conjuntos referentes às combinações de variáveis na acurácia de três algoritmos supervisionados; iii) avaliar a influência do tamanho da série de dados utilizada para treinamento dos modelos na estimativa da ETo; e iv) selecionar o conjunto que represente os melhores resultados para serem aplicados em um experimento de campo.

3.2. Material e Métodos

A presente pesquisa foi conduzida dentro das seguintes etapas: (1) coleta da base de dados brutos (temperaturas máxima, média e mínima, umidade relativa média, velocidade do vento e radiações solares global e líquida); (2) pré-processamento dos dados; (3) análise das variáveis mais relevantes para os modelos e (4) separação em amostra para treino, validação e teste, desenvolvimento de modelos preditivos, avaliação e comparação da acurácia dos modelos.

3.2.1. Área de estudo e coleta dos dados

Os dados coletados da série histórica referem-se à localidade de Piracicaba, no interior do estado de São Paulo, com latitude 22° 43' 30", longitude 47° 38' 01" e altitude de 546 m, área territorial de 1.378 km² e integrante de importante região produtora de cana-de-açúcar e milho do estado.

O clima da região é Aw, segundo a classificação climática de Köppen (ALVARES, SENTELHAS e DIAS, 2022), com temperatura média anual de 21,9 °C, umidade relativa média anual é de 86,2%, enquanto a velocidade do vento média no ano é de 1,46 m s⁻¹. Anualmente, chove cerca de 1257,8 mm. A região recebe os menores índices de radiação solar global durante os meses de maio, junho e julho (13,62 MJ m⁻² dia⁻¹, 12,48 MJ m⁻² dia⁻¹ e 13,32 MJ m⁻² dia⁻¹, respectivamente).

Os dados climáticos diários (Tabela 2) foram obtidos a partir de uma série histórica para o período referente de 01/01/2001 a 31/12/2021 (21 anos), do posto meteorológico da ESALQ/USP, localizado nas dependências do Departamento de Engenharia de Biosistemas (LEB) na cidade de Piracicaba, SP.

Tabela 2. Variáveis climáticas coletadas na base de dados do posto meteorológico da ESALQ/USP.

Variável	Tipo de variável	Descrição
Tmax	Contínua	Temperatura máxima (°C)
Tmin	Contínua	Temperatura mínima (°C)
Tmed	Contínua	Temperatura média (°C)
Uz	Contínua	Velocidade do vento ($m s^{-1}$)
URmed	Contínua	Umidade relativa média (%)
Rs	Contínua	Radiação solar global ($MJ m^{-2} dia^{-1}$)
Rn	Contínua	Radiação solar líquida ($MJ m^{-2} dia^{-1}$)

O período de 21 anos foi escolhido devido à qualidade dos sensores presentes no posto meteorológico, uma vez que a mistura de dados provenientes de sensores mais antigos com novos pode implicar em uma queda da qualidade dos dados e, conseqüente, redução na precisão dos modelos preditivos. Além disso, o ano de início da coleta de dados se refere ao primeiro ano em que a medição da Rn, importante variável na estimativa eficiente da ETo, foi realizada no posto meteorológico da ESALQ.

Pela Figura 1 pode-se conferir os valores das médias mensais das variáveis coletadas para o período referente a 01/01/2001 a 31/12/2021, as quais caracterizam o clima em Piracicaba, SP. Em uma escala mensal, a temperatura do ar variou entre 18,08 °C (julho) a 24,59 °C (fevereiro), com média de 21,94 °C. A umidade relativa do ar e a velocidade do vento foram em média de aproximadamente 80,53 % e $1,28 m s^{-1}$, respectivamente. O maior índice de radiação solar foi observado no mês de dezembro ($21,92 MJ m^{-2} dia^{-1}$) e o menor em junho ($12,99 MJ m^{-2} dia^{-1}$).

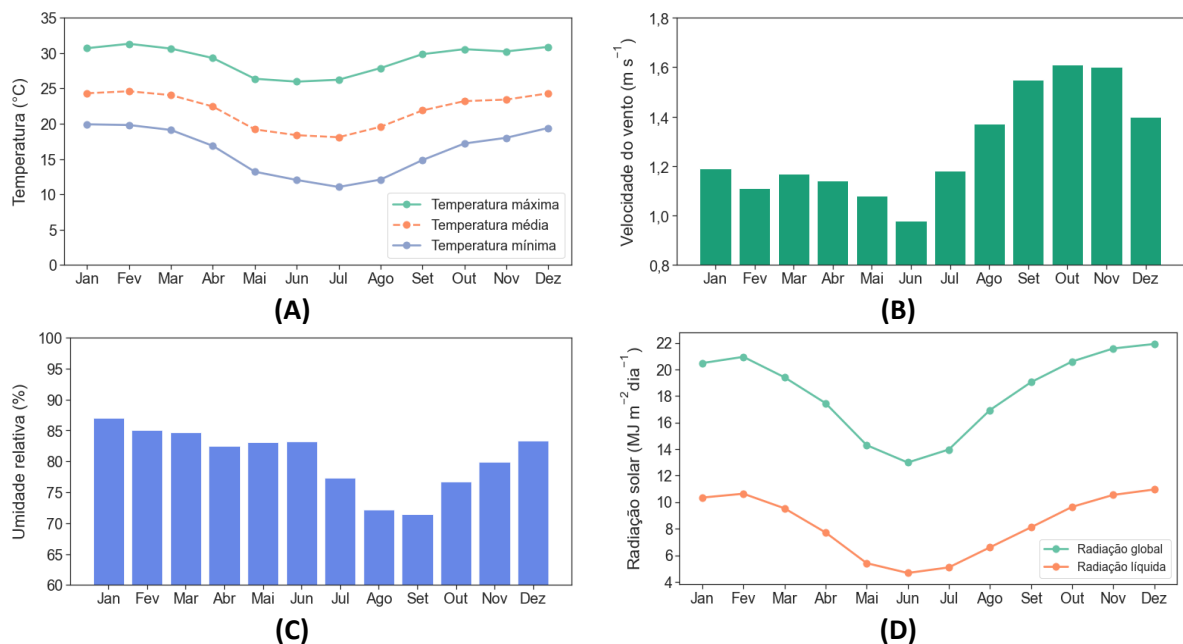


Figura 1. Valores das médias mensais das temperaturas máxima, média e mínima (A), velocidade do vento (B), umidade relativa (C), radiação solar global e líquida (D) durante o período de 01 de janeiro de 2001 – 31 de dezembro de 2021 em Piracicaba, SP.

Para se obter a melhor combinação de variáveis de entrada do modelo 13 conjuntos representando diferentes cenários foram aplicados aos algoritmos e avaliados a sua acurácia, com isso foi possível avaliar quais dados climáticos possuem maior correlação e relevância na predição da ETo para a região de Piracicaba e selecionar o conjunto utilizado no experimento de campo (Tabela 3). A construção dos cenários iniciou-se com a temperatura média do ar, pois essa variável geralmente está disponível na maioria das estações meteorológicas.

Tabela 3. Combinações de variáveis climáticas para compor os dados de entrada dos algoritmos.

Cenário	Dados de entrada
1	Tmed
2	Tmed + URmed
3	Tmed + Uz
4	Tmed + Rs
5	Tmed + Rn
6	Tmed + URmed + Uz
7	Tmed + URmed + Rs
8	Tmed + URmed + Rn
9	Tmed + Uz + Rs
10	Tmed + Uz + Rn
11	Tmed + Rs + Rn
12	Tmed + URmed + Uz + Rs
13	Tmed + URmed + Uz + Rn

Com o objetivo de verificar se o tamanho da série história, em que o desenvolvimento dos algoritmos se baseia, afeta o desempenho dos modelos, os cenários foram treinados e testados com base em uma série de 10 anos de dados climáticos (01/01/2012 a 31/12/2021) e outra de 21 anos (01/01/2001 a 31/12/2021).

3.2.2. Determinação da ETo pela abordagem do modelo FAO-56 Penman-Monteith

Por se tratar de um aprendizado supervisionado, os valores da ETo estimados pela equação da FAO-56 Penman-Monteith (Eq. 1) (Allen et al., 1998) foram utilizados como valores-alvo de referência, servindo como outputs no treinamento e teste para o desenvolvimento e avaliação dos modelos de *machine learning*. Portanto, foram determinados os valores da ETo diária para a base de dados coletada durante o período referente de 2001 a 2021.

$$ET_o = \frac{0,408\Delta(R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0,34u_2)} \quad (1)$$

Em que:

ET_o = evapotranspiração de referência, $mm\ d^{-1}$;

R_n = saldo de radiação à superfície da cultura, $MJ\ m^{-2}\ d^{-1}$;

G = densidade do fluxo de calor do solo, $MJ\ m^{-2}\ d^{-1}$;

T = temperatura do ar a 2 m de altura, $^{\circ}C$;

u_2 = velocidade do vento a 2 m de altura, $m\ s^{-1}$;

e_s = pressão de vapor de saturação, kPa ;

e_a = pressão parcial de vapor, kPa ;

Δ = declividade da curva de pressão de vapor de saturação, $kPa\ ^{\circ}C^{-1}$; e

γ = coeficiente psicrométrico, $kPa\ ^{\circ}C^{-1}$.

O método de PM é considerado o mais representativo para determinação indireta da ET_o e tem sido amplamente recomendado como método padrão pela FAO, uma vez que combina os efeitos energéticos com aerodinâmicos. Também vem sendo utilizado para calibrar outras abordagens de cálculo da ET_o .

A equação de PM é influenciada e controlada por vários parâmetros climáticos, aerodinâmicos e de resistência da superfície. Sua aplicação, porém, é dificultada por necessitar de diversas variáveis climáticas de entrada, bem como dados de alta qualidade e integração de constantes empíricas, sendo que as estações meteorológicas que satisfazem esses requisitos de observação são limitadas, especialmente em países em desenvolvimento. Nesses casos, as variáveis são frequentemente incompletas e nem sempre confiáveis. Além disso, a parte mais complexa para o uso dessa equação refere-se à obtenção da resistência da superfície do dossel. Por isso, é necessário explorar métodos mais simples para estimativa da ET_o com razoável precisão.

Pela Tabela 4 pode-se conferir os dados requisitados para FAO-PM56 em comparação com modelos baseados em dados. Pode ser visto que o método padrão depende de vários parâmetros que não são facilmente acessíveis.

Tabela 4. Dados necessários para aplicação dos métodos de determinação da ET_o (adaptado de Raza et al., 2020).

Fatores necessários	Tmax	Tmin	Uz	URmed	Rn	N	e_s	e_a	e_{max}	e_{min}	Outros fatores (Δ, γ e G)
FAO-PM56	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
<i>Data driven</i>	✓	✓	✓	✓	✓						

3.2.3. Pré-processamento e desenvolvimento dos modelos

As estações meteorológicas são susceptíveis a “falhas” em suas coletas de dados climáticos diários, por isso realizou-se uma etapa de pré-processamento dos dados para eliminação de valores nulos e padronização das variáveis climáticas.

O conjunto de dados construído apresentou 7535 observações e 11 variáveis. A divisão do banco de dados seguiu a proporção de 75% para treino e 25% para teste.

Foram calculadas as correlações entre as variáveis climáticas (dados de entrada dos modelos) com a variável-alvo (ETo), através do coeficiente de correlação de Pearson (Eq. 2), com o objetivo de analisar se existe alguma relação entre as variáveis e entender como a variabilidade de uma afeta a outra e para contribuir na escolha do melhor conjunto de variáveis que serão usadas como as *features* dos modelos.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

Em que n é o número de observações da amostra, \bar{x} é a média de x e \bar{y} é a média de y .

Os dados obtidos contêm características de dimensões e escalas distintas, o que afeta negativamente a modelagem e pode levar a um resultado tendencioso das previsões, uma vez que as variáveis que são medidas em escalas diferentes não contribuem igualmente para o ajuste e aprendizado do modelo e podem acabar criando um viés. Para lidar com esse problema, a padronização a partir do método *StandardScaler* da biblioteca Scikit-learn se apresenta como uma técnica de dimensionamento que torna os dados livres de escala, na qual a ideia principal é normalizar ($\mu = 0$ e $\sigma = 1$), individualmente, cada variável antes de aplicar o modelo de aprendizado de máquina, redimensionando a distribuição de valores para que a média dos valores observados seja 0 e o desvio-padrão 1.

A padronização foi realizada transformando os dados de acordo com a Eq. 3.

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

Em que x é o dado original, μ é a média e σ o desvio-padrão.

Após o desenvolvimento dos modelos, os dados também foram submetidos a uma validação cruzada (*cross validation*) com o objetivo de evitar o *overfitting*. Para isso, os dados brutos foram separados em 10 pacotes diferentes, onde 10% dos dados dentro de cada pacote foram guardados de modo a não serem utilizados durante o treino e teste, permitindo assim que o modelo fosse validado com condições de variáveis de entrada nunca antes processadas pelos modelos.

A montagem do banco de dados, pré-processamento e o treinamento e teste dos algoritmos foram realizados com o uso da biblioteca Scikit-Learn, uma biblioteca de código aberto

especializada em aprendizado de máquinas, da linguagem de programação Python versão 3.9, em ambiente de desenvolvimento Jupyter Notebooks.

3.2.4. Algoritmos aplicados

3.2.4.1. Support Vector Regression

O *Support vector regression* (SVR), aplicação do SMV à análise de regressão, utiliza um conjunto de funções matemáticas que são definidas como o kernel, sendo que a sua função é pegar os dados como entrada e transformá-los no formato necessário, ou seja, transformar dados linearmente inseparáveis em dados linearmente separáveis.

O algoritmo utilizou um subconjunto do conjunto de dados fornecido para construir uma equação adequada para o hiperplano. A equação generalizada para o hiperplano pode ser representada como mostra a Eq. 4:

$$f(x) = w^T \varphi(x) + b \quad (4)$$

em que: $f(x)$ é o vetor de saída, x é o vetor de entrada, w é o vetor do peso das características, $\varphi(x)$ é a função kernel que mapeia o vetor de entrada x em um espaço de características (*feature space*) de dimensão superior, onde os dados treinados apresentam linearidade, e b é um termo atribuído ao desvio.

As Eqs. 5 e 6 foram utilizadas para estimar o peso (w) e o termo de desvio (b) ao minimizar a função de perda (ou de erro). Para isso, é necessário minimizar o erro de previsão, ou seja, a distância entre os resultados previstos e os desejados. O SVR adotou uma função de perda ϵ -sensível, penalizando as previsões que estão mais distantes do que ϵ do dado de saída desejada. Várias funções de perda podem ser adotadas, incluindo a demonstrada na Eq. 5:

$$L_{\epsilon}(y, f(x, w)) = \begin{cases} 0, & \text{se } |y - f(x, w)| \leq \epsilon \\ |y - f(x, w)| - \epsilon, & \text{caso contrário} \end{cases} \quad (5)$$

Com base na Eq. 5, obtém-se o problema de otimização na Eq. 6:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (6)$$

em que, $\frac{1}{2} \|w\|^2$ é o termo de regularização, C é o fator de penalidade, L_{ϵ} é a função de perda ϵ -sensível, ϵ é a margem de tolerância do SVM, $\xi_i + \xi_i^*$ são as variáveis de folga adicionadas para proteção contra outliers.

O desempenho do SVR depende das funções kernel, aplicadas pela simplicidade de seu cálculo e pela sua capacidade de representar espaços abstratos. Uma função kernel K é utilizada para

mapear dados de entrada (conjunto de treinamento) do seu espaço original para um novo espaço de maior dimensão, chamado de espaço de características, com o objetivo de converter não-linearidade em linearidade e efetuar a separação dos dados (transformação do conjunto de treinamento para que ele seja linearmente separável). A estrutura do algoritmo pode ser verificada pela Figura 2.

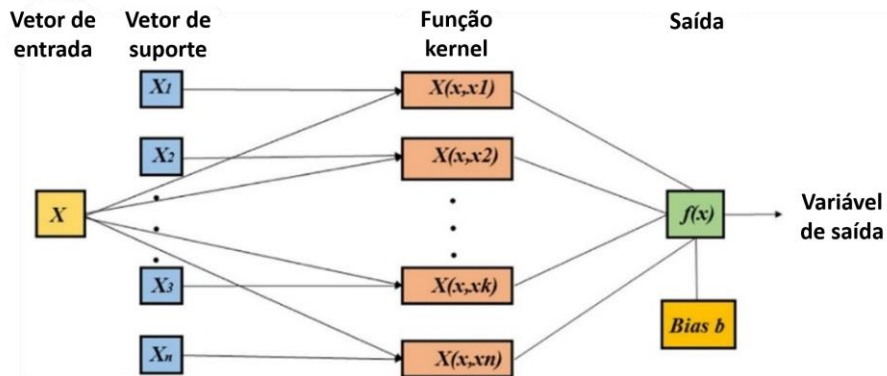


Figura 2. Estrutura do algoritmo SVR (adaptado de Mohammadi e Mehdizadeh, 2020; Vapnik e Chervonenkis, 1974).

A função kernel utilizada foi a *Radial Basis Function* (RBF), a qual é definida de forma semelhante a uma distribuição Gaussiana, projetando os dados de entrada em um espaço de dimensão infinita. Sua estrutura é demonstrada pela Eq. 7.

$$K(x_1, x_2) = e^{-\gamma \|x_1 - x_2\|^2}, \gamma > 0 \quad (7)$$

Em que: x_1 e x_2 representam a posição do vetor de entrada e γ é o parâmetro de ajuste do modelo.

3.2.4.2. Histogram-based Gradient Boosting Regression

Gradient boostings (GB) são formas de algoritmos de aprendizado de máquina baseados em modelos aditivos, caracterizados como uma metodologia *ensemble* (*ensemble method*), em que vários modelos preditivos “fracos” são criados e combinados com o objetivo de se obter um melhor desempenho como um todo.

O algoritmo pode ser pensado como uma soma F_m de estimadores fracos construídos de forma sequencial que prediz um vetor resposta y através da redução dos erros quadrados médios da fase de treinamento anterior. Se o modelo for treinado com M estimadores, o modelo final pode ser interpretado como apresentado na Eq. 8.

$$\hat{y}_i = F_M(x_i) = \sum_{m=1}^M F_m(x_i) \quad (8)$$

Em que, \hat{y}_i é o vetor resposta previsto, F_M é o valor predito pelo estimador fraco, x_i é o vetor de entrada, M é o número de estimadores usados na previsão do vetor resposta e F_m a soma ponderada do estimador fraco.

O estimador h_m foi adicionado para minimizar a soma das perdas L_m e um novo modelo F_m foi então construído com o objetivo de melhorar o seu desempenho com base no estimador fraco anterior F_{m-1} , como indicado nas Eqs. 9 e 10:

$$F_m(x) = F_{m-1}(x) + h_m(x) \quad (9)$$

$$h_m = \arg \min_m L_m = \arg \min_x \sum_{i=1}^n l(y_i, F_{m-1}(x_i) + h(x_i)) \quad (10)$$

Em que, $l(y_i, F_{m-1}(x_i))$ é a função de perda específica.

O *Histogram-based gradient boost* (HGBR), GB baseado em histograma, é uma abordagem do aprendizado de máquina introduzida recentemente por Chen e Guestrin (2016), caracterizada por agrupar variáveis contínuas em compartimentos (bins) discretos e utilizar esses compartimentos para construir histogramas das variáveis durante o treinamento e determinar as falhas dos modelos.

Essa técnica do ML é uma modificação do GB capaz de acelerar a fase de treino e reduzir o consumo da memória do modelo, tornando-se um algoritmo mais eficiente.

3.2.4.3. Multilayer Perceptron Regressor

Multi-layer perceptron regressor (MLPR) é um algoritmo baseado em redes neurais artificiais, composto por camadas de entrada e saída e uma ou mais camadas ocultas com vários neurônios empilhados juntos.

As entradas foram combinadas com os pesos iniciais em uma soma ponderada e submetidas à função de ativação, sendo que cada combinação foi propagada para a próxima camada, ou seja, cada camada alimentou a próxima com o seu resultado. Nessa abordagem, os dados fluem em uma única direção, ou seja, para frente, das camadas de entrada para a(s) camada(s) oculta(s) para a camada de saída.

Porém, se o algoritmo computasse apenas as somas ponderadas em cada neurônio, propagasse os resultados para a camada de saída e parasse nessa etapa, não teria a capacidade de aprender quais são os pesos que minimizam a função de custo, daí a necessidade da utilização da técnica da retropropagação.

A retropropagação consiste no mecanismo de aprendizado que permite o algoritmo ajustar iterativamente os pesos com o objetivo de minimizar a função de custo. Esse processo possui como

requisito que a função que combina entradas e pesos em um neurônio e a função de limite, por exemplo a ReLU, sejam diferenciáveis.

Em cada iteração, depois que as somas ponderadas foram propagadas por todas as camadas, o gradiente do erro quadrático médio (MSE) foi calculado em todos os pares de entrada e saída. Para propagá-lo de volta, os pesos da primeira camada oculta foram atualizados com o valor do gradiente. Esse processo continuou até que o gradiente para cada par de entrada e saída tenha convergido. Assim, a rede foi treinada mediante o ajuste dos pesos.

Matematicamente, os neurônios (α) podem ser descritos pela Eq. 11.

$$\alpha = \Phi \left(\sum_i w_i x_i + b \right) \quad (11)$$

Em que: Φ é a função de ativação, w_i é um vetor de peso, x_i é uma unidade de entrada e b é o desvio atribuído a cada camada de neurônios.

A função de ativação utilizada na presente pesquisa foi a ReLU, enquanto o peso foi determinado pelo otimizador ADAM, um método de otimização estocástico e o número máximo de iterações foi igual a 500. O algoritmo MLPR pode ser conferido pela Figura 3.

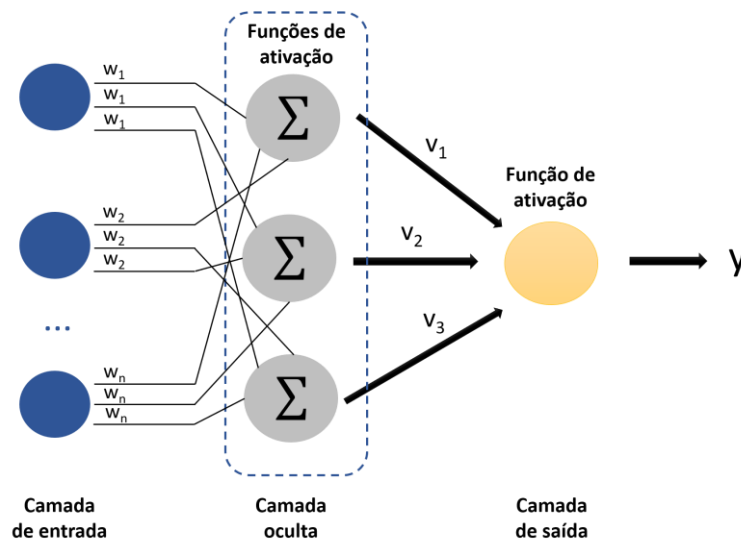


Figura 3. Estrutura do algoritmo MLPR (adaptado de Gerón, 2019).

3.2.5. Avaliação da performance dos modelos

Para avaliar o desempenho dos modelos nos momentos de treino, teste e validação, foram utilizados parâmetros estatísticos: R^2 , empregado para medir a variação entre os valores preditos

pelo modelo e os valores reais (Eq. 12) e erro quadrado médio (RMSE), utilizado para determinação da diferença entre os valores preditos e reais de ETo. (Eq. 13).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (13)$$

Em que, y_i é o valor observado, \hat{y}_i o valor predito, \bar{y} o valor médio observado e n o número de observações.

3.3. Resultados e Discussão

3.3.1. Correlação entre as variáveis do banco de dados

Pela Figura 4 pode-se conferir o mapa de calor, técnica de visualização de dados que evidencia a importância de um fenômeno através da cor em duas dimensões, referente ao banco de dados em estudo, o qual tem como objetivo analisar a relação entre as variáveis climáticas da região de Piracicaba e a ETo, a fim de compreender quais dos elementos presentes mostrou maior e menor influência a partir do cálculo da correlação entre as variáveis.

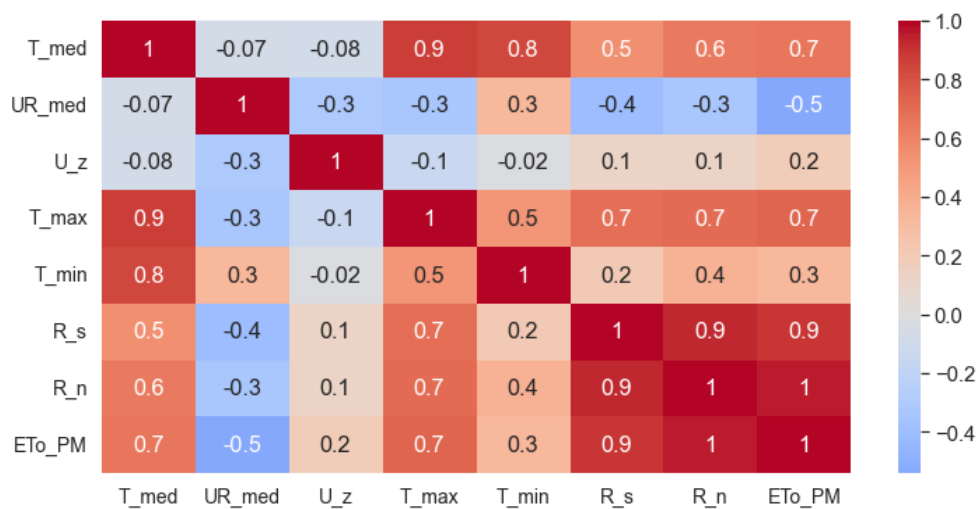


Figura 4. Mapa de calor apresentando a correlação entre as variáveis climáticas e a ETo determinada pelo método de Penman-Monteith.

Os resultados conferidos pela Figura 4 constataam que todos os dados selecionados apresentaram correlação significativa com a ETo, corroborando com diversos estudos que analisaram a influência de variáveis meteorológicas no comportamento da ETo, como a temperatura média (Liuzzo et al. 2016; Ariza-Villaverde et al. 2019; Tamaddun et al. 2019), radiação solar (Chatzithomas e Alexandris 2015, Gao et al., 2017), velocidade do vento (Liuzzo et al. 2016, Valipour et al. 2017) e umidade relativa do ar (Chatzithomas e Alexandris 2015; Ariza-Villaverde et al. 2019). Também pode-se concluir que a ETo está de fato diretamente relacionada aos elementos climáticos, como destacado por Allen et al. (1998), o qual propôs que a estimativa quantitativa desse processo poderia ser feita utilizando somente dados provenientes do clima.

É possível analisar que as correlações mais significativas com a ETo, ou seja, aquelas mais próximas de -1 ou 1, estão relacionadas à radiação solar (0,9 e aproximadamente 1 para radiação solar global e líquida, respectivamente), sendo essa considerada o elemento climático de maior importância na estimativa da taxa de ETo, como demonstrado por Chatzithomas e Alexandris (2015). A energia radiante do sol é a fonte primária de energia para a vida terrestre, sendo que a sua distribuição sobre a terra é o principal determinante das variações dos fatores climáticos, como alterações na temperatura média, uma vez que a energia solar absorvida pela atmosfera toma forma de calor ou energia cinética provocando o movimento do ar (vento).

As temperaturas máxima e média do ar também apresentaram elevada correlação com a ETo (iguais a 0,7 e 0,6, respectivamente). Segundo estudos de Aminzadeh et al. (2016), Ariza-Villaverde et al. (2019) e Dong e Dai (2017), a temperatura possui um papel determinante, pois a incidência da radiação solar no solo colabora para o processo de aquecimento e liberação de calor latente para a atmosfera, bem como na transferência de energia para a cultura na forma de fluxo de calor sensível, resultando no aumento da evaporação.

A umidade relativa do ar, por sua vez, apresentou valor de correlação negativa, com um coeficiente de -0,5, por isso, pode-se afirmar que à medida que a UR aumenta, a taxa de ETo tende a diminuir. Enquanto que a velocidade do vento assumiu terceira ordem de importância, sendo que essa variável normalmente apresenta maior relevância em regiões áridas, onde pode desempenhar um papel crítico em determinadas mudanças na evaporação e ETo.

Segundo Liuzzo et al. (2016), mudanças significativas na velocidade do vento provocam altas variações relativas na taxa de ETo se ocorrerem durante os meses de verão (estação de crescimento), quando as temperaturas também são mais altas. A velocidade do vento em Piracicaba, conforme conferido na Figura 1B, varia pouco ao longo do ano, atingindo seu máximo em $1,71 \text{ m s}^{-1}$ e seu mínimo em $1,06 \text{ m s}^{-1}$, o que explica, então, a baixa correlação entre essa variável e a ETo.

Filho et al. (2015) constataram, no estado da Paraíba, que a radiação solar e temperatura média foram as variáveis de maior influência sobre a ETo, enquanto que a velocidade do vento não

obteve efeito significativo e a umidade relativa apresentou um efeito negativo, reafirmando os resultados encontrados na presente pesquisa.

O fator determinante na estimativa da ETo se difere em diferentes regiões e em diferentes escalas. Gao et al. (2017), por exemplo, analisaram que, na China, a radiação solar é a variável com maior influência, seguida da umidade relativa.

Portanto, de forma geral, para a região de Piracicaba, a taxa de evapotranspiração deverá ser maior, quanto maior for a disponibilidade de radiação solar, maior a temperatura média do ar e o vento e quanto menor for a umidade relativa do ar.

3.3.2. Avaliação dos conjuntos

Os valores de ETo estimados pelo método da FAO-PM foram considerados os valores de referência para avaliar a aplicação dos modelos e conjuntos propostos durante os períodos de treino, teste e validação. Os indicadores estatísticos, R^2 e RMSE, referentes aos algoritmos construídos com base em série histórica de 10 anos, podem ser verificados pela Tabela 5 para o SVR, a Tabela 6 para o HGBR e a Tabela 7 para o MLPR.

Tabela 5. Performance comparativa do modelo SVR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia^{-1}) com base em série histórica de 10 anos de dados.

Cenário	Treino		Teste		Validação	
	R^2	RMSE (mm dia^{-1})	R^2	RMSE (mm dia^{-1})	R^2	RMSE (mm dia^{-1})
1	0,43	0,98	0,45	0,99	0,34	1,03
2	0,78	0,61	0,78	0,62	0,74	0,41
3	0,55	0,87	0,55	0,89	0,47	0,83
4	0,81	0,57	0,81	0,58	0,72	0,44
5	0,91	0,39	0,91	0,41	0,89	0,17
6	0,80	0,57	0,81	0,58	0,74	0,41
7	0,94	0,31	0,93	0,34	0,91	0,14
8	0,99	0,15	0,99	0,15	0,98	0,03
9	0,87	0,47	0,86	0,49	0,81	0,3
10	0,93	0,34	0,93	0,35	0,91	0,14
11	0,92	0,37	0,91	0,4	0,89	0,18
12	0,96	0,25	0,95	0,29	0,92	0,13
13	0,99	0,11	0,99	0,12	0,98	0,03

Tabela 6. Performance comparativa do modelo HGBR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia⁻¹) com base em série histórica de 10 anos de dados.

Cenário	Treino		Teste		Validação	
	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)
1	0,45	0,96	0,44	1,00	0,33	1,06
2	0,84	0,52	0,77	0,64	0,73	0,42
3	0,62	0,79	0,52	0,92	0,44	0,89
4	0,86	0,48	0,81	0,58	0,74	0,41
5	0,94	0,33	0,90	0,41	0,88	0,19
6	0,87	0,46	0,80	0,60	0,76	0,38
7	0,97	0,23	0,93	0,34	0,92	0,13
8	0,99	0,12	0,98	0,16	0,98	0,03
9	0,92	0,37	0,86	0,49	0,82	0,28
10	0,96	0,26	0,93	0,36	0,91	0,14
11	0,95	0,28	0,91	0,40	0,89	0,17
12	0,98	0,19	0,95	0,30	0,94	0,10
13	0,99	0,09	0,99	0,13	0,99	0,02

Tabela 7. Performance comparativa do modelo MLPR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia⁻¹) com base em série histórica de 10 anos de dados.

Cenário	Treino		Teste		Validação	
	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)
1	0,43	0,98	0,45	0,99	0,36	1,01
2	0,78	0,61	0,79	0,61	0,72	0,44
3	0,55	0,87	0,56	0,88	0,46	0,84
4	0,81	0,56	0,82	0,57	0,74	0,41
5	0,91	0,39	0,91	0,41	0,89	0,17
6	0,80	0,57	0,81	0,58	0,73	0,42
7	0,94	0,32	0,94	0,34	0,91	0,14
8	0,99	0,15	0,99	0,15	0,98	0,03
9	0,87	0,47	0,86	0,49	0,83	0,28
10	0,93	0,34	0,93	0,35	0,91	0,15
11	0,91	0,38	0,91	0,40	0,89	0,18
12	0,96	0,26	0,95	0,29	0,91	0,14
13	0,99	0,12	0,99	0,12	0,99	0,02

Os indicadores estatísticos, R² e RMSE, referentes aos algoritmos construídos com base em série histórica de 21 anos, podem ser verificados pela Tabela 8 para o SVR, a Tabela 9 para o HGBR e a Tabela 10 para o MLPR.

O treinamento dos algoritmos utilizando uma série histórica de 21 anos de dados climáticos evidenciou ainda mais a diferença da relevância entre Rs e Rn para a estimativa da ETo, onde a presença da variável Rn levou a um desempenho superior dos modelos. Além disso, foi possível

observar de forma mais contundente que o acréscimo da Uz nos cenários 12 e 13 não resultou em diferenças significativas na performance do algoritmos.

Tabela 8. Performance comparativa do modelo SVR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia⁻¹) com base em série histórica de 21 anos de dados.

Cenário	Treino		Teste		Validação	
	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)
1	0,47	0,90	0,48	0,90	0,44	0,83
2	0,74	0,62	0,73	0,64	0,71	0,43
3	0,55	0,83	0,56	0,82	0,51	0,72
4	0,86	0,47	0,85	0,48	0,82	0,27
5	0,92	0,34	0,92	0,36	0,91	0,13
6	0,77	0,59	0,76	0,61	0,71	0,42
7	0,93	0,34	0,92	0,35	0,89	0,16
8	0,99	0,14	0,99	0,14	0,99	0,02
9	0,88	0,43	0,88	0,44	0,84	0,24
10	0,94	0,31	0,94	0,32	0,92	0,12
11	0,93	0,33	0,92	0,35	0,91	0,13
12	0,94	0,30	0,93	0,32	0,90	0,15
13	0,99	0,10	0,99	0,11	0,99	0,02

Tabela 9. Performance comparativa do modelo HGBR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia⁻¹) com base em série histórica de 21 anos de dados.

Cenário	Treino		Teste		Validação	
	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)
1	0,49	0,88	0,47	0,91	0,43	0,84
2	0,79	0,56	0,72	0,66	0,71	0,43
3	0,62	0,77	0,55	0,83	0,50	0,74
4	0,88	0,42	0,85	0,48	0,82	0,27
5	0,94	0,31	0,92	0,36	0,91	0,14
6	0,83	0,51	0,76	0,62	0,73	0,39
7	0,95	0,29	0,92	0,36	0,90	0,15
8	0,99	0,12	0,99	0,15	0,99	0,02
9	0,91	0,36	0,88	0,43	0,85	0,23
10	0,95	0,26	0,94	0,32	0,92	0,11
11	0,95	0,28	0,92	0,34	0,91	0,14
12	0,96	0,25	0,93	0,32	0,92	0,12
13	0,99	0,09	0,99	0,12	0,99	0,01

Tabela 10. Performance comparativa do modelo MLPR entre treino, teste e validação dos cenários referentes a diferentes combinações de variáveis climáticas de entrada para determinação da ETo diária (mm dia⁻¹) com base em série histórica de 21 anos de dados.

Cenário	Treino		Teste		Validação	
	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)	R ²	RMSE (mm dia ⁻¹)
1	0,47	0,90	0,48	0,90	0,41	0,87
2	0,75	0,62	0,73	0,64	0,68	0,47
3	0,55	0,83	0,56	0,82	0,50	0,74
4	0,86	0,47	0,86	0,47	0,81	0,29
5	0,92	0,34	0,92	0,35	0,91	0,14
6	0,77	0,59	0,76	0,61	0,69	0,45
7	0,93	0,34	0,92	0,35	0,89	0,16
8	0,99	0,14	0,99	0,14	0,99	0,02
9	0,88	0,43	0,88	0,44	0,84	0,24
10	0,94	0,31	0,93	0,32	0,91	0,13
11	0,93	0,33	0,92	0,35	0,90	0,16
12	0,94	0,31	0,93	0,32	0,88	0,17
13	0,99	0,11	0,99	0,11	0,99	0,02

Os três modelos obtiveram desempenho semelhante na estimativa da ETo em 13 diferentes cenários tanto para 10 anos quanto 21 anos de série histórica. Foi observado que o R² e o RMSE variaram entre 0,33 a 0,99 e de 0,01 a 1,06 mm dia⁻¹, respectivamente, sendo que a presença ou ausência de elementos meteorológicos nos conjuntos de entrada impactou significativamente a performance do modelo.

O cenário 1 pode ser considerado um modelo tendo como base apenas a variável temperatura, o qual produziu a pior estimativa em todos os algoritmos (R²_{médio} = 0,43 e RMSE_{médio} = 0,94 mm dia⁻¹). Os cenários 2, 3, 4 e 5 foram formados adicionando à Tmed, a URmed, a Uz, a Rs e a Rn, respectivamente, representando modelos utilizando duas variáveis de entrada. Do ponto de vista dos indicadores estatísticos, 2 (R²_{médio} = 0,75 e RMSE_{médio} = 0,55 mm dia⁻¹), 3 (R²_{médio} = 0,53 e RMSE_{médio} = 0,83 mm dia⁻¹), 4 (R²_{médio} = 0,82 e RMSE_{médio} = 0,46 mm dia⁻¹) e 5 (R²_{médio} = 0,91 e RMSE_{médio} = 0,30 mm dia⁻¹), obtiveram desempenhos superiores ao cenário 1. Portanto, pode-se inferir que a ordem dos fatores que influenciam a ETo, além da temperatura média do ar, em Piracicaba é: Rn, Rs, URmed e Uz.

Foi observado que a robustez e a eficiência dos algoritmos foram aumentando à medida em que um maior número de variáveis eram adicionadas aos conjuntos de entrada, o que também foi evidenciado por Antonopoulos e Antonopoulos (2017) e Fan et al. (2018). Nesse contexto, também foram propostos cenários em que os algoritmos eram alimentados com três variáveis de entrada, onde foi possível analisar que a combinação de temperatura do ar, radiação solar e umidade relativa do ar resulta em melhores performances dos modelos.

Nos cenários 12 e 13 foram considerados um conjunto completo de variáveis meteorológicas quase semelhante ao requerido pelo método da FAO-PM, o qual gerou os melhores resultados ($R^2_{\text{médio}} = 0,94$ e $\text{RMSE}_{\text{médio}} = 0,23 \text{ mm dia}^{-1}$; $R^2_{\text{médio}} = 0,99$ e $\text{RMSE}_{\text{médio}} = 0,08 \text{ mm dia}^{-1}$, respectivamente), indicando que os três algoritmos foram capazes de encontrar correlações mais próximas aos valores estimados pela FAO-PM no momento em que reuniu-se quatro variáveis relacionados à temperatura, umidade relativa, vento e radiação solar, as quais influenciam diretamente a ETo (Figura 4).

Para a região de Piracicaba, como a Uz possui baixa correlação com a ETo ($= 0,2$), os resultados dos cenários 8 e 13 são semelhantes ($R^2_{\text{médio}} = 0,99$ e $\text{RMSE}_{\text{médio}} = 0,10 \text{ mm dia}^{-1}$; $R^2_{\text{médio}} = 0,99$ e $\text{RMSE}_{\text{médio}} = 0,08 \text{ mm dia}^{-1}$, respectivamente), evidenciado que essa variável possui menor peso na estimativa da ETo. Porém, quando utilizada nos algoritmos em conjunto com mais três variáveis de entrada (13), resultou em menor RMSE do que o apresentado pelo cenário 8, portanto, pode-se concluir que a sua presença e interação com a Tmed, URmed e Rn resultam em melhor capacidade preditiva dos algoritmos, corroborando com resultados encontrados por Chen et al. (2020).

Em relação aos tipos de radiação solar, constatou-se que a Rn possui maior influência na ETo do que a Rs, uma vez que os índices estatísticos são superiores quando a Rn faz parte da combinação dos dados de entrada, isso é evidenciado comparando-se os cenários 4 e 5, 7 e 8, 9 e 10 e 12 e 13.

Além disso, os cenários considerando a Rn (5, 8, 10, 11 e 13) alcançaram performances superiores em relação aos que não utilizaram essa variável como variável de entrada, comprovando a importância desse elemento meteorológico para a estimativa da ETo em Piracicaba. A importância da Rn também foi destacada por Chatzithomas e Alexandris (2015) e Wang et al. (2019).

Em todos os cenários propostos, apesar das diferenças serem pequenas, o algoritmo HGBR apresentou melhores índices estatísticos na predição da ETo.

A partir dos resultados apresentados, é possível perceber que as estimativas feitas para os 13 cenários seguiram o mesmo padrão tanto para algoritmos treinados com 10 anos de dados quanto 21 anos, por isso o tamanho da série não afetou o desempenho dos modelos.

Esta seção resume a interessante capacidade dos algoritmos SVR, HGBR e MLPR em estimar a ETo para a localidade de Piracicaba sob diferentes conjuntos de dados de entrada e tamanhos de série histórica.

3.3.3. Avaliação dos algoritmos para experimento de campo

Na seção anterior foram testados diferentes cenários para a predição da ETo com o objetivo de analisar qual seria a combinação de variáveis climáticas que resultaria nos menores valores de RMSE e maiores R^2 , ou seja, no cenário mais “preciso”. Tal cenário foi escolhido para ser testado em um experimento de campo, uma vez que o manejo da irrigação da cultura do milho pode ser realizado com base nos valores estimados da ETo.

Nesse contexto, o cenário 13 foi o escolhido, em que as variáveis de entrada são: Tmed, URmed, Uz e Rn.

A utilização de algoritmos de *machine learning* com quantidade reduzida de variáveis de entrada mostrou efeito positivo e significativo como uma alternativa para a estimativa da ETo sem o uso de modelos empíricos (Tabelas 11 e 12). O modelo HGBR apresentou o melhor desempenho na predição da ETo para as condições climáticas de Piracicaba nos períodos de treino e validação, já durante o teste o SVR mostrou-se superior. Isso indicou que, para as condições em estudo, o modelo HGBR demonstrou ser mais suscetível ao sobreajuste, ou seja, o rendimento do modelo decresce quando submetido a novos dados de Tmed, URmed, Uz e Rn. Vale destacar que as diferenças nos índices estatísticos entre os três algoritmos são mínimas.

Tabela 11. Performance comparativa dos algoritmos na estimativa da ETo, com base em 10 anos de dados, nos períodos de treino, teste e validação referentes ao cenário escolhido para o experimento de campo.

Modelo	Treino		Teste		Validação	
	R^2	RMSE (mm dia ⁻¹)	R^2	RMSE (mm dia ⁻¹)	R^2	RMSE (mm dia ⁻¹)
SVR	0,992	0,113	0,992	0,119	0,984	0,025
HGBR	0,996	0,086	0,990	0,132	0,988	0,019
MLPR	0,991	0,122	0,991	0,122	0,987	0,021

Tabela 12. Performance comparativa dos algoritmos na estimativa da ETo, com base em 21 anos de dados, nos períodos de treino, teste e validação referentes ao cenário escolhido para o experimento de campo.

Modelo	Treino		Teste		Validação	
	R^2	RMSE (mm dia ⁻¹)	R^2	RMSE (mm dia ⁻¹)	R^2	RMSE (mm dia ⁻¹)
SVR	0,993	0,100	0,993	0,107	0,987	0,019
HGBR	0,995	0,086	0,991	0,117	0,990	0,014
MLPR	0,992	0,108	0,993	0,107	0,988	0,017

O bom desempenho do algoritmo SVR também já foi demonstrado por outros autores, como Fan et al. (2018), Tang et al. (2019) e Wen et al. (2015), os quais destacam ainda a estabilidade apresentada por esse modelo durante os períodos de treino, teste e validação.

Embora tenha alcançado resultado satisfatório, verificou-se uma inferioridade na precisão do MLPR, reforçando o resultado encontrado por Wu e Fan (2019), os quais utilizaram apenas a temperatura como variável de entrada para determinação da ETo em região subtropical úmida da China, onde o MLPR obteve desempenho inferior ao produzido pelo SVR.

A confiabilidade dos modelos de regressão em fazer previsões da ETo foi garantida pelos altos e baixos valores de R^2 e RMSE, respectivamente. Os valores de R^2 indicaram que esses modelos podem fazer estimativas para novos dados climáticos com uma precisão de 99% por meio de uma combinação linear de variáveis relacionadas à temperatura, radiação solar, umidade e vento.

O uso de modelos clássicos de aprendizado de máquina para a predição da ETo com quantidade limitada de dados climáticos, atraiu a atenção de diversos pesquisadores nos últimos anos, os quais encontraram como resultado o desempenho semelhante ou até superior desses algoritmos em relação aos métodos padrões de determinação da ETo, demonstrando o quão promissora é a metodologia apresentada na presente pesquisa (DOU e YANG, 2018; FENG et al., 2017; GRANATA, 2019; YIN et al., 2017).

Uma comparação mais abrangente entre os valores da ETo calculados a partir do método da FAO-PM e os valores estimados por cada algoritmo com base nas séries de dados de 10 e 21 anos, pode ser conferida pela Figura 5. Nela é possível analisar que os modelos geraram estimativas pouco dispersas e com elevado R^2 . O MLPR e o SVR produziram algumas estimativas mais dispersas do que o HGBR, especialmente para valores de ETo menores do que 3 mm dia^{-1} . Além disso, pode-se perceber que as estimativas seguiram o mesmo padrão tanto para algoritmos treinados com 10 anos de dados quanto 21 anos, ou seja, o tamanho da série não afetou o desempenho dos modelos.

A distribuição dos valores reais da ETo com base no método FAO-PM e os valores preditos pelos algoritmos utilizados com base em séries de 10 e 21 anos de dados podem ser conferidos pela Figura 6A e 6B, respectivamente. Pode-se observar que o SVR e HGBR tendem a subestimar a ETo quando esta é maior que 5 mm dia^{-1} . Enquanto que, a mediana e a amplitude interquartílica dos três modelos se encontram bem próximas àquelas calculadas pelo método padrão, evidenciando que os algoritmos se apresentam como uma alternativa para a determinação da ETo na região de Piracicaba.

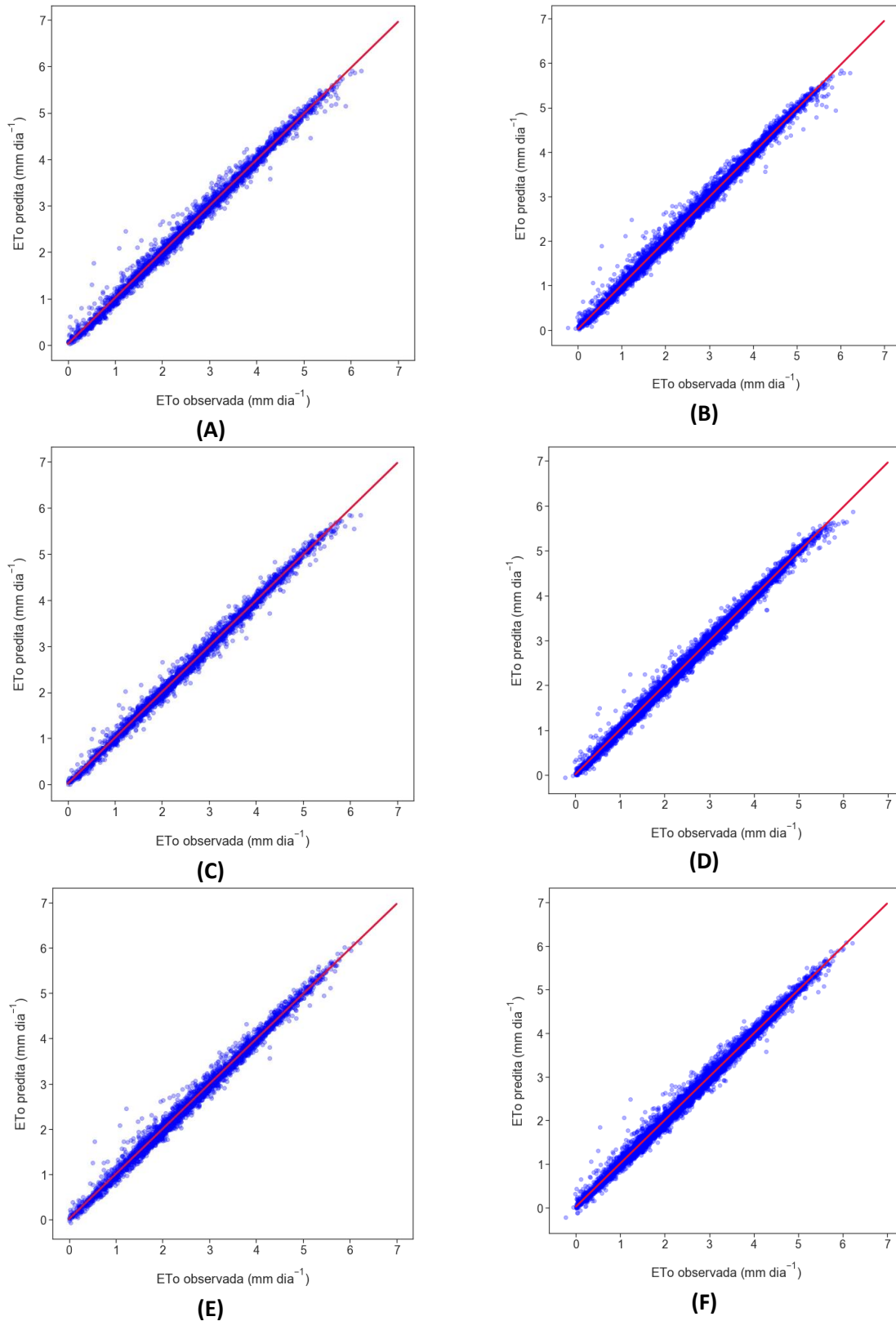
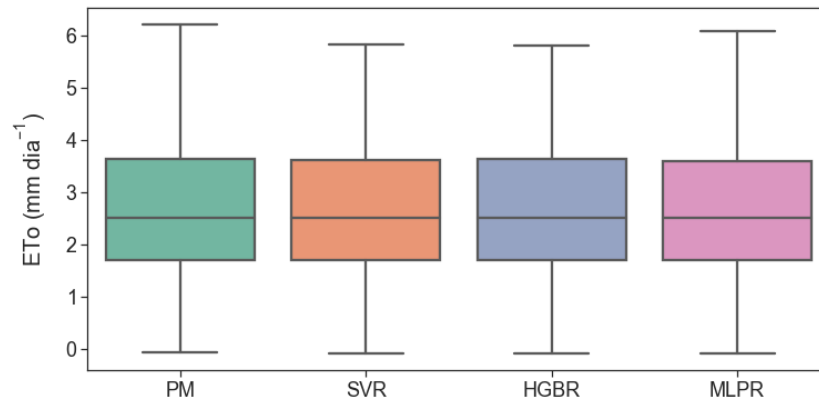
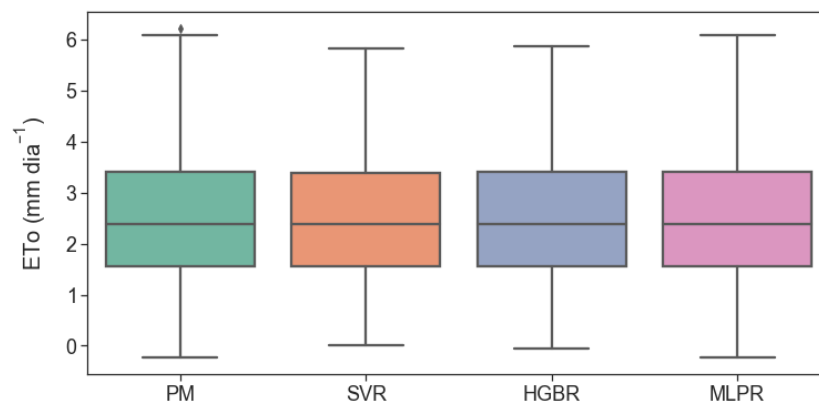


Figura 5. Comparação entre os valores de ETo observados usando o método da FAO-PM e os valores preditos pelos algoritmos com base em séries de 10 e 21 anos de dados, respectivamente: SVR 10 anos (A), SVR 21 anos (B), HGBR 10 anos (C), HGBR 21 anos (D), MLPR 10 anos (E) e MLPR 21 anos (F).

Também é possível analisar como o tamanho da série de dados não alterou o padrão de distribuição dos valores, uma vez que o boxplot seguiu o mesmo padrão tanto para algoritmos treinados com 10 anos de dados quanto 21 anos.



(A)



(B)

Figura 6. Boxplot referente aos valores da ETo diária calculados a partir do método padrão (FAO-PM) e estimados pelos algoritmos em estudo para a região de Piracicaba com base em 10 (A) e 21 anos de dados (B).

Por uma questão de concisão e legibilidade, pela Figura 7 pode-se conferir a ETo média mensal predita pelos algoritmos com base em série de 10 e 21 anos de dados em comparação aos valores observados usando o método da FAO-PM, assim é possível verificar o comportamento das estimativas ao longo do ano.

Pode-se observar que o comportamento de PM e dos valores preditos mostraram-se sempre muito próximos, onde os valores estimados pelo algoritmo HGBR são os que mais se aproximaram dos valores calculados pelo método padrão, uma vez que as linhas praticamente se sobrepõem.

As predições da ETo médias mensais realizadas pelo MLPR foram as que apresentaram maiores diferenças em relação ao PM. É possível observar uma tendência de subestimativa nos meses de março e abril, enquanto que em setembro e outubro ocorre uma superestimativa nos

valores observados. Logo, pode ser concluído que esse algoritmo possui uma capacidade inferior de se relacionar de maneira adequada com as variáveis de entrada escolhidas, assim, frente a possíveis eventos extremos, seu poder de determinação da ETo é reduzido.

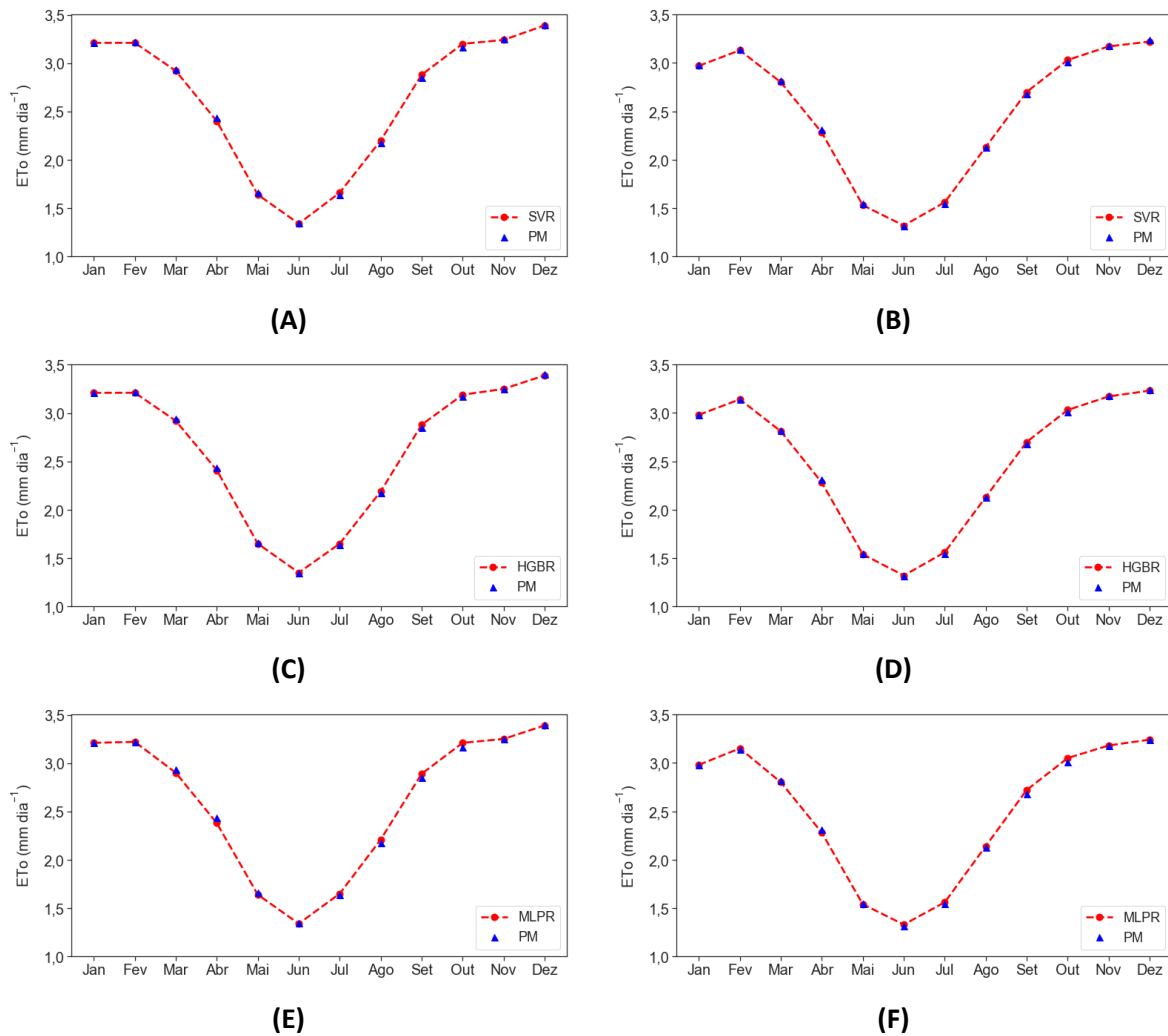


Figura 7. Evapotranspiração média mensal calculada a partir do método padrão (PM) e dos algoritmos treinados e testados: SVR 10 anos (A), SVR 21 anos (B), HGBR 10 anos (C), HGBR 21 anos (D), MLPR 10 anos (E) e MLPR 21 anos (F).

Uma vez obtendo esses resultados, foi possível notar que os três modelos foram capazes de encontrar correlações próximas ao método FAO-PM, utilizando quatro elementos meteorológicos (Tmed, URmed, Rs e Uz).

É possível perceber também que as estimativas seguiram o mesmo padrão tanto para algoritmos treinados com 10 anos de dados quanto 21 anos. Portanto, com o objetivo de melhorar o desempenho de simulação dos modelos, foram escolhidos os algoritmos desenvolvidos com 10 anos de dados para utilização no manejo de irrigação do experimento de campo.

3.4. Conclusões

Diante dos resultados obtidos pode-se concluir que:

- a. As variáveis meteorológicas obtidas para a região de Piracicaba apresentaram correlações significativas com a ETo, especialmente Rn, Rs, Tmed e URmed. Por isso, elas podem ser utilizadas para estimar quantitativamente esse processo, provando que a ETo está de fato diretamente relacionada aos elementos climáticos;
- b. Em relação à construção de modelos preditivos, observou-se que quanto maior a quantidade de variáveis de entrada, maior a chance do algoritmo produzir um desempenho satisfatório quanto à estimativa da ETo. Os resultados obtidos indicaram que o conjunto contendo Tmed, URmed, Uz e Rn como variáveis de entrada, resultou nas melhores performances dos algoritmos de aprendizado de máquina para as condições estudadas;
- c. O tamanho da série histórica utilizada para construção dos algoritmos não afetou o desempenho dos modelos; e
- d. As estimativas indicaram que todos os modelos aplicados (SVR, HGBR e MLPR) foram capazes de identificar relações lineares e não lineares entre ETo e dados meteorológicos, resultando em desempenho satisfatório da predição da ETo utilizando quatro variáveis de entrada.

Portanto, pode-se concluir que os algoritmos de aprendizado de máquina adotados são ferramentas de alto potencial para a predição da ETo, nas condições climáticas de Piracicaba, e assim, se apresentam como alternativas para a determinação desse processo frente às dificuldades relatadas sobre o método padrão (FAO-PM).

Referências

- ALLEN, R. G.; PEREIRA, L. S.; HOWELL, T. A.; JENSEN, M. E. Evapotranspiration information reporting: I. Factors governing measurement accuracy. **Agricultural Water Management**, v. 98, n. 6, p. 899–920, 2011.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. **Crop evapotranspiration: guidelines for computing crop water requirements**. Rome: FAO, 300 p., 1998 (FAO irrigation and drainage paper no. 56).
- ALVARES, C. A.; SENTELHAS, P. C.; DIAS, H. B. Southeastern Brazil inland tropicalization: Köppen system applied for detecting climate change throughout 100 years of meteorological observed data. **Theoretical and Applied Climatology**, v. 149, p. 1431–1450, 2022.
- AMINZADEH M.; RODERICK M. L.; OR, D. A generalized complementary relationship between actual and potential evaporation defined by a reference surface temperature. **Water Resources Research**, v. 52, p. 385–406, 2016.

- ANTONOPOULOS, V. S.; ANTONOPOULOS, A. V. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate. **Computers and Electronics in Agriculture**, v. 132, p. 86–96, 2017.
- ARIZA-VILLAVERDE, A.B.; PAVÓN-DOMÍNGUEZ, P.; CARMONA-CABEZAS, R.; DE RAVÉ, E. G.; JIMÉNEZ-HORNERO, F. J. Joint multifractal analysis of air temperature, relative humidity and reference evapotranspiration in the middle zone of the Guadalquivir river valley. **Agricultural and Forest Meteorology**, v. 278, 107657, 2019.
- CHATZITHOMAS, C. D.; ALEXANDRIS, S. G. Solar radiation and relative humidity based, empirical method, to estimate hourly reference evapotranspiration. **Agricultural Water Management**, v. 152, p. 188–197, 2015.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. **In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining**, p. 785-794, 2016.
- CHEN, Z.; ZHU, Z.; JIANG, H.; SUN, S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. **Journal of Hydrology**, v. 591, p. 125–286, 2020.
- CHIA, M. Y.; HUANG, Y. F.; KOO, C. H. Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. **Computers and Electronics in Agriculture**, v. 175, p. 1–12, 2020.
- DJAMAN, K.; O'NEILL, M.; OWEN, C.K.; SMEAL, D.; KOUDAHE, K.; WEST, M.; ALLEN, S.; LOMBARD, K.; IRMAK, S. Crop evapotranspiration, irrigation water requirement and water productivity of maize from meteorological data under semiarid climate. **Water (Switzerland)**, v. 10, n. 4, p. 1–17, 2018.
- DONG, B.; DAI, A. The uncertainties and causes of the recent changes in global evapotranspiration from 1982 to 2010. **Climate Dynamics**, v. 49, p. 279–296, 2017.
- DOU, X.; YANG, Y. Modeling evapotranspiration response to climatic forcings using data-driven techniques in grassland ecosystems. **Advances in Meteorology**, v. 18, 1824317, 2018.
- FALAMARZI, Y.; PALIZDAN, N.; HUANG, Y. F.; LEE, T. S. 2014. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). **Agricultural Water Management**, v. 140, p. 26–36, 2014.
- FAN, J.; YUE, W.; WU, L.; ZHANG, F.; CAI, H.; WANG, X.; LU, X.; XIANG, Y. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. **Agricultural and Forest Meteorology**, v. 263, p. 225–241, 2018.
- FENG, Y.; CUI, N.; GONG, D.; ZHANG, Q.; ZHAO, L. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. **Agricultural Water Management**, v. 193, p. 163–173, 2017.

- FERREIRA, L. B.; da CUNHA, F. F.; de OLIVEIRA, R. A.; FERNANDES FILHO, E. I. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – A new approach. **Journal of Hydrology**, v. 572, p. 556–570, 2019.
- FILHO, A. I.; BORGES, P. F.; ARAÚJO, L. S.; PEREIRA, A. R.; LIMA, E. M.; SILVA, L. S.; JUNIOR, C. V. S. Influência das variáveis climáticas sobre a evapotranspiração. **Gaia Scientia**, v. 9, p. 62–66, 2015.
- GAO, Z.; HE, J.; DONG, K.; LI, X. Trends in reference evapotranspiration and their causative factors in the West Liao River basin, China. **Agricultural and Forest Meteorology**, v. 232, p. 106–117, 2017.
- GÉRON, A. *Mãos à Obra: Aprendizado de Máquina com Scikit-Learn & TensorFlow*. Traduzido por Rafael Contatori, 1 ed., Alta Books, Rio de Janeiro, RJ, Brasil, 2019.
- GRANATA, F.; GARGANO, R.; de MARINIS, G. Artificial intelligence based approaches to evaluate actual evapotranspiration in wetlands. **Science of the Total Environment**, v. 703, p. 1–15, 2020.
- HOLMES, J. W. Measuring evapotranspiration by hydrological methods. **Agricultural Water Management**, v. 8, n. 1–3, p. 29–40, 1984.
- HU, X.; SHI, L.; LIN, G.; LIN, L. Comparison of physical-based, data-driven and hybrid modeling approaches for evapotranspiration estimation. **Journal of Hydrology**, v. 601, p. 1–13, 2021.
- JING, W.; YASEEN, Z. M.; SHAHID, S.; SAGGI, M. K.; TAO, H.; KISI, O.; SALIH, S. Q.; AL-ANSARI, N.; CHAU, K. W. Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions. **Engineering Applications of Computational Fluid Mechanics**, v. 13, p. 811–823, 2019.
- KRISHNA, P. Evapotranspiration and agriculture-A review. **Agricultural reviews**, v. 40, p. 1–11, 2019.
- LIUZZO, L.; VIOLA, F.; NOTO, L. V. Wind speed and temperature trends impacts on reference evapotranspiration in Southern Italy. **Theoretical and Applied Climatology**, v. 123, p. 43–62, 2016.
- MOHAMMADI, B.; MEHDIZADEH, S. Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. **Agricultural Water Management**, v. 237, p. 1–13, 2020.
- RAZA, A.; SHOAB, M.; FAIZ, M. A.; BAIG, F.; KHAN, M. M.; ULLAH, M. K.; ZUBAIR, M. Comparative assessment of reference evapotranspiration estimation using conventional method and machine learning algorithms in four climatic regions. **Pure and Applied Geophysics**, v. 177, n. 9, p. 4479–4508, 2020.
- TAMADDUN, K. A.; KALRA, A.; BERNARDEZ, M.; AHMAD, S. Effects of ENSO on temperature, precipitation, and potential evapotranspiration of north india's monsoon: an analysis of trend and entropy. **Water**, v. 11, 189, 2019.

- TANG, D.; FENG, Y.; GONG, D.; HAO, W.; CUI, N. Evaluation of artificial intelligence models for actual crop evapotranspiration modeling in mulched and non-mulched maize croplands. **Computers and Electronics in Agriculture**, v. 152, p. 375–384, 2018.
- VALIPOUR, M.; MA, G. S.; KHOSHRAVESH, M. Estimation and trend evaluation of reference evapotranspiration in a humid region. **Italian Journal of Agrometeorology**, v. 1, p. 19–36, 2017.
- VAPNIK, V.; CHERVONENKIS, A. **Theory of pattern recognition**, Nauka, Moscow, 1974.
- WANG, S.; LIAN, J.; PENG, Y.; HU, B.; CHEN, H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. **Agricultural Water Management**, v. 221, p. 220–230, 2019.
- WANNIARACHCHI, S.; SARUKKALIGE, R. A Review on Evapotranspiration Estimation in Agricultural Water Management: Past, Present, and Future. **Hydrology** 9, 123, 2022.
- WEN, X.; SI, J.; HE, Z.; WU, J.; SHAO, H.; YU, H. Support-Vector-Machine-Based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. **Water Resources Management**, v. 29, p. 3195–3209, 2015.
- WU, L.; FAN, J. Comparison of neuron-based, kernel-based, tree-based and curve-based machine learning models for predicting daily reference evapotranspiration. **PLoS ONE**, v. 14, e0217520, 2019.
- YIN, Z.; FENG, Q.; YANG, L.; DEO, R. C.; WEN, X.; SI, J.; XIAO, S. Future Projection with an Extreme-Learning Machine and Support Vector Regression of Reference Evapotranspiration in a Mountainous Inland Watershed in North-West China. **Water**, v. 9, 880, 2017.

4. USO DE MACHINE LEARNING E MODELO DSSAT/CERES-MAIZE NA ESTIMATIVA DA PRODUTIVIDADE DE MILHO EM PIRACICABA, SP

Resumo

O milho se caracteriza como um produto fundamental para a agricultura brasileira, cultivado em todas as regiões do país, sendo que a sua produção pode ser influenciada pela variabilidade climática, fertilidade do solo, tratos culturais, data de semeadura, manejo de irrigação, fatores bióticos, entre outros. Outro aspecto relevante é o uso de modelos, amplamente utilizados em termos mundiais, como pesquisa operacional e estratégica e suporte à decisão na produção agrícola e gestão de recursos. O modelo DSSAT (*Decision Support System for Agrotechnology Transfer*) é um dos sistemas de modelo de cultura mais utilizados para avaliação do desenvolvimento de plantas pelo qual é possível simular processos fisiológicos de fotossíntese, respiração e distribuição de matéria seca. Assim sendo, o objetivo da presente pesquisa foi avaliar algoritmos de aprendizado supervisionado na estimativa da produtividade de milho para a região de Piracicaba, tendo como base dados climáticos e simulações pelo modelo DSSAT. Para isso, foram obtidos 30 anos de variáveis meteorológicas do posto meteorológico da ESALQ/USP, além de dados da cultura e do solo da região, e foram realizadas simulações pelo modelo DSSAT, já calibrado para tal situação, em três épocas de semeadura diferentes. Após obtenção das produtividades pelo modelo DSSAT, foram treinados e testados, dois algoritmos de aprendizado de máquina – *extra-trees* (EXT) e *random forest* (RF) para predição do rendimento de milho. Pelos resultados obtidos, pode-se notar que produção de milho simulada pelo DSSAT para as condições de Piracicaba variou conforme a época de plantio e que as variáveis meteorológicas apresentaram correlações significativas com a produtividade. Em relação aos algoritmos treinados e testados, os valores de RMSE e R^2 indicaram que os modelos aplicados apresentaram desempenho satisfatório na predição do rendimento de grãos de milho na região de Piracicaba em comparação com o DSSAT, apresentando-se como sendo uma alternativa aos modelos clássicos de crescimento e desenvolvimento de culturas. Portanto, os algoritmos de aprendizado de máquina podem auxiliar em um sistema de decisão, diante de variáveis de entrada, demonstrando uma capacidade de prever o desenvolvimento da cultura do milho para a região de Piracicaba, SP.

Palavras-chave: Rendimento de grãos, *Random forest*, *Extra-trees*, Predição da produtividade

Abstract

Maize is characterized as a fundamental product for Brazilian agriculture, cultivated in all regions of the country, its production can be influenced by climate variability, soil fertility, cultural practices, sowing date, irrigation management, biotic factors, among others. In the context of growing food demand, it becomes necessary to search for tools that enable decision-making based on reliable information. In this sense, crop models have been developed and used worldwide as operational and strategic research, decision support in agricultural production and resource management. The DSSAT (*Decision Support System for Agrotechnology Transfer*) model is one of the most widely used crop modeling systems for evaluating plant development, through which it is possible to simulate physiological processes such as photosynthesis, respiration and dry matter distribution. Therefore, the objective of this research was to evaluate supervised learning algorithms in estimating maize yield in the Piracicaba region based on climate data and from DSSAT simulations. For this purpose, 30 years of meteorological variables from the ESALQ/USP weather station were obtained, as well as crop and soil data of the region, and simulations for three different sowing seasons were performed in DSSAT. After obtaining maize

yields from DSSAT, two machine learning algorithms – extra-trees (EXT) and random forest (RF) - were trained and tested for yield prediction. Results showed that maize production simulated by DSSAT for Piracicaba conditions varied according to planting season, and that meteorological variables showed significant correlations with productivity. Regarding the algorithms, the RMSE and R^2 values indicated that the applied models demonstrated satisfactory performance in predicting maize grain yield in the Piracicaba region compared to DSSAT, presenting themselves as an alternative to the classical crop growth and development models. Therefore, machine learning algorithms can provide a decision system, based on input variables, with the ability to predict maize development in the Piracicaba, SP, region.

Keywords: Grain yield, Random forest, Extra-trees, Yield prediction

4.1. Introdução

O milho (*Zea mays* L.) constitui-se em um dos mais importantes cereais cultivados, consumidos e estudados no mundo, sendo responsável por fornecer produtos utilizados para alimentação humana e animal, bem como, matérias-primas para a indústria. Devido à multiplicidade de suas aplicações, a cultura do milho assume papel relevante em termos socioeconômicos (GREAVES e WANG, 2017).

No Brasil, foram cultivados 21,58 milhões de hectares e produzidos 113,13 milhões de toneladas de milho na safra 2021/22, com produtividade de 5.242 kg ha^{-1} , sendo as maiores áreas observadas nos estados do Mato Grosso, Paraná e Mato Grosso do Sul, os quais foram responsáveis pela produção de 41,62, 16,42 e 12,46 milhões de toneladas, respectivamente (CONAB, 2022). Segundo Kipkulei et al. (2022), a produção da cultura de milho é influenciada pela variabilidade climática, fertilidade do solo, tratos culturais, data de semeadura, manejo de irrigação, fatores bióticos, entre outros.

As projeções dos desafios globais para o planeta nos próximos 50 anos mostram que os principais problemas para a humanidade serão aspectos relacionados à água, energia e alimentos. Nesse contexto da crescente demanda por alimentos, torna-se necessária a busca por ferramentas capazes de otimizar cada vez mais os processos agrícolas, possibilitando que a tomada de decisão seja pautada em informações seguras (OLIVEIRA et al., 2018).

Dentre essas ferramentas, destacam-se modelos capazes de simular a interação solo-planta-atmosfera e auxiliar na previsão da safra de determinada cultura. Esses modelos utilizam variáveis climáticas da região, estado nutricional e sanitário das plantas, características de solo e tratos culturais como dados de entrada para estimar o rendimento da produção e fornecer orientações para o gerenciamento agrícola, como o agendamento de irrigação e adubação (MACCARTHY et al., 2017; MALIK et al., 2019; FU et al., 2020; XIAO et al., 2020).

Um desses modelos é o DSSAT (*Decision Support System for Agrotechnology Transfer*, DSSAT version 4.6, JONES et al., 2003; HOOGENBOOM et al., 2014), o qual é um dos principais sistemas que empregam modelos de simulação de crescimento e desenvolvimento de culturas ao longo do tempo, sendo composto por ferramentas de suporte que auxiliam na manipulação de dados de solo, clima, condições experimentais e informações sobre o genótipo. O DSSAT inclui o CERES-MAIZE (*Crop Environment Resource Synthesis*), específico para simular o rendimento da cultura de milho sob condições ambientais distintas.

O modelo vem sendo amplamente utilizado para atender a uma variedade de aplicações, como a otimização de práticas agronômicas (e.g., manejo racional da irrigação) para intensificar a produção (Kadiyala et al., 2015; Mubeen et al., 2016; Oliveira et al., 2018), avaliação do desempenho de várias cultivares em diferentes condições ambientais (Chisanga et al., 2021; Feleke et al., 2021), análise dos impactos das mudanças climáticas na produção agrícola (Jiang et al., 2021; Oliveira et al., 2018) e avaliação de estratégias para lidar com condições limitadas de solo e clima (Malik e Dechmi, 2019).

No entanto, tais modelos geralmente requerem elevada quantidade de parâmetros de entrada, o que torna mais difícil a sua calibração precisa (XI et al., 2017). Por isso, diferentes abordagens baseadas no aprendizado de máquina para estimar o crescimento e desenvolvimento de culturas utilizando menos dados de entrada surgiram como alternativas promissoras e ferramentas complementares para a modelagem comumente usada (BASSO e LIU, 2019; CRANE-DOESCH, 2018; MISHRA et al., 2016; MUPANGWA et al., 2020).

Ao contrário dos modelos de simulação de culturas, o *machine learning* inclui métodos nos quais o sistema “aprende” uma função de transferência para prever a saída desejada com base nas entradas fornecidas, em vez de o pesquisador fornecer a função de transferência (SHAHHOSSEINI et al., 2021).

Embora geralmente essas abordagens sejam utilizadas separadamente, estudos recentes estão combinando-as para obter uma melhor estimativa da produção, onde as saídas dos modelos de crescimento de cultura estão sendo utilizadas como dados de entrada para alimentação de um modelo de aprendizado de máquina, como observado em Everingham et al. (2016), Pagani et al. (2016) e Feng et al. (2019).

Em uma pesquisa recente, Shahhosseini et al. (2021) projetaram uma abordagem híbrida (modelagem da colheita + ML) a partir de simulações no modelo APSIM (*Agricultural Production Systems sIMulator*) com o objetivo de melhorar a estimativa do rendimento de milho na região chamada de cinturão do milho dos Estados Unidos (*US Maize Belts*) localizada na região central do país. Os autores analisaram que adicionar as variáveis simuladas pelo APSIM como dados de entrada em modelos de ML pode diminuir o RMSE de 7 a 20%.

Com base na importância da estimativa da produção para o planejamento agrônomo, a hipótese da presente pesquisa é que os dados de saída dos modelos de crescimento de cultura podem alimentar algoritmos de *machine learning* para construção de modelos preditivos do rendimento agrícola, tornando o processo de determinação mais simples e menos oneroso.

Uma vez construídos e alimentados esses algoritmos, a segunda hipótese desse estudo é que algoritmos de aprendizado de máquina podem tornar-se uma alternativa aos modelos clássicos de estimativa do crescimento e produção de culturas.

De modo a testar as hipóteses citadas, os objetivos deste estudo foram: i) simular e avaliar a produção de milho em três datas de semeadura na cidade de Piracicaba, SP, para um período de 30 anos; ii) usar os dados simulados e variáveis climáticas para treino e teste de algoritmos de aprendizado de máquina para estimar a produção de milho; e iii) avaliar a performance dos modelos de ML na estimativa do rendimento agrícola.

4.2. Material e Métodos

A presente pesquisa seguiu as etapas demonstradas pela estrutura conceitual (Figura 8).

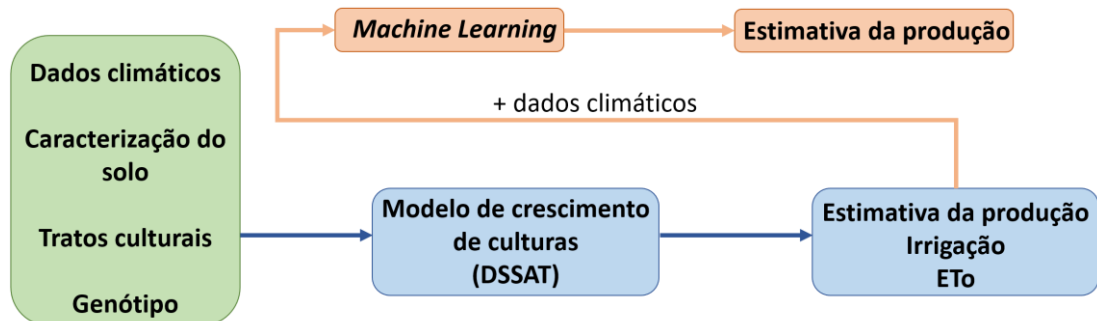


Figura 8. Estrutura conceitual do objetivo desse estudo.

4.2.1. Descrição sobre o modelo CERES-MAIZE

O modelo CERES-MAIZE, parte do sistema modular denominado *Crop System Model (CSM)* da plataforma DSSAT – *Decision Support System for Agrotechnology Transfer*, simula dinamicamente o desenvolvimento do sistema radicular, o crescimento e senescência das folhas e caule, acumulação de biomassa e o crescimento e a produção do milho, em função das propriedades físicas e química do solo, das condições climáticas, efeitos do genótipo e práticas de manejo. Também é caracterizado por considerar processos de balanço hídrico no solo, produção de matéria seca, interceptação da

radiação pelo dossel e partição de biomassa (JONES et al., 2003; HOOGENBOOM et al., 2010; HOOGENBOOM et al., 2014).

Os dados de entrada requeridos para simulação no modelo DSSAT/CERES-MAIZE incluem informações sobre a caracterização do solo, dados diários meteorológicos, condições iniciais de água e nitrogênio e manejo da cultura, como data e profundidade de semeadura, cultivar, população de plantas e aplicações de fertilizantes e irrigação (CARDOSO, FARIA e FOLEGATTI, 2004).

4.2.2. Propriedades físicas do solo

Para a condução da pesquisa foi utilizado um Latossolo Vermelho Amarelo (LVA), fase arenosa, série “Sertãozinho”, localizado em área experimental pertencente ao Departamento de Engenharia de Biosistemas LEB/ESALQ/USP, na cidade de Piracicaba, SP.

Os dados de entrada do solo relevantes ao modelo são: classificação, granulometria (teores de argila, silte e areia), coloração do solo, drenagem, profundidade, declividade, CTC (capacidade de troca de cátions) e pH em água.

Na Tabela 13 pode-se conferir os dados de entrada utilizados no DSSAT referentes à caracterização do solo.

Tabela 13. Caracterização do solo utilizado na pesquisa: Latossolo Vermelho Amarelo.

Profundidade cm	Horizonte	Argila	Silte %	Areia	CTC cmolc kg ⁻¹	Densidade do solo kg dm ⁻³
0 – 10	Ap	17,1	7,8	75,1	2,5	1,53
10 -20	AP	17,5	8,0	74,5	2,5	1,50
20 -30	AB	17,5	8,0	74,5	2,3	1,50
30 - 80	Bw	17,0	8,6	74,4	2,1	1,69

O Latossolo Vermelho Amarelo apresenta drenagem moderadamente boa, declividade baixa e escoamento superficial moderadamente baixo, garantindo baixo risco de erosão e baixas perdas significativas de nutrientes na percolação. Como é um solo caracterizado por ser profundo e poroso, apresenta condições adequadas para o bom desenvolvimento radicular em profundidade (EMBRAPA, 2006).

4.2.3. Dados meteorológicos

Os dados meteorológicos diários de temperatura máxima e mínima do ar (°C), radiação solar global (MJ m⁻² dia⁻¹), velocidade do vento (m s⁻¹), precipitação (mm) e umidade relativa do ar

(%) para a cidade de Piracicaba (latitude 22° 43' 30" e longitude 47° 38' 01") foram obtidos diretamente do site do posto meteorológico da ESALQ/USP, para o período de 01/01/1992 a 31/12/2021 (série de 30 anos).

Pela Figura 9 pode-se conferir as médias mensais das variáveis coletadas para a normal climatológica de 1992 a 2021.

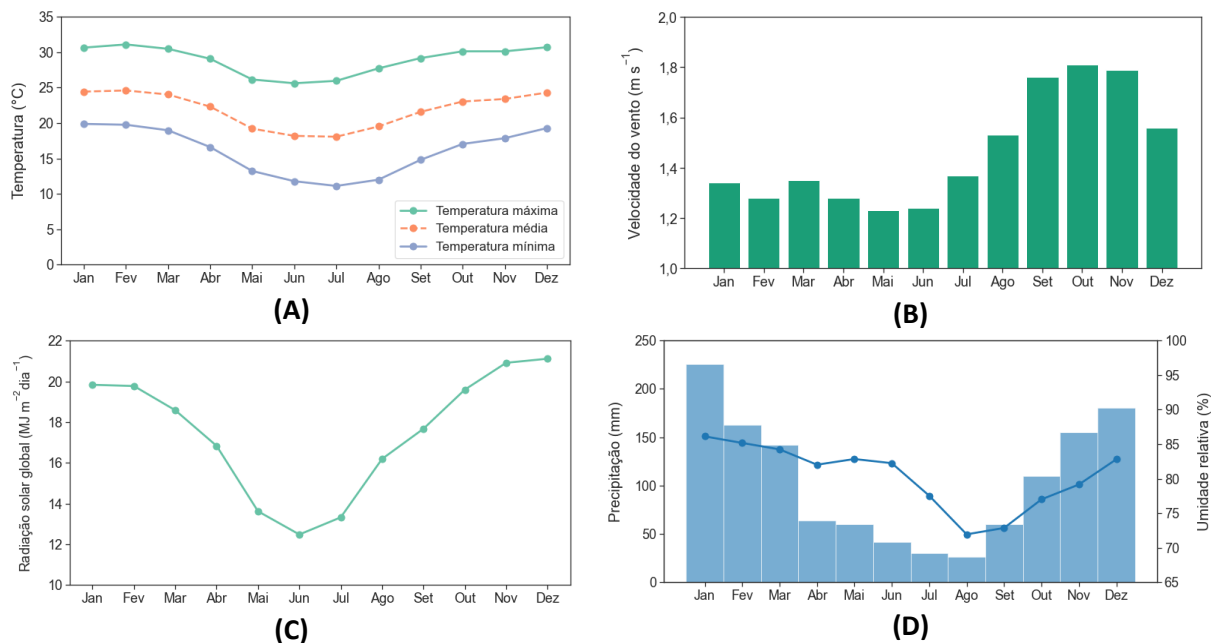


Figura 9. Médias mensais da temperatura máxima, média e mínima (A), velocidade do vento (B), radiação solar global (C), precipitação e umidade relativa do ar (D) referentes ao período entre 1992 – 2021 (30 anos).

O clima da região é Aw, segundo a classificação climática de Köppen (ALVARES, SENTELHAS e DIAS, 2022), no qual a temperatura do mês mais quente (fevereiro) é de 24,6 °C e do mês mais frio (julho) de 18,1 °C, com média anual de 21,9 °C. Em relação à precipitação, o total de chuvas para o mês mais chuvoso (janeiro) é de 225,7 mm e para o mês mais seco (agosto) de 26,3 mm.

Anualmente, chove cerca de 1257,8 mm na região de Piracicaba. A umidade relativa média anual é de 86,2%, enquanto a velocidade do vento média no ano é de 1,46 m s⁻¹. Durante os meses de maio, junho e julho, a região recebe os menores índices de radiação solar global, afetando diretamente a atividade fotossintética das plantas (13,62 MJ m⁻² dia⁻¹, 12,48 MJ m⁻² dia⁻¹ e 13,32 MJ m⁻² dia⁻¹, respectivamente).

4.2.4. Dados do genótipo

Para simular o crescimento e desenvolvimento do milho, o modelo CERES-MAIZE necessita ter seus coeficientes genéticos calibrados. Os coeficientes são: P1 – soma térmica desde a

emergência até ao final do período juvenil; P2 – coeficiente de sensibilidade ao fotoperíodo; e P5 – graus-dia acumulados desde o florescimento até o ponto de maturidade fisiológica; definem a fenologia da cultura, enquanto que G2 – número potencial de grãos por planta; e G3 – taxa potencial de crescimento do grão; se relacionam ao rendimento dos grãos. O coeficiente PHINT refere-se à quantidade de graus-dia necessários para emissão de uma folha.

Esses parâmetros podem ser obtidos mediante parametrizações acompanhando experimentos de campo e análises biométricas. No caso dessa pesquisa, os coeficientes foram obtidos na literatura, no qual Bender (2017) calibrou os parâmetros genéticos para uma cultivar definida como “média” de milho na 1ª safra e na 2ª safra, visando a calibração e o ajuste do modelo para uma cultivar representativa para todo o território brasileiro (Tabela 14).

Tabela 14. Coeficientes genéticos calibrados para uma cultivar média (1ª e 2ª safra) (BENDER, 2017).

Safra	Coeficientes Genéticos					
	P1	P2	P5	G2	G3	PHINT
1ª safra	290,60	0,50	907,10	749,90	9,40	46,79
2ª safra	285,20	0,50	914,70	857,00	6,07	44,10

4.2.5. Outras informações simuladas

Foram simuladas três datas de semeadura: 1º de fevereiro (milho 2ª safra), 26 de maio e 1º de outubro (milho 1ª safra), com o objetivo de avaliar a influência da época de plantio na produtividade final do milho. A população de plantas foi definida como 12 plantas m⁻², com espaçamento entre linha de 60 cm e profundidade de semeadura de 5 cm com sementes secas distribuídas em linhas.

Nas simulações com o modelo DSSAT/CERES-MAIZE, a evapotranspiração de referência (ET_o) foi estimada pelo método de Penman-Monteith FAO-56 (ALLEN et al., 1998). A fotossíntese foi determinada em função da eficiência da radiação solar global. No cálculo do balanço hídrico foi utilizado o modelo de Ritchie (1998), enquanto que a evaporação de água do solo foi estimada de acordo com o método de Suleiman-Ritchie (2003). Por fim, a infiltração de água no solo seguiu o método do número da curva de escoamento superficial, definido pelo serviço de conservação do solo do Departamento de Agricultura dos Estados Unidos (SOIL CONSERVATION SERVICE, 1972).

Para essa pesquisa foi simulado somente o balanço hídrico, enquanto que a influência de nitrogênio, fósforo, potássio e doenças não foi levada em consideração, portanto parte-se do princípio de que a planta não sofrerá estresse por déficit nutricional ou doenças.

A irrigação foi definida como automática quando requerida, ou seja, é a irrigação necessária para que o solo se mantenha sempre acima de 60% da CAD (Capacidade de Água Disponível do solo). A água foi aplicada utilizando-se o sistema de aspersão.

4.2.6. Criação do banco de dados, pré-processamento e desenvolvimento do modelo

Após realização das simulações no DSSAT, foram extraídos os dados de entrada (Tabela 15) necessários para treinamento e teste de algoritmos de *machine learning* supervisionado para predição da produtividade de milho (dado de saída).

As variáveis climáticas presentes na Tabela 15 (Tmax, Tmin, Rs, Precip, URmin e Uz) correspondem aos dados de entrada necessários para a simulação do crescimento e produção de milho no DSSAT.

Tabela 15. Dados gerados a partir das simulações no DSSAT utilizados para construção do banco de dados.

Variável	Tipo de variável	Descrição
Tmax	Contínua	Temperatura média máxima (°C), plantio – colheita
Tmin	Contínua	Temperatura média mínima (°C), plantio – colheita
Tmed	Contínua	Temperatura média (°C), plantio – colheita
Rs	Contínua	Radiação média solar global ($\text{MJ m}^{-2} \text{ dia}^{-1}$), plantio – colheita
Precip	Contínua	Precipitação total (mm), plantio – colheita
URmin	Contínua	Umidade média relativa mínima (%), plantio – colheita
Uz	Contínua	Velocidade média do vento (km h^{-1}), plantio – colheita
Irrig	Contínua	Irrigação total (mm), plantio – colheita
ETo	Contínua	Evapotranspiração total (mm), plantio – colheita
Data	Discreta	Data de semeadura
Prod	Contínua	Produtividade de grãos (kg ha^{-1}), colheita

O conjunto de dados construído apresentou 90 observações e 10 variáveis, uma vez que as simulações no DSSAT foram feitas para 30 anos e três datas de semeadura. A divisão do banco de dados seguiu a proporção de 75% para treino e 25% para teste.

Foram calculadas as correlações entre os dados de entrada com a variável-alvo (Prod), pelo coeficiente de correlação de Pearson (Eq. 14), com o objetivo de analisar se existe alguma relação entre as variáveis e entender como a variabilidade de uma afeta a outra. O resultado desse cálculo pode ser conferido a partir de um mapa de calor.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (14)$$

Em que n é o número de observações da amostra, \bar{x} é a média de x e \bar{y} é a média de y .

Os dados obtidos contêm características de dimensões e escalas distintas, o que afeta negativamente a modelagem e pode levar a um resultado tendencioso das previsões, uma vez que as variáveis que são medidas em escalas diferentes não contribuem igualmente para o ajuste e aprendizado do modelo e podem acabar criando um viés. Para lidar com esse problema, a padronização a partir do método *StandardScaler* da biblioteca Scikit-learn se apresenta como uma técnica de dimensionamento que torna os dados livres de escala, na qual a ideia principal é normalizar ($\mu = 0$ e $\sigma = 1$), individualmente, cada variável antes de aplicar o modelo de aprendizado de máquina, redimensionando a distribuição de valores para que a média dos valores observados seja 0 e o desvio-padrão 1. A padronização foi realizada transformando os dados de acordo com a Eq. 15.

$$z = \frac{x - \mu}{\sigma} \quad (15)$$

Em que μ é a média e σ o desvio-padrão.

Após o desenvolvimento dos modelos, os dados também foram submetidos a uma validação cruzada (*cross validation*) com o objetivo de evitar o overfitting. Para isso, os dados brutos foram separados em 10 pacotes diferentes, onde 10% dos dados dentro de cada pacote foram guardados de modo a não serem utilizados nos momentos de treino e teste, permitindo assim que o modelo fosse validado com condições de variáveis de entrada nunca processadas pelos modelos.

A montagem do banco de dados, pré-processamento e o treinamento e teste dos algoritmos foram realizados com o uso da biblioteca Scikit-Learn, uma biblioteca de código aberto especializada em aprendizado de máquinas, da linguagem de programação Python versão 3.9, em ambiente de desenvolvimento Jupyter Notebooks.

4.2.7. Algoritmos aplicados

4.2.7.1. Random Forest

Random forest (RF) é um algoritmo de aprendizado supervisionado baseado na utilização de diversos algoritmos menores (árvores de decisão), o qual usa um conjunto de métodos (*ensemble methods*) para resolver problemas de regressão e classificação.

O algoritmo operou combinando métodos de aprendizado conjunto com a estrutura da árvore de decisão, a qual representa um conjunto de condições ou restrições que são organizadas hierarquicamente e aplicadas sucessivamente de uma raiz à uma folha da árvore, com o objetivo de construir aleatoriamente uma infinidade de árvores a partir do conjunto de dados e calcular a média dos resultados de cada árvore, o que geralmente leva à uma previsão forte. Logo, o *Random Forest* atua como um algoritmo estimador que agrega o resultado de muitas árvores de decisão e, em seguida, gera o resultado mais ideal. Um preditor de regressão de RF pode ser expresso como a Eq. 16:

$$\hat{f}_{\text{RF}}^C = \frac{1}{C} \sum_{i=1}^C T_i(x) \quad (16)$$

em que x representa a variável de entrada vetorizada, C é o número de árvores e $T_i(x)$ é uma única árvore de regressão construída com base em um subconjunto de variáveis de entrada e amostras obtidas por meio do *bootstrapping*.

Durante o processo de desenvolvimento de uma floresta aleatória, em cada nó, um pequeno grupo de variáveis de entrada foi selecionado aleatoriamente para realizar a divisão. No caso de M variáveis de entrada, um número $m < M$ é identificado de modo que em cada nó m variáveis são escolhidas aleatoriamente e a melhor divisão dessas m é utilizada para dividir o nó. O valor de m é mantido constante durante o crescimento da floresta. O critério de divisão da árvore de regressão é baseado na escolha da variável de entrada que apresenta o menor Índice de Gini (Eq. 17).

$$I_G(t_{X(x_i)}) = 1 - \sum_{j=1}^m f(t_{X(x_i)}, j)^2 \quad (17)$$

Em que $f(t_{X(x_i)}, j)$ é a proporção de amostras com o valor x_i pertencente à folha j com nó t .

Para cada árvore de decisão construída, um novo conjunto de treino foi gerado por meio do método *bootstrapping*, substituindo, então, os dados originais. Por isso, nem todas as amostras são selecionadas para o treinamento de uma árvore e alguns dados de treinamento podem ser usados repetidamente. As amostras que não são selecionadas são incluídas como parte de outro subconjunto denominado de observações *out-of-bag* (OBB), o qual, geralmente, representa um terço do conjunto total.

Pela Figura 10, pode-se observar a estrutura do algoritmo *Random Forest*.

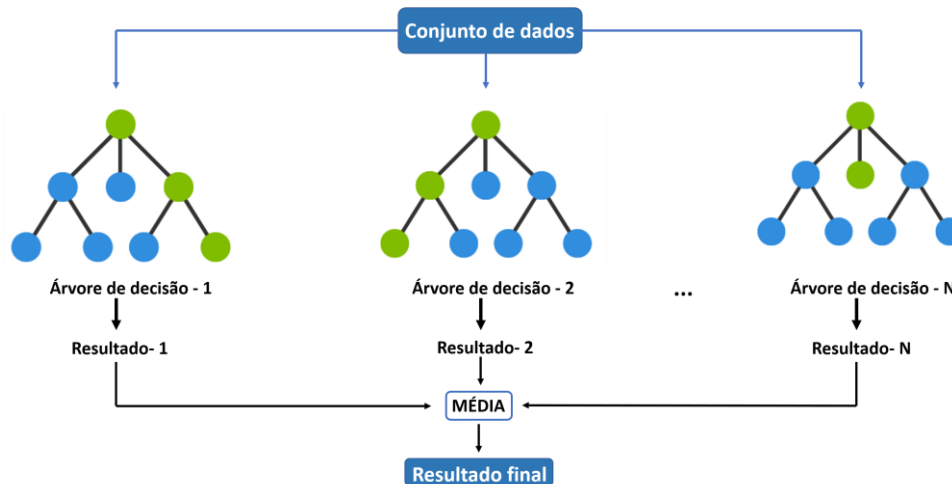


Figura 10. Estrutura do *Random Forest* (adaptado de Ahmad et al., 2018).

Para a presente pesquisa, foi selecionado 500 como o número de árvores a serem construídas usando amostras dos dados de entrada originais, a partir do método *bootstrapping*. Com isso, uma árvore de regressão foi criada para cada uma das amostras. Em cada nó, os preditores foram amostrados aleatoriamente e a melhor divisão dessas variáveis foi selecionada. Novos dados foram então previstos agregando as previsões.

4.2.7.1.1. Feature importance

A *feature importance* é um importante recurso do algoritmo *Random forest* utilizado para avaliar o quanto uma variável de entrada contribui, individualmente, no desempenho do modelo, mostrando qual característica (*feature*) possui maior influência estatística na determinação da variável-alvo. Portanto, ao treinar uma árvore, pode-se calcular o quanto cada recurso contribui para diminuir a impureza ponderada.

O envolvimento de múltiplas características no treinamento de algoritmos pode levar a uma redução significativa na eficiência do processamento, principalmente devido à redundância de informações. Para resolver esses problemas e reduzir a dimensionalidade dos dados, o método da *feature importance* foi implementado por Li et al. (2019).

A importância de uma característica é calculada como a diminuição na impureza do nó ponderada pela probabilidade de atingir esse nó. A probabilidade do nó pode ser calculada pelo número de amostras que chegam ao nó, dividido pelo número total de amostras.

O processo no *Random forest* para realizar a *feature importance* seguiu as seguintes etapas: primeiro, uma árvore de regressão foi construída para gerar as denominadas observações *out-of-bag* (OOB). Em seguida, com base nos dados OOB, o algoritmo consegue investigar a relevância dos

dados de entrada e obter uma pontuação, a qual é expressa pela medida chamada de *mean decrease accuracy* (MDA) ou *mean decrease in impurity* (MDI), sendo que quanto maior o valor da MDA, maior a importância da variável no modelo. O MDI calcula para cada variável a diminuição média na impureza que introduziu em todas as árvores de decisão ao construí-las.

O princípio é transformar o valor de um parâmetro da característica em um número aleatório, calcular seu impacto na precisão do modelo e medir a importância desse parâmetro com base no valor obtido da MDA a partir de vários cálculos (GENUER et al., 2010).

4.2.7.2. Extra-Trees Regressor

O algoritmo de aprendizado de máquina supervisionado *Extra-Trees* (*Extremely Randomized Trees*) pode ser caracterizado como uma extensão altamente aleatória do modelo *Random Forest*, seu principal diferencial é que esse algoritmo é caracterizado por ser extremamente aleatório.

O processo de desenvolvimento do *Extra-Trees* (EXT) é similar ao descrito para o RF, sendo que a diferença entre os dois está na quantidade de processos aleatórios utilizados em cada algoritmo. As etapas descritas para o RF, também ocorreram no EXT, porém, no último, um novo fator de aleatoriedade foi adicionado.

Ao contrário do RF, as EXT não utilizam a etapa de *bagging* (agregação de *bootstrap*) de árvores para gerar o subconjunto de treinamento para cada árvore. Todo o conjunto de treinamento é usado para treinar todas as árvores de decisão do conjunto. Portanto, no EXT a aleatoriedade não vem da amostragem por *bootstrap*, mas sim das divisões aleatórias dos dados.

Depois de selecionar aleatoriamente as variáveis candidatas para o primeiro nó, os dados existentes em cada uma dessas variáveis foram separados em duas partes baseando-se em um valor ideal/limite (*thresholds*), trazendo maior precisão ao modelo. Para encontrar esse valor de divisão, o algoritmo define alguns pontos de forma a dividir o conjunto em todas as combinações possíveis. Em seguida, o melhor valor correspondente para dividir cada nó foi aleatoriamente selecionado.

A escolha de um valor de divisão feita pelo EXT foi realizada de forma randômica, enquanto o RF utilizou apenas a definição de variáveis e amostras como etapas randômicas, partindo para otimização dos cálculos em busca da melhor divisão. Com cada variável com o seu valor ideal para a divisão definido, o algoritmo tentou, então, escolher, usando o Índice de Gini, a melhor dessas variáveis para o nó inicial.

4.2.8. Avaliação da performance dos modelos

Para avaliar o desempenho dos modelos nos momentos de treino, teste e validação, foram utilizados parâmetros estatísticos: R^2 , empregado para medir a variação entre os valores preditos pelo modelo e os valores reais (Eq. 18) e erro quadrado médio (RMSE), utilizado para determinação da diferença entre os valores preditos e reais de ETo. (Eq. 19).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (18)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (19)$$

Em que, y_i é o valor observado, \hat{y}_i o valor predito, \bar{y} o valor médio observado e n o número de observações.

4.3. Resultados e Discussão

4.3.1. Produtividade de milho simulada pelo DSSAT

Pela Tabela 16 pode-se conferir os valores de produtividade (kg ha^{-1}) de grãos de milho para as condições simuladas pelo modelo DSSAT em três datas de semeadura diferentes na região de Piracicaba. A partir dos resultados obtidos, é possível analisar que diferentes épocas de plantio influenciaram no rendimento de grãos de milho.

Valores semelhantes de rendimento de grãos podem ser encontrados em Oliveira (2015), no qual o DSSAT foi utilizado para estimativa da produtividade de milho em Piracicaba utilizando diferentes lâminas de irrigação, visando o manejo racional da água.

Tabela 16. Produtividade de grãos para as condições simuladas no modelo DSSAT do cultivo de milho em Piracicaba em três datas diferentes de semeadura.

Ano	Data da semeadura		
	01/fev	26/mai	01/out
	kg ha ⁻¹		
1992	4.280	542	6.376
1993	7.159	906	4.562
1994	4.708	482	6.989
1995	4.818	641	8.368
1996	5.589	1.047	9.013
1997	8.934	2.175	6.698
1998	10.764	5.088	8.280
1999	10.073	3.549	6.176
2000	7.820	2.261	9.218
2001	6.740	3.317	12.057
2002	8.522	1.852	8.055
2003	7.413	874	8.153
2004	10.112	2.428	10.009
2005	5.698	4.130	10.417
2006	10.103	3.231	9.682
2007	8.560	2.034	9.760
2008	9.389	3.576	7.181
2009	8.130	5.878	9.801
2010	9.506	3.858	9.110
2011	11.191	3.215	15.114
2012	10.975	6.959	11.983
2013	7.335	6.350	10.981
2014	8.670	3.373	5.689
2015	8.203	2.714	8.952
2016	7.872	1.948	10.972
2017	7.337	2.087	10.240
2018	8.477	1.412	9.861
2019	7.433	1.877	9.743
2020	8.468	1.067	8.973
2021	7.593	1.915	10.793
Média	8.062,40	2.692,87	9.106,87

Os dados simulados mostraram que a produtividade foi reduzida quando o milho foi semeado no final do mês de maio, com média de 2.692,87 kg ha⁻¹, sendo considerada muito baixa quando comparada com as obtidas nas semeaduras de fevereiro (8.062,40 kg ha⁻¹) e de outubro (9.106,87 kg ha⁻¹), épocas consideradas ideais para o início do cultivo.

A mudança da data de semeadura de maio para fevereiro resultou em um aumento na produtividade de aproximadamente 5.400 kg ha⁻¹ e a alteração de maio para outubro provocou um aumento de aproximadamente 6.415 kg ha⁻¹.

As produtividades inferiores observadas nos três tratamentos ao longo dos anos (e.g. 1992, 1993, 1994, 1997, 1999 e 2014 para o plantio em 01/out) estão relacionadas, principalmente, com

menores índices pluviométricos durante o ciclo, o que ocasiona uma diminuição na ETc, uma vez que o milho demanda um consumo mínimo de 350-500 mm para garantir um rendimento satisfatório, por isso o plantio em períodos chuvosos é recomendado.

Na simulação do modelo CERES-MAIZE, a taxa de crescimento do milho depende especialmente da temperatura, onde a temperatura mínima afeta a quantidade de biomassa que é destinada aos grãos. Para o rendimento do milho, o número de grãos por planta é uma função do número de grãos potenciais por planta e da taxa média de crescimento da cultura desde o espigamento até o enchimento dos grãos (DEJONGE et al., 2012).

Segundo Andrade (1995), apesar de apresentar um elevado potencial produtivo, o milho possui acentuada sensibilidade a estresses bióticos e abióticos, logo, a fim de buscar a sua máxima capacidade produtiva, o seu cultivo exige rigoroso planejamento e manejo.

Na Tabela 17 pode-se conferir as condições médias meteorológicas que o ciclo de cultivo apresentou quando semeado em cada uma das três datas em estudo. A partir dela, pode-se analisar que as produtividades apresentaram tendências de acréscimos com o aumento dos valores das variáveis climáticas, especialmente Tmin, Tmed, Rs e precipitação, o que pode ser observado nos plantios do milho safrinha e safra. Logo, rendimentos de colheita superiores estão diretamente relacionados com condições ambientais favoráveis para o melhor desenvolvimento e crescimento do milho.

De acordo com Fancelli e Dourado Neto (2000), os fatores climáticos de maior influência sobre o milho são a radiação solar, a precipitação e a temperatura, o que explica os resultados observados na presente pesquisa.

Vale ressaltar que por ser uma planta C4, o milho depende bastante da disponibilidade de energia, por isso maiores valores de radiação solar no período do cultivo resultam em maiores produtividades, uma vez que a capacidade da planta em produzir fitomassa seca está diretamente relacionada com a quantidade de energia luminosa disponível, isso pode ser percebido no ciclo do milho safra (plantado em outubro e colhido entre os meses de fevereiro e junho) (COSTA, 1997).

Tabela 17. Variáveis climáticas médias e produtividades durante o ciclo de cultivo do milho para as três datas de semeadura simuladas.

Data de semeadura	Tmax	Tmin	Tmed	Precip	Rs	Uz	URmin	Prod
		°C		mm	MJ m⁻² dia⁻¹	m s⁻¹	%	kg ha⁻¹
01/fev	28,70	16,40	22,55	456,32	16,63	1,28	54,42	8.062,40
26/mai	27,54	13,20	20,37	266,60	15,51	1,51	45,88	2.692,87
01/out	30,38	18,51	24,45	678,97	20,40	1,62	53,34	9.106,87

A partir da Figura 11 pode-se conferir a variabilidade da produtividade de milho simulada com base em dados climáticos entre 1992 e 2021, onde, novamente, pode-se destacar produtividades inferiores quando o ciclo de cultivo inicia-se no final de maio, período em que as temperaturas são mais baixas ($T_{min} = 13,20\text{ }^{\circ}\text{C}$), bem como o volume de chuva ($Precip = 266,60\text{ mm}$) e a radiação solar ($R_s = 15,51\text{ MJ m}^{-2}\text{ dia}^{-1}$), o que dificulta o ideal estabelecimento da cultura, uma vez que nesse momento as plantas apresentam alta sensibilidade à estresses e eventos extremos podem determinar toda a produção.

Portanto, os resultados obtidos sugerem que a escolha da época de plantio é de extrema relevância para atingir elevadas produtividades de milho, pois, desse modo, é possível evitar épocas com potenciais riscos climáticos, como períodos de veranicos e geadas. Além disso, pode-se diminuir o custo da produção devido à diminuição de gastos com irrigação, fertilizantes e herbicidas (RUGIRA et al., 2021).

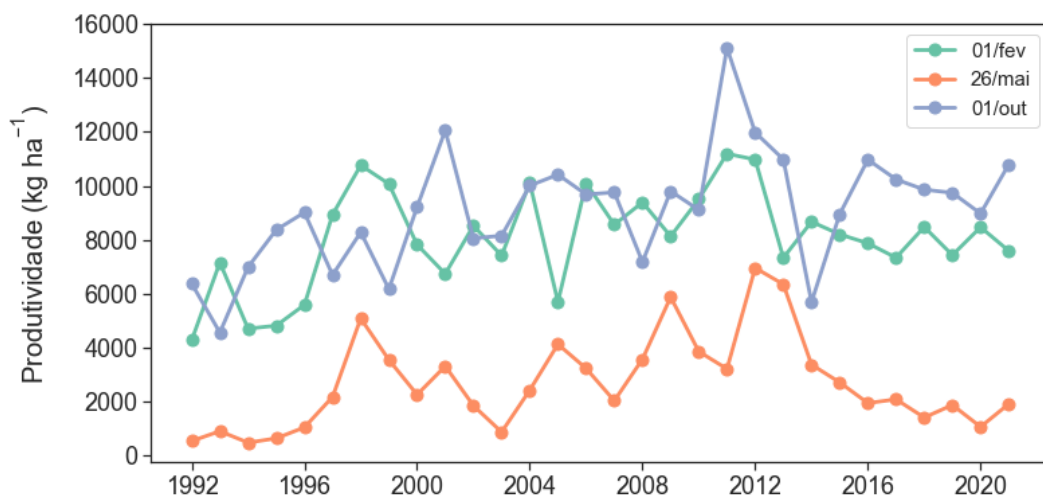


Figura 11. Produtividade de milho simulada pelo modelo DSSAT com base em dados climáticos entre 1992 e 2021 para a região de Piracicaba em três épocas diferentes de plantio.

A Figura 12 apresenta a disposição dos dados simulados, os quais foram utilizados como valores-alvo de referência, servindo como inputs no treinamento e teste para o desenvolvimento e avaliação dos modelos de *machine learning*. A partir dessa figura é possível observar que os dados seguem a distribuição normal.

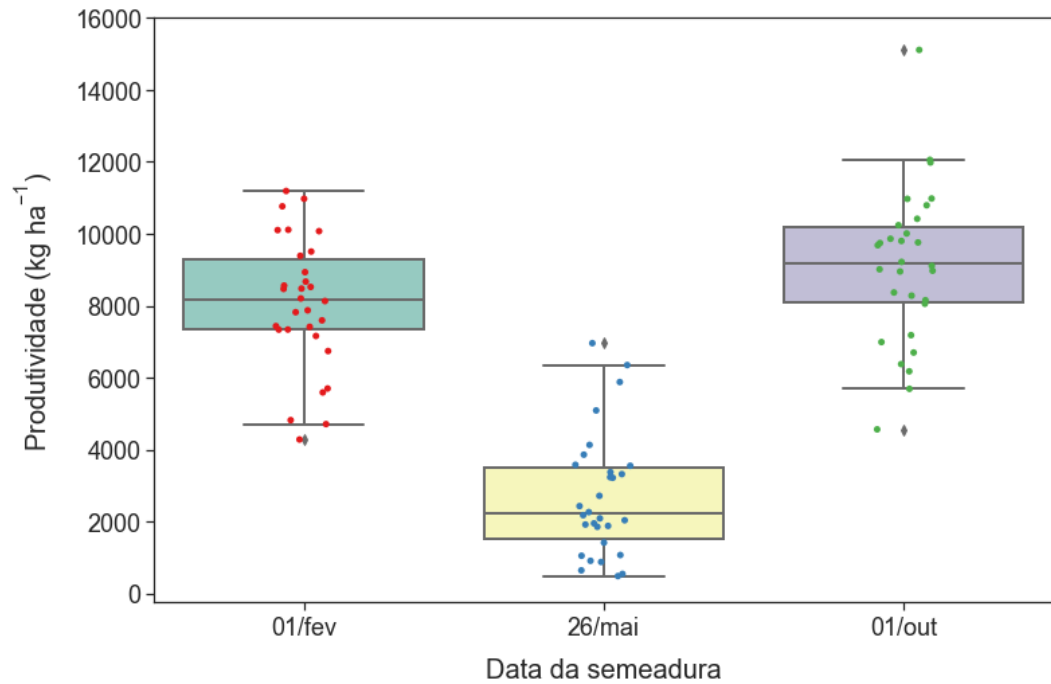


Figura 12. Boxplot das produtividades de milho simuladas pelo modelo DSSAT para as três épocas de semeadura em estudo.

4.3.2. Correlação entre as variáveis de entrada e produtividade

Pela Figura 13 pode-se conferir o mapa de calor referente ao banco de dados em estudo, o qual tem como objetivo analisar a relação entre as variáveis de entrada dos modelos e a produtividade de milho simulada pelo DSSAT, a fim de compreender quais dos elementos presentes mostrou maior e menor influência na produção a partir do cálculo da correlação entre as variáveis. A correlação fornece pistas importantes sobre a significância das variáveis preditoras.

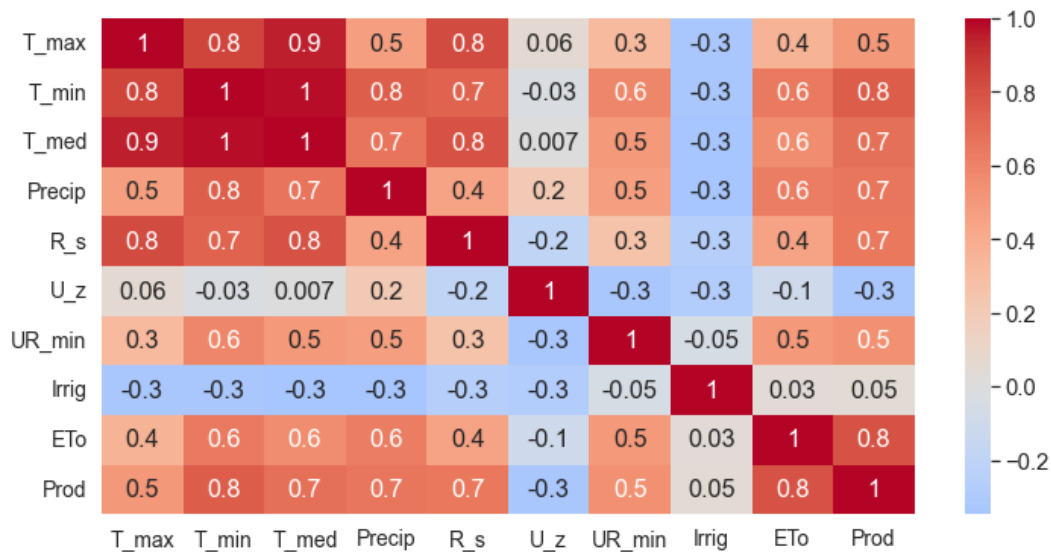


Figura 13. Mapa de calor apresentando a correlação entre as variáveis de entrada e a produtividade de grãos de milho simulada pelo DSSAT.

Os resultados conferidos na Figura 13 constataam que a maioria dos dados selecionados apresentaram correlação significativa com a produtividade de milho, com exceção da irrigação, a qual como foi utilizada somente para complementar a chuva para que o solo se mantivesse sempre acima de 60% da CAD, por isso os valores das lâminas foram baixos.

É possível analisar que as correlações mais significativas com o rendimento de grãos de milho, ou seja, aquelas mais próximas de -1 ou 1, estão relacionadas à ETo, temperaturas mínima e média, radiação solar e precipitação (iguais a 0,8, 0,8, 0,7, 0,7 e 0,7, respectivamente). Como já demonstrado no Capítulo 1 da presente pesquisa, a ETo, para as condições da região de Piracicaba, é fortemente correlacionada com as variáveis climáticas, especialmente, temperatura, radiação solar e umidade relativa, logo, por representar o clima como um todo, apresenta correlação significativa com a produtividade.

Segundo Fancelli e Dourado Neto (2000), a interação de tempo e condições climáticas a que uma cultura de milho é exposta determina a intensidade com que ela pode expressar seu potencial genético. Nesse contexto, os fatores climáticos que mais afetam o crescimento e a produção de milho são a radiação solar, a precipitação e a temperatura, corroborando com os resultados conferidos na Figura 13.

Como uma planta C4, o milho tem uma alta taxa fotossintética e responde ao aumento da intensidade da luz com altos rendimentos, uma vez que a capacidade da planta em produzir fitomassa seca está diretamente relacionada com a quantidade de energia luminosa disponível (COSTA, 1997). O número de grãos por planta é determinado por volta da época de espigamento, e estudos mostram a existência de uma forte relação entre o número de grãos (principal fator

determinante da produtividade) e a taxa de crescimento da planta, a qual é basicamente impulsionada pela fotossíntese (ANDRADE et al., 1999).

Como pode ser conferido na Figura 13, o rendimento do milho também é fortemente dependente da quantidade e frequência das chuvas, bem como a sua distribuição temporal e espacial, sendo que as plantas tendem a apresentar extrema sensibilidade ao déficit hídrico especialmente durante a floração até o início da fase de enchimento dos grãos, por isso a deficiência de água nessas etapas provoca elevada redução da produtividade.

A temperatura é um importante fator no adequado crescimento de milho. Pode-se observar que a temperatura máxima, para as condições em estudo, apresentou menor correlação com a produtividade do que as temperaturas média e máxima. Isso provavelmente se deve ao fato de que altas temperaturas aceleram a perda de água do solo, o que afeta negativamente o desenvolvimento do milho desde a germinação até o crescimento vegetativo e subsequente enchimento de grãos (OMOYO et al., 2015).

Além disso, de acordo com Bai et al. (2016), a elevação da temperatura máxima pode acelerar o desenvolvimento da cultura e encurtar a duração do ciclo de cultivo, bem como afetar as taxas de fotossíntese, respiração e enchimento de grãos.

Diferentemente da correlação, a qual corresponde à um termo estatístico referente ao quão próximas duas variáveis estão, em termos de ter uma relação linear uma com a outra, a *feature importance* é um recurso do algoritmo *Random forest* utilizado para avaliar o quanto uma variável de entrada contribui, individualmente, para o desempenho do modelo.

Nesse contexto, é possível analisar, pela Figura 14, que a temperatura mínima, para as condições em estudo, apresentou a maior contribuição estatística na determinação da variável-alvo, seguida pela evapotranspiração de referência, radiação solar global e temperatura média, sendo que estas variáveis também mostraram alta correlação com a produtividade, como observado na Figura 13.

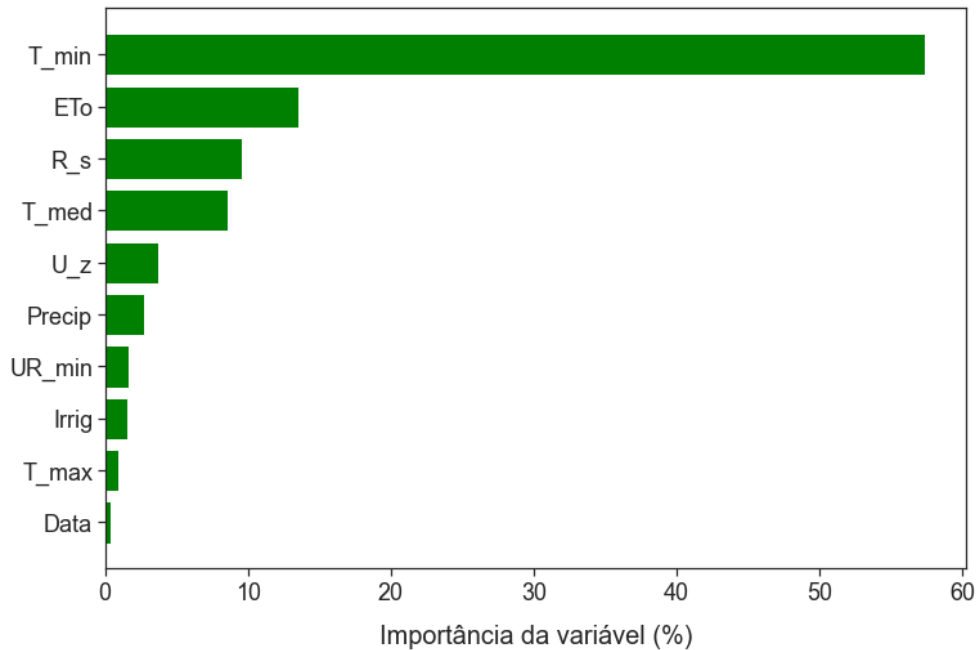


Figura 14. Contribuição de cada variável para o desempenho do modelo (*Random forest*) na determinação da produtividade de milho.

A elevada importância da temperatura mínima na estimativa da produtividade de milho pode ser explicada por essa cultura ser termofílica, a qual se beneficia de temperaturas mínimas mais altas, sobretudo nas fases de pré-floração e floração, sendo esse fator climático determinante no sucesso da produção (YIN et al., 2017).

Em estudo realizado por Fei et al. (2020), foi possível observar que o aumento da temperatura mínima em maio foi benéfico para a germinação e o crescimento vegetativo, enquanto que o seu aumento em setembro proporcionou condições mais ideais para o enchimento de grãos e evitou danos causados pela geada precoce. Nesse caso, a temperatura mínima foi classificada pelos autores como o principal fator responsável pelo aumento do potencial da produção de milho.

A elevada contribuição da ETo provavelmente se deve ao fato de essa variável representar o clima como um todo, uma vez que está altamente correlacionada com as variáveis climáticas, principalmente temperatura, radiação solar e umidade relativa do ar.

A data de semeadura apresentou baixa relevância para estimativa do rendimento de grãos, o que não está de acordo com a evidente importância que a escolha da época de plantio possui para alcançar elevadas produtividades. Isso, provavelmente, se deve ao fato de que somente três datas diferentes foram simuladas, assim, a baixa variabilidade de dados, nesse caso, se reflete diretamente no momento do treino dos modelos que não possuem dados em quantidade suficiente para criar uma relação mais precisa entre essa variável de entrada e a variável-resposta (produtividade).

Portanto, devido às correlações significativas, pode-se dizer que a utilização de modelos eficientes para obtenção da estimativa de safra está na dependência do conhecimento da relação entre as demandas dessa cultura e as condições climáticas do ambiente em que ela está condicionada. Por isso, os elementos do clima podem ser utilizados para ajuste de modelos de simulação de desenvolvimento e crescimento de culturas (RUGIRA et al., 2021).

4.3.3. Avaliação dos algoritmos

Os resultados presentes na Tabela 18 confirmaram que a utilização de algoritmos de *machine learning* com variáveis climáticas como dados de entrada para a estimativa da produtividade de milho na região de Piracicaba apresentou efeito positivo e significativo, se estabelecendo como uma alternativa viável e prática ao modelo DSSAT, a qual pode contribuir nas tomadas de decisões em relação ao planejamento e gestão da safra.

Tabela 18. Performance comparativa dos algoritmos nos períodos de treino, teste e validação referentes à estimativa da produtividade de milho em Piracicaba.

Modelo	Treino		Teste		Validação	
	R ²	RMSE (kg ha ⁻¹)	R ²	RMSE (kg ha ⁻¹)	R ²	RMSE (kg ha ⁻¹)
EXT	0,973	554,64	0,836	1.381,02	0,827	1.461,34
RF	0,971	566,94	0,791	1.560,69	0,788	1.613,63

A confiabilidade dos modelos de regressão em fazer previsões do rendimento de grãos foi garantida pelos altos valores de R², os quais indicaram que o EXT e o RF podem fazer estimativas para novos dados com uma precisão de 82,7% e 78,8%, respectivamente, por meio de uma combinação linear de variáveis relacionadas às condições climáticas. O melhor desempenho do EXT em comparação com o RF pode ser explicado, possivelmente, devido ao seu alto fator de aleatoriedade.

Os valores de RMSE maiores que 1.000 kg ha⁻¹ observados na presente pesquisa durante as fases de teste e validação, também foram identificados em outros estudos em que algoritmos de aprendizado de máquina foram utilizados para predição da produtividade de milho, como relatado por Kang et al. (2020) e Meng et al. (2021).

O modelo CNN-RNN (*Convolutional Neural Network – Recurrent Neural Network*) desenvolvido por Khaki et al. (2020) para prever o rendimento de grãos na região do cinturão do milho norte-americano, apresentou erros de previsão de 988 kg ha⁻¹ e 1.107 kg ha⁻¹. Nessa mesma pesquisa, os autores também testaram o uso do RF e encontraram erros de 943 kg ha⁻¹ e 1.813 kg ha⁻¹ durante as fases de treino e validação, respectivamente. Shahhosseini et al. (2021), utilizando o RF

para estimar o rendimento da safra no centro-oeste dos Estados Unidos, observaram RMSE de 1.024 kg ha⁻¹ e precisão de 76,7% na etapa de teste.

De acordo com Zhang e Ma (2012), os modelos caracterizados por serem *ensemble models*, como o RF e EXT, foram desenvolvidos com o objetivo de capturar relações complexas nos dados de entrada. Foi provado que a combinação bem diversificada de estimadores do ML pode resultar em um modelo de desempenho superior. Isso explica as performances satisfatórias dos dois algoritmos aqui estudados em estimar a produtividade de milho com base apenas nos dados climáticos.

Por dependerem apenas de variáveis climáticas e apresentarem desempenho consistente, a utilização de algoritmos como o RF e EXT tornam-se uma alternativa ao uso de modelos clássicos de crescimento e produção de culturas, os quais exigem um grande volume de dados de entrada em relação ao solo e à cultura. Uma outra alternativa sugerida seria utilizar ambas as metodologias de forma híbrida, onde Shahhosseini et al. (2021) demonstraram que a introdução de variáveis simuladas pelo modelo APSIM em algoritmos de aprendizado de máquina como dados de entradas para uma tarefa de previsão, pode reduzir o erro de previsão em 7 e 20%.

Na Tabela 19 pode-se conferir os valores médios de produtividade de grãos de milho (kg ha⁻¹) preditos pelos algoritmos EXT e o RF e simulados pelo DSSAT em três datas de semeadura diferentes na região de Piracicaba para o período entre 1992 e 2021. É possível observar que as estimativas feitas pelos algoritmos se aproximaram das simuladas pelo modelo DSSAT e seguiram a tendência de estimar produtividades elevadas para o milho semeado em fevereiro e outubro, épocas em que as condições ambientais se mostraram mais favoráveis para o cultivo.

Tabela 19. Produtividades médias, para o período de 30 anos, preditas pelos algoritmos de aprendizado de máquina em comparação com as médias simuladas pelo modelo DSSAT.

Modelo	Data da semeadura		
	01/fev	26/mai	01/out
	kg ha ⁻¹		
EXT	7.704,91	2.961,04	8.904,84
DSSAT	8.062,40	2.692,87	9.106,87
RF	7.787,84	2.911,49	8.977,81

Uma comparação mais abrangente entre os valores das produtividades estimadas por cada algoritmo e os valores simulados pelo modelo DSSAT pode ser conferida pela Figura 15, na qual pode-se verificar o comportamento da estimativa do rendimento de grãos com base em dados climáticos ao longo dos 30 anos em estudo. É possível analisar que os algoritmos seguiram o padrão simulado pelo DSSAT e, portanto, foram capazes de identificar relações lineares e não lineares entre a produtividade e as condições climáticas. Pela Figura 15 também pode-se perceber que os algoritmos obtiveram desempenhos semelhantes ao longo dos 30 anos simulados.

A Figura 15A mostra a redução da precisão dos modelos em anos como 1993, 1998, 1999, 2011 e 2012, especialmente quando foram simuladas produtividades acima do valor médio (8.062,40 kg ha⁻¹). É possível observar uma tendência de subestimativa quando a produção de milho aparenta ser superior à média. Já pela Figura 15B, pode-se conferir que em apenas um ano o desempenho dos algoritmos não foram satisfatórios para a semeadura em maio, onde a produtividade foi superestimada em 2015.

Por último, pode-se analisar pela Figura 15C que para o plantio em outubro ocorreram mais erros de estimativa, como observados em 1993, 2001, 2003, 2011, 2012 e 2014. Esses erros chegaram a ser maiores que 2.000 kg ha⁻¹, impactando significativamente no desempenho geral do modelo.

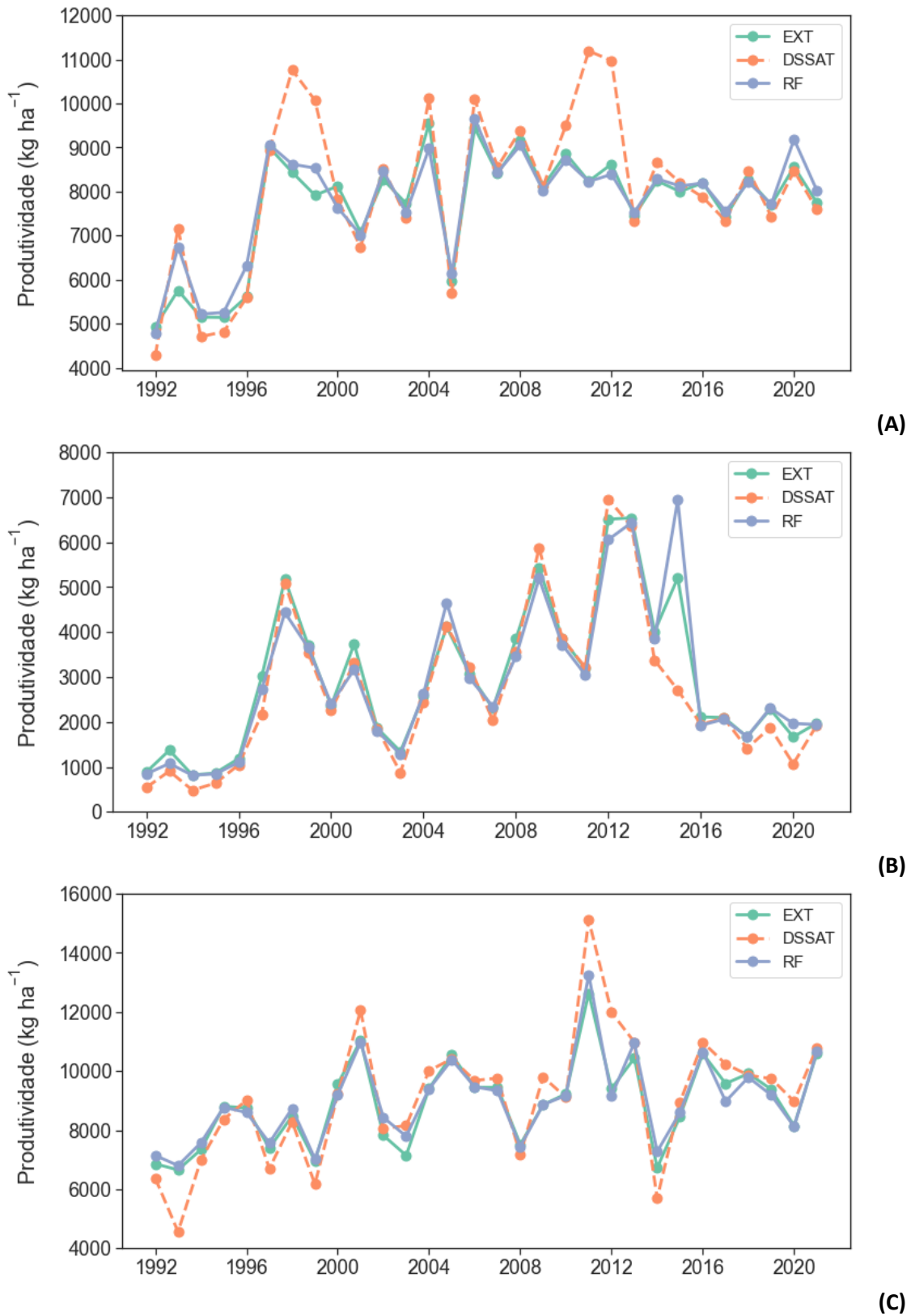


Figura 15. Produtividades de milho estimadas pelos algoritmos EXT e RF em comparação com as simuladas pelo modelo DSSAT para a região de Piracicaba durante o período entre 1992-2021 para três épocas de semeadura diferentes: 01/fev (A), 26/mai (B) e 01/out (C).

Logo, é possível observar que alguns anos tiveram maior diferença entre os valores simulados pelo DSSAT e os preditos pelos algoritmos do que outros. Isso pode ocorrer principalmente devido à influência de fatores externos, como a capacidade de água disponível no solo em relação aos estádios fenológicos da cultura e outros aspectos relacionados ao solo, ou seja, variáveis que não foram consideradas na análise, o que pode ser confirmado quando é verificado que não existiram eventos climáticos extremos nesses anos que justificassem tamanho erro de previsão, nesse caso, isso significaria que os algoritmos não seriam capazes de identificar como mudanças extremas no clima impactariam na produção.

As estimativas foram menos consistentes para a semeadura em maio, época em que ocorrem produtividades baixas, corroborando com os maiores erros observados quando a produtividade simulada pelo DSSAT é maior do que o rendimento médio da região.

Por isso, mesmo com os altos valores de R^2 apresentados pela Tabela 19, pode ser concluído que a inclusão de dados do solo como variáveis de entrada pode contribuir para maior estabilidade das estimativas dos modelos.

Portanto, a presente pesquisa procurou propor uma estrutura para previsão da produtividade de milho a partir da construção de um modelo abrangente impulsionado pelas condições climáticas. Os resultados aqui obtidos contribuem para uma ampla compreensão da resposta do rendimento do milho a diversas variáveis ambientais.

4.4. Conclusões

Diante dos resultados obtidos pode-se concluir que:

- a. As simulações do modelo DSSAT permitiram observar que o plantio de milho no final do mês de maio resultou em uma redução da produtividade quando comparado ao plantio em fevereiro e outubro, os quais representam o cultivo do milho safrinha e safra, respectivamente. Logo, rendimentos de colheita superiores estão diretamente relacionados com condições ambientais favoráveis em que as plantas são submetidas;
- b. As variáveis meteorológicas obtidas para a região de Piracicaba apresentaram correlações significativas com o rendimento de grãos de milho. Portanto, pode-se afirmar que a produtividade da cultura depende da relação intrínseca estabelecida entre a planta e as condições ambientais em que ela está condicionada, especialmente a disponibilidade hídrica, temperatura e radiação solar. Assim, dados climáticos podem ser utilizados para estimar a produção; e

- c. As estimativas consistentes indicaram que os algoritmos aplicados foram capazes de identificar relações lineares e não lineares entre o rendimento de grãos de milho e dados meteorológicos, resultando em desempenho satisfatório da predição da produtividade com base apenas em variáveis climáticas.

Portanto, pode-se concluir que os algoritmos de aprendizado de máquina adotados apresentaram-se como sendo ferramentas de alto potencial de utilização, e mostraram-se como alternativa aos modelos clássicos de crescimento e produção de culturas, para a predição da produtividade de milho, nas condições climáticas de Piracicaba, podendo auxiliar, de maneira eficiente, os tomadores de decisão.

Referências

- AHMAD, M. W.; MOURSHED, M.; REZGUI, Y. Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression. **Energy**, v. 164, p. 465–474, 2018.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. 1998. **Crop evapotranspiration: guidelines for computing crop water requirements**. Rome: FAO, 1998, 300 p. (FAO irrigation and drainage paper no. 56).
- ALVARES, C. A.; SENTELHAS, P. C.; DIAS, H. B. Southeastern Brazil inland tropicalization: Köppen system applied for detecting climate change throughout 100 years of meteorological observed data. **Theoretical and Applied Climatology**, v. 149, p. 1431–1450, 2022.
- ANDRADE, F. H. Analysis of growth and yield of maize, sunflower and soybean grown at Balance, Argentina. *Fields Crops Research*, v. 41, p. 1–12, 1995.
- ANDRADE, F. H.; VEGA, C.; UHART, S.; CIRILO, A.; CANTARERO, M.; VALENTINUZ, O. Kernel number determination in maize. **Crop Science**, v. 39, p. 453–459, 1999.
- BAI, H.; TAO, F.; XIAO, D.; LIU, H.; ZHANG, H. Attribution of yield change for rice-wheat rotation system in China to climate change, cultivars and agronomic management in the past three decades. **Climatic Change** 135, p. 539–553, 2016.
- BASSO, B.; LIU, L. Chapter four - seasonal crop yield forecast: methods, applications, and accuracies. **Advances in Agronomy** 154, 201–255, 2019.
- BENDER, F. D. **Mudanças climáticas e seus impactos na produtividade da cultura do milho e estratégias de manejo para minimização de perdas em diferentes regiões brasileiras**. Tese. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 183 p., 2017.

- CARDOSO, C. O.; FARIA, R. T. D. E.; FOLEGATTI, M. V. Aplicação do modelo CERES-MAIZE na análise de estratégias de irrigação para milho “safrinha” em Londrina-PR. **Engenharia Agrícola Jaboticabal**, v. 24, n. 1, p. 37–45, 2004.
- CHISANGA, C. B.; PHIRI, E.; CHINENE, V. R. N. Evaluating APSIM-and-DSSAT-CERES-Maize models under rainfed conditions using zambian rainfed maize cultivars. **Nitrogen**, v. 2(4), p. 392–414, 2021.
- CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos, v. 10, Safra 2022/23 – Primeiro levantamento**, Brasília, p. 1–77, 2022.
- COSTA, L. C. Modelagem e simulação em agrometeorologia. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA 10, Piracicaba, 1997. **Anais...** Piracicaba: SBA, p. 3–6, 1997.
- CRANE-DROESCH, A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. **Environmental Research Letters**, v. 13, n. 11, 114003, 2018.
- DEJONGE, K. C.; ASCOUGH, J. C., II.; AHMADI, M.; ANDALES, A. A.; ARABI, M. Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments. **Ecological Modelling**, 231, 113–125, 2012.
- EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Sistema brasileiro de classificação de solos**. 2.ed. Rio de Janeiro: Embrapa Solos, Brasília: Sistema de Produção de Informação, 306 p., 2006.
- EVERINGHAM, Y.; SEXTON, J.; SKOCAJ, D.; INMAN-BAMBER, G. Accurate prediction of sugarcane yield using a random forest algorithm. **Agronomy for Sustainable Development**, v. 36, 27, 2016.
- FANCELLI, A. L.; DOURADO NETO, D. Produção de milho. **Livraria e Editora Agropecuária**, Guaíba, 360 p., 2000.
- FEI, L.; MEIJUN, Z.; JIAQI, S.; ZEHUI, C.; XIAOLI, W.; JIUCHUN, Y. Maize, wheat and rice production potential changes in China under the background of climate change. **Agricultural Systems**, v. 182, p. 1190-1208, 2020.
- FELEKE, H. G.; SAVAGE, M. J.; TESFAYE, K. Calibration and validation of APSIM–Maize, DSSAT CERES–Maize and AquaCrop models for Ethiopian tropical environments. **South African Journal of Plant and Soil**, v. 38(1), p. 36–51, 2021.
- FENG, P.; WANG, B.; LIU, D. L.; WATERS, C.; YU, Q. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south eastern Australia. **Agricultural and Forest Meteorology**, v. 275, p. 100–113, 2019.
- FU, C.; WANG, J.; GONG, S.; ZHANG, Y.; WANG, C.; MO, Y. Optimization of irrigation and fertilization of drip-irrigated corn in the chernozem area of north-east China based on the CERES-Maize model. **Irrigation and Drainage**, v. 69, p. 714–731, 2020.

- GENUER, R.; POGGI, J.-M.; TULEAU-MALOT, C. Variable selection using random forests. **Pattern Recognition Letters**, v. 31, p. 2225–2236, 2010.
- GREAVES, G. E.; WANG, Y. M. Effect of regulated deficit irrigation scheduling on water use of corn in southern Taiwan tropical environment. **Agricultural Water Management**, v. 188, p. 115–125, 2017.
- HOOGENBOOM, G.; JONES, J. W.; PORTER, C. H.; WILKENS, P. W.; BOOTE, K. J.; HUNT, L. A.; TSUJI, G. Y. **Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.5**. University of Hawaii, Honolulu, Hawaii, 2010.
- HOOGENBOOM, G.; JONES, J. W.; WILKENS, P. W.; PORTE, C. H.; BOOTE, K. J.; HUNT, L. A.; SINGH, U.; LIZASO, J. L.; WHITE, J. W.; URYASEV, O.; ROYCE, F. S.; OGOSHI, R.; GIJSMAN, A. J.; TSUJI, G. Y. **Decision Support System for Agrotechnology Transfer (DSSAT): version 4.6**. DSSAT Foundation, Prosser, Washington, 2014.
- JIANG, R.; HE, W.; HE, L.; YANG, J. Y.; QIAN, B.; ZHOU, W.; HE, P. Modelling adaptation strategies to reduce adverse impacts of climate change on maize cropping system in Northeast China. **Scientific Reports**, v. 11, 810, 2021.
- JONES, J. W.; HOOGENBOOM, G.; PORTER, C. H.; BOOTE, K. J.; BATCHELOR, W. D.; HUNT, L. A.; WILKENS, P. W.; SINGH, U.; GIJSMAN, A. J.; RITCHIE, J. T. The DSSAT cropping system model. **European Journal of Agronomy**, v. 18, p. 235–265, 2003.
- KADIYALA, M. D. M.; JONES, J.W.; MYLAVARAPU, R. S.; LI, Y. C.; REDDY, M. D. Identifying irrigation and nitrogen best management practices for aerobic rice–maize cropping system for semi-arid tropics using CERES-rice and maize models. **Agricultural Water Management**, v. 149, p. 23–32, 2015.
- KANG, Y. H.; OZDOGAN, M.; ZHU, X. J.; YE, Z. W.; HAIN, C.; ANDERSON, M. Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest. **Environmental Research Letters**, v. 15, n. 6, 2020.
- KHAKI, S.; WANG, L.; ARCHONTOULIS, S. V. A CNN-RNN framework for crop yield prediction. **Frontiers in Plant Science**, v. 10, 1750, 2020.
- KIPKULEI, H. K.; BELLINGRATH-KIMURA, S.D.; LANA, M. et al. Modelling cropland expansion and its drivers in Trans Nzoia County, Kenya. **Modeling Earth Systems and Environment**, v. 8, p. 5761–5778, 2022.
- LI, G.; LI, J.; ZHANG, L.; XIN, Y.; HUAWEI, D. Feature selection method incorporating ant colony algorithm and random forest. **Computer Science**, v. 46, p. 212–215, 2019.
- MACCARTHY, D. S.; ADIKU, S. G.; FREDUAH, B. S.; GBEFO, F. Using CERES-Maize and ENSO as decision support tools to evaluate climate-sensitive farm management practices for maize production in the northern regions of Ghana. **Frontiers in Plant Science** 8:31, 2017.

- MALIK, W.; DECHMI, F. DSSAT modelling for best irrigation management practices assessment under Mediterranean conditions. **Agricultural Water Management**, v. 216, p. 27–43, 2019.
- MALIK, W.; ISLA, R.; DECHMI, F. DSSAT-CERES-maize modelling to improve irrigation and nitrogen management practices under Mediterranean conditions. **Agricultural Water Management**, v. 213, p. 298–308, 2019.
- MENG, L.; LIU, H.; USTIN, S.; ZHANG, X. Predicting maize yield at the plot scale of different fertilizer systems by multi-source data and machine learning methods. **Remote Sensing**, v. 13, 3760, 2021.
- MISHRA, S.; MISHRA, D.; SANTRA, G. H. Applications of machine learning techniques in agricultural crop production: A review paper. **Indian Journal of Science and Technology**, v. 9(38), p. 1–14, 2016.
- MUBEEN, M.; AHMAD, A.; WAJID, A.; KHALIQ, T.; HAMMAD, H. M.; SULTANA, S. R.; AHMAD, S.; FAHAD, S.; NASIM, W. Application of CSM-CERES-Maize model in optimizing irrigated conditions. **Outlook on Agriculture**, v. 45(3), p. 173–184, 2016.
- MUPANGWA, W.; CHIPINDU, L.; NYAGUMBO, I.; MKUHLANI, S.; SISITO, G. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. **SN Applied Sciences** 2, 952, 2020.
- OLIVEIRA, L. A. **Manejo da irrigação no cultivo da cana-de-açúcar e milho sob cenários futuros: aplicação dos modelos DSSAT/CANEGRO e CERES-MAIZE**. Dissertação. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 94 p., 2015.
- OLIVEIRA, L. A.; MIRANDA, J. H.; COOKE, R. A. C. Water management for sugarcane and corn under future climate scenarios in Brazil. **Agricultural Water Management**, v. 201, p. 199-206, 2018.
- OMOYO, N. N.; WAKHUNGU, J.; OTENG, S. Effects of climate variability on maize yield in the arid and semi-arid lands of lower eastern Kenya. **Agriculture & Food Security**, v. 4, p. 1–13, 2015.
- PAGANI, V.; STELLA, T.; GUARNERI, T.; FINOTTO, G.; van DEN BERG, M.; MARIN, F. R.; ACUTIS, M.; CONFALONIERI, R. Forecasting sugarcane yields using agro-climatic indicators and Canegro model: a case study in the main production region in Brazil. **Agricultural Systems**, v. 154, p. 45–52, 2017.
- RITCHIE, J. T. Soil water balance and plant water stress. In: TSUGI, G. Y.; HOOGENBOOM, G.; THORNTON, P. K. (Eds.). **Understanding options for agricultural production**. Dordrech: Kluwer Academic; International Consortium for Agricultural Systems Applications, p. 41–53, 1998.
- RUGIRA, P.; MA, J.; ZHENG, L.; WU, C.; LIU, E. Application of DSSAT CERES-MAIZE to identify the optimum irrigation management and sowing dates on improving maize yield in northern China. **Agronomy**, v. 11, 674, 2021.
- SHAHHOSSEINI, M.; HU, G.; HUBER, I.; ARCHONTOULIS, S. V. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. **Scientific Reports**, v. 11, 1606, 2021.

- SOIL CONSERVATION SERVICE. Hydrology section 4: In: ____ . **National engineering handbook**. Washington: USDA, NRCS. Chap. 4/10, 1972.
- SULEIMAN, A. A; RITCHIE, J. T. Modeling soil water redistribution during second-stage evaporation. **Soil Science Society of American Journal**, v. 67, p. 377–386, 2003.
- XI, M. L.; LU, D. A.; GUI, D. W.; QI, Z. M.; ZHANG, G. N. Calibration of an agricultural-hydrological model (RZWQM2) using surrogate global optimization. **Journal of Hydrology**, 544, p. 456–466, 2017.
- XIAO, D.; LI LIU, D.; WANG, B.; FENG, P.; WATERS, C. Designing high-yielding maize ideotypes to adapt changing climate in the North China Plain. **Agricultural Systems**, v. 181, 102805, 2020.
- YIN, W.; CHEN, G.; FENG, F.; GUO, Y.; HU, F.; CHEN, G.; ZHAO, C.; YU, A.; CHAI, Q. Straw retention combined with plastic mulching improves compensation of intercropped maize in arid environment. **Field Crops Research**, v. 204, p. 42–51, 2017.
- ZHANG, C.; MA, Y. Ensemble Machine Learning: Methods And Applications. Germany: **Springer**, 329 p., 2012.

5. MANEJO DA IRRIGAÇÃO E ESTIMATIVA DA PRODUTIVIDADE DE MILHO PARA A REGIÃO DE PIRACICABA, SP, UTILIZANDO MACHINE LEARNING

Resumo

Os desafios da agricultura contemporânea para os próximos 50 anos incluem questões relacionadas à água, energia e alimentos. O milho caracteriza-se como um produto fundamental para o agronegócio brasileiro, cultivado em todas as regiões do país. A sua produção depende diretamente da relação estabelecida entre a planta e os fatores climáticos do ambiente em que se encontra. O manejo da irrigação baseado na demanda de água da cultura, representada pela taxa da evapotranspiração, é amplamente utilizado. A evapotranspiração de referência (ET_o) pode ser determinada por diversos métodos indiretos, sendo a equação da FAO-56 Penman-Monteith (PM) entendida como a mais precisa. Como alternativa a esse método, diversos estudos mostraram a eficiência de algoritmos de *machine learning* em estimar a ET_o, bem como o rendimento da produção, contudo, esses modelos dificilmente são colocados à prova em situações reais de campo. Assim sendo, o objetivo da presente pesquisa foi aplicar os modelos preditivos de *machine learning* em um experimento de campo para realização do manejo da irrigação e estimativa da produtividade do milho. Para tal, o experimento foi conduzido no Departamento de Engenharia de Biosistemas na Escola Superior de Agricultura “Luiz de Queiroz”, ESALQ/USP, localizada em Piracicaba, SP. O delineamento experimental foi inteiramente casualizado, com quatro tratamentos e cinco repetições. Os tratamentos consistiram nos modelos utilizados para o cálculo da ET_o, os quais determinaram a quantidade de água evapotranspirada pela cultura que deve ser reposta. Ao final do ciclo, foram feitas avaliações biométricas e as variáveis foram submetidas à análise de variância. Pelos resultados obtidos, pode-se notar que os modelos desenvolvidos apresentaram desempenhos satisfatórios na predição da ET_o durante o ciclo de cultivo do milho e do rendimento de grãos. Também foi constatado que não houve diferença significativa para todas as variáveis relacionadas ao crescimento, desenvolvimento e produção do milho em relação ao tratamento utilizado para o manejo da irrigação. Portanto, os algoritmos de aprendizado de máquina adotados mostraram-se como ferramentas expressivas para predição da ET_o durante o manejo da irrigação e da produtividade de milho nas condições de Piracicaba, sendo alternativas à modelos e métodos clássicos utilizados para uma gestão eficiente do sistema produtivo.

Palavras-chave: Evapotranspiração de referência, Rendimento agrícola, Gestão hídrica, Coeficiente de cultura

Abstract

The challenges of contemporary agriculture for the next 50 years include issues related to water, energy and food. Maize is characterized as a fundamental product for Brazilian agribusiness, cultivated in all regions of the country. Its production depends directly on the relationship established between plant and climatic factors of the environment in which it is grown. Irrigation management based on crop water demand, represented by the evapotranspiration rate, is widely used. Reference evapotranspiration (ET_o) can be determined by various indirect methods, with the FAO-56 Penman-Monteith equation (PM) considered the most accurate. As an alternative to this method, several studies have shown the efficiency of machine learning algorithms in estimating ET_o and crop yield. However, these models are rarely tested in real field situations. Therefore, the objective of this research was to apply predictive machine learning models in a field experiment to manage irrigation and estimate maize productivity. For this purpose, the experiment was conducted at the Department of Biosystems Engineering at the “Luiz de Queiroz”

College of Agriculture, ESALQ/USP, located in Piracicaba, SP. The experimental design was completely randomized, with four treatments and five replications. The treatments consisted of the models used to calculate ETo, which determined the amount of water evapotranspired by the crop that needs to be replenished. When the cultivation cycle ended, biometric evaluations were performed and the variables were subjected to analysis of variance. Results showed that the developed models exhibited satisfactory performance in predicting ETo and grain yield during the maize cultivation cycle. It was also found that there was no significant difference for all variables related to maize growth, development and production in relation to the treatment used for irrigation management. Therefore, the adopted machine learning algorithms proved to be significant tools for predicting ETo during irrigation management and maize productivity in Piracicaba conditions, providing alternatives to classical models and methods used for an efficient management of the production system.

Keywords: Reference evapotranspiration, Agricultural yield, Water management, Crop coefficient

5.1. Introdução

As projeções dos desafios globais para os próximos 50 anos mostram que os principais problemas para a humanidade serão aspectos relacionados à água, energia e alimentos. Nesse cenário da crescente demanda por alimentos, torna-se necessária a busca por ferramentas capazes de otimizar cada vez mais os processos agrícolas, possibilitando aumento da produtividade de modo sustentável (OLIVEIRA et al., 2018).

O milho (*Zea mays* L.) constitui-se em um dos mais importantes cereais cultivados, consumidos e estudados no mundo, sendo responsável por fornecer produtos utilizados para a alimentação humana e animal e matérias-primas para a indústria, tais como combustíveis, óleos, bebidas, polímeros, entre outros. Devido à multiplicidade de suas aplicações, a cultura do milho assume papel relevante em termos socioeconômicos (GREAVES e WANG, 2017; MIRANDA, 2018).

A cadeia produtiva do milho constitui-se em uma das mais importantes do cenário agrícola brasileiro, correspondendo a 37% da produção nacional de grãos. A principal região produtora é a Centro-Oeste, seguida pelas regiões Sul, Nordeste, Sudeste e Norte. No Brasil, foram cultivados 21,58 milhões de hectares e produzidos 113,13 milhões de toneladas de milho na safra 2021/22, com produtividade de 5.242 kg ha⁻¹, sendo as maiores áreas observadas nos estados do Mato Grosso, Paraná e Mato Grosso do Sul, os quais foram responsáveis pela produção de 41,62, 16,42 e 12,46 milhões de toneladas, respectivamente. A principal região produtora no Brasil é a Centro-Oeste, seguida pelas regiões Sul, Nordeste, Sudeste e Norte (CONAB, 2022).

A produção de milho depende diretamente da relação estabelecida entre a planta e os fatores climáticos do ambiente em que se encontra. Nesse contexto, a disponibilidade e a gestão hídrica ocupam posição de destaque no ciclo de cultivo, uma vez que o excesso ou a falta de água causam injúrias e redução da produtividade das plantas, especialmente quando ocorrem nos estádios de florescimento e enchimento de grãos. Por isso, o conhecimento e entendimento em

relação à quantidade de água a ser fornecida para a cultura tornam-se essenciais, a fim de se obter um manejo consciente, o qual permitirá que as plantas se desenvolvam de forma satisfatória (KIPKULEI et al., 2022; PALHARES, 2003; TAIZ e ZEIGER, 2013).

Nesse sentido, a reposição de água ao solo por meio da irrigação, na quantidade e no momento oportuno, apresenta-se como fator decisivo para maximizar o processo produtivo. Para tal, o manejo da irrigação baseado na necessidade de água da cultura, definida pela demanda evapotranspirométrica, é amplamente utilizado (BERNARDO et al., 2019).

A evapotranspiração (ET) representa a soma dos componentes da evaporação da água da superfície do solo e da transpiração das plantas via liberação de moléculas de água pelo estômato. Esse processo consiste no principal parâmetro para determinação do consumo de água das plantas, por isso a sua estimativa proporciona a quantidade ideal de água a ser aplicada na cultura pelo sistema de irrigação (JING et al., 2019; WANG et al., 2019).

A evapotranspiração de referência (ET_o) pode ser determinada por diversos métodos indiretos, sendo a equação da FAO-56 Penman-Monteith (PM) entendida como a mais precisa e amplamente recomendada. No entanto, de acordo com Yamaç e Todorovic (2020), a utilização dessa abordagem é dificultada por necessitar de quantidade e qualidade significativas dos dados, bem como da integração de constantes empíricas, sendo que as estações meteorológicas que satisfazem esses requisitos de observação são limitadas, especialmente em países em desenvolvimento. Assim, torna-se necessário o desenvolvimento de modelos mais simples para determinação da ET_o com razoável precisão.

Como alternativa à essa abordagem, diferentes técnicas de *machine learning* vêm sendo aplicadas à estimativa da ET_o para o desenvolvimento de métodos práticos e de baixo custo, como apresentado em Chen et al., 2020; Chia et al., 2020; Granata, 2019; Mohammadi e Mehdizadeh, 2020; Raza et al., 2020; e Wang et al., 2019. Essas técnicas podem ser utilizadas para determinação da ET_o diária, a qual é empregada para a realização do manejo de irrigação.

Além do manejo da irrigação, modelos capazes de simular a interação solo-planta-atmosfera e auxiliar na previsão da safra de determinada cultura também se destacam como ferramentas essenciais para otimização dos processos agrícolas. Por isso, diferentes abordagens baseadas no aprendizado de máquina para estimar o crescimento e desenvolvimento de plantas vêm sendo amplamente difundidas (BASSO e LIU, 2019; CRANE-DOESCH, 2018; MISHRA et al., 2016; MUPANGWA et al., 2020).

Diversos estudos demonstraram o desempenho satisfatório de algoritmos de *machine learning* para determinação da ET_o, bem como para estimativa do rendimento do cultivo. Contudo, esses modelos dificilmente são colocados à prova em experimentos de campo, o que permite avaliar

a resposta da cultura em tempo real ao manejo da irrigação utilizando as técnicas alternativas em comparação com o método padrão.

Por isso, a hipótese da presente pesquisa é que o uso de um manejo de irrigação baseado em um modelo preditivo de aprendizado de máquina, possa ser uma alternativa na determinação do volume de água aplicado, sem alterar as variáveis de crescimento e de produção de plantas de milho.

De modo a testar a hipótese citada, os objetivos dessa pesquisa foram: i) aplicar os modelos preditivos de *machine learning* em um experimento de campo para realização do manejo da irrigação do milho; ii) observar a ocorrência de diferença significativa nas variáveis de crescimento e de produção de plantas de milho sob os diferentes tratamentos em experimento de campo; e iii) avaliar o desempenho dos modelos preditivos de produção do milho com base nos dados observados em campo.

5.2. Material e Métodos

5.2.1. Localização da área experimental

O experimento foi conduzido em área experimental do Departamento de Engenharia de Biosistemas (LEB/ESALQ/USP) na cidade de Piracicaba, São Paulo, latitude 22° 43' 30" e longitude 47° 38' 01" numa altitude de 547 m e clima Aw na classificação de Köppen (Figura 16).



Figura 16. Localização e vista da área experimental no Departamento de Engenharia de Biosistemas na ESALQ/USP.

5.2.2. Preenchimento dos recipientes e instalação dos tensiômetros

Foram utilizados vinte vasos plásticos com capacidade de 50 L, altura de 0,38 m, diâmetro superior de 0,45 m e inferior de 0,40 m, os quais foram alocados sobre blocos de concreto sem espaçamento entre vasos e 1 m entre fileiras. Estes foram preenchidos com Latossolo Vermelho

Amarelo (LVA), fase arenosa, série “Sertãozinho”, coletados no município de Piracicaba, SP e classificados conforme EMBRAPA (2006).

Pela Tabela 20 pode-se conferir os valores dos resultados da análise granulométrica do solo, realizada após terem sido destorroados, seco ao ar e peneirados em uma malha de 2 mm, seguindo a recomendação do Manual de Métodos de Análises do Solo (EMBRAPA, 2011). A análise química do solo encontra-se na Tabela 21, realizada a partir de amostras desse material, com o objetivo de conhecer as condições iniciais do solo.

Tabela 20. Caracterização física do Latossolo Vermelho Amarelo (LVA).

Características	LVA
Granulometria (g kg ⁻¹)	
Areia grossa	290
Areia fina	520
Silte	27
Argila	163
Classe Textural*	Franco-arenosa
Densidade do solo (kg dm ⁻³)	1,48
Densidade de partículas (kg dm ⁻³)	2,60
Porosidade (%)	43,07
Condutividade hidráulica do solo saturado (K _o , cm h ⁻¹)	9,5

*Classificação segundo USDA

Tabela 21. Caracterização química do Latossolo Vermelho Amarelo (LVA).

Parâmetros	LVA
pH (CaCl ₂)	5,18
Matéria Orgânica (g dm ⁻³)	15,7
Fósforo (P - mg dm ⁻³)	18,4
	mmol _c dm ⁻³
Potássio (K)	2,24
Cálcio (Ca)	22,7
Magnésio (Mg)	8,7
Hidrogênio (H) + Alumínio (Al)	22,4
Soma de bases trocáveis (SB)	33,6
Alumínio (Al)	< 0,1
Capacidade de troca de cátions (CTC)	56

Baseando-se nos resultados apresentados nas Tabelas 20 e 21 e de posse das recomendações do Boletim Técnico 100, procedeu-se a interpretação de adubação recomendada para a cultura de milho no Estado de São Paulo.

Com os blocos já dispostos na área, os vasos começaram a ser preenchidos. O preenchimento dos recipientes foi dividido em camadas, na qual na primeira (parte inferior) foi adicionada uma subcamada constituída de brita e coberta por uma manta geotêxtil, com intuito de

evitar perdas de solo pelo dreno instalado no fundo dos vasos. Após essa subcamada, as próximas foram preenchidas com solo utilizando um recipiente auxiliar de volume conhecido. A cada camada adicionada, o solo foi adensado com intuito de obter a densidade real de um Latossolo Vermelho Amarelo ($1,45 \text{ g cm}^{-3}$) (Figura 17).

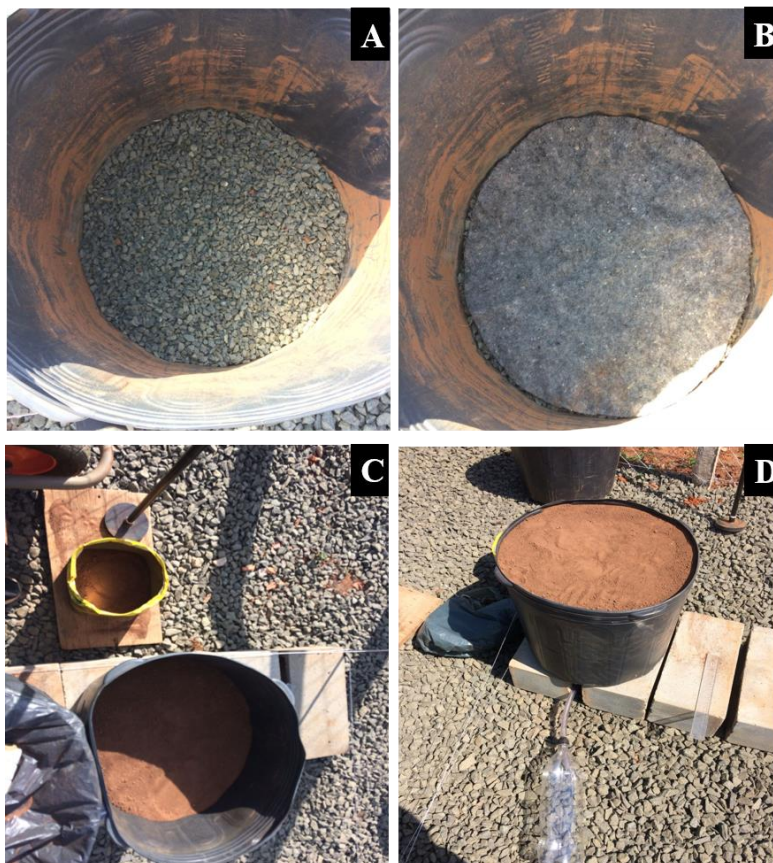


Figura 17. Camada de brita (A), coberta por uma manta geotêxtil (B). Vaso de 50 L sendo preenchido com Latossolo Vermelho Amarelo com a utilização de um recipiente auxiliar de volume conhecido e um instrumento para compactar o solo (C). Vaso completamente preenchido e com recipiente instalado para obtenção do volume drenado (D).

Os vasos foram adaptados com válvula de drenagem. Cada válvula foi conectada a um reservatório específico por meio da mangueira, com o intuito de coletar e medir o volume drenado.

Com o objetivo de monitorar a umidade do solo durante todo o ciclo do milho, tensiômetros foram instalados a 20 cm de profundidade (Figura 18). A medição do potencial mátrico foi realizada a cada dois dias com utilização do tensímetro e a umidade volumétrica do solo foi obtida a partir da curva de retenção de água no solo elaborada em laboratório.



Figura 18. Tensiômetros instalados a 20 cm de profundidade (A). Tensímetro utilizado para medir o potencial mátrico do solo (B).

5.2.3. Parâmetros da curva de retenção de água no solo

Três amostras de solos com estrutura deformada foram coletadas na camada superficial dos vasos com a utilização do anel de Kopeck e levadas ao laboratório para a obtenção dos parâmetros de ajuste da curva de retenção de água no solo (Eqs. 20 e 21). Foram estabelecidos nove pontos para elaboração da curva, sendo os valores de umidade volumétrica para as tensões de 0 (saturação), 10, 20, 40 e 100 cm.c.a. utilizando-se a mesa de tensão e de 300, 500, 1000, 5000 e 15000 cm.c.a. com a câmara de Richards (KLUTE, 1986).

$$\theta(\Psi_m) = \theta_r + \frac{(\theta_s - \theta_r)}{[1 + (\alpha \cdot |\Psi_m|)^n]^m} \quad (20)$$

$$m = 1 - \frac{1}{n}, n > 1 \quad (21)$$

Em que:

$\theta(\Psi_m)$ = umidade volumétrica do solo em função do potencial mátrico do solo (Ψ_m), $\text{cm}^3 \text{cm}^{-3}$;

θ_s = umidade volumétrica do solo saturado, $\text{cm}^3 \text{cm}^{-3}$;

θ_r = umidade volumétrica residual do solo, $\text{cm}^3 \text{cm}^{-3}$;

m, n = parâmetros de regressão da equação; adimensional;

α = parâmetro com dimensão igual ao inverso da tensão, cm^{-1} ;

Ψ_m = potencial mátrico, cm.

As médias da umidade com base em volume das três repetições foram calculadas e inseridas no programa computacional RETC (REtention Curve), conforme a sequência detalhada na Figura 19, para realização de seu ajuste numérico segundo o modelo de van GENUCHTEN (1980), com o objetivo de caracterizar as propriedades hidráulicas do solo em estudo.

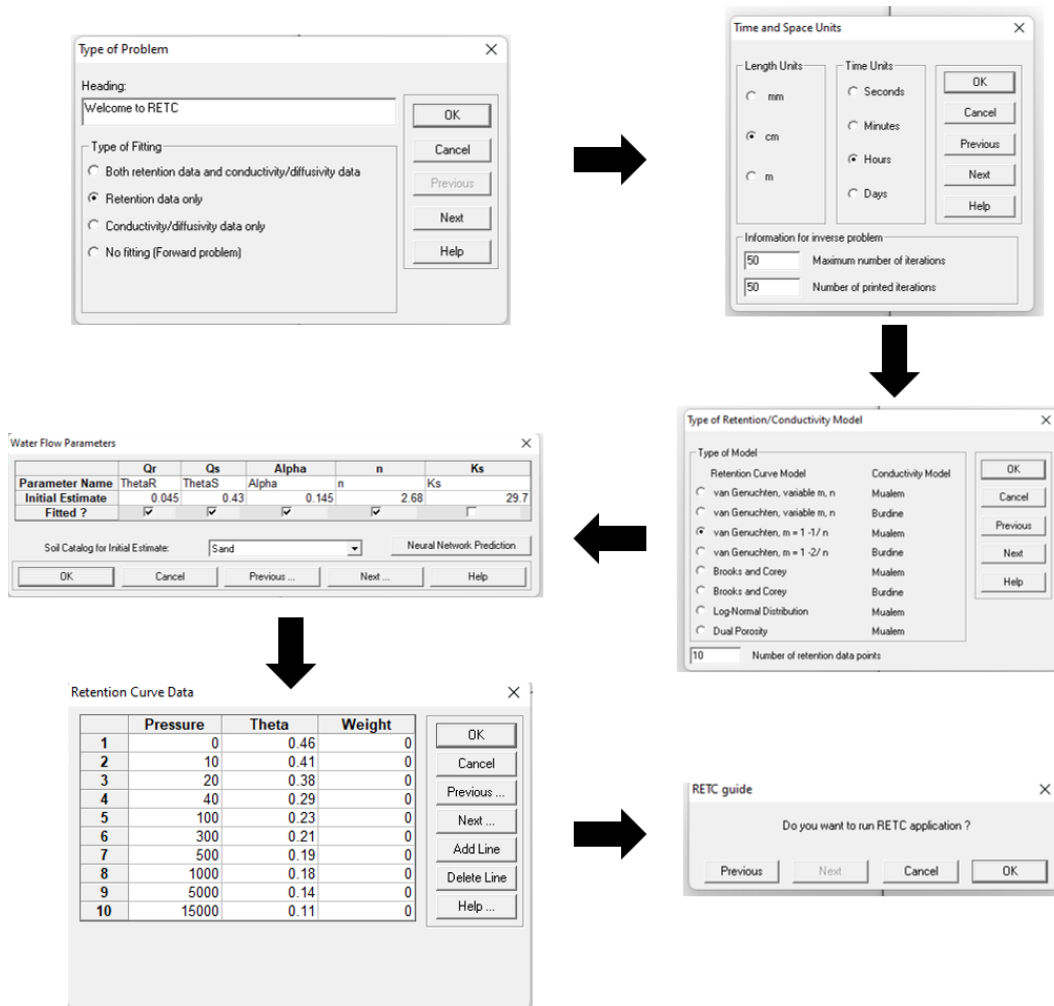


Figura 19. Etapas executadas no software RETC para elaboração da curva de retenção de água no solo.

Pela Figura 20, pode-se observar a curva de retenção de água ajustada pelo programa RETC para o Latossolo Vermelho Amarelo a partir dos valores de umidade volumétrica observados.

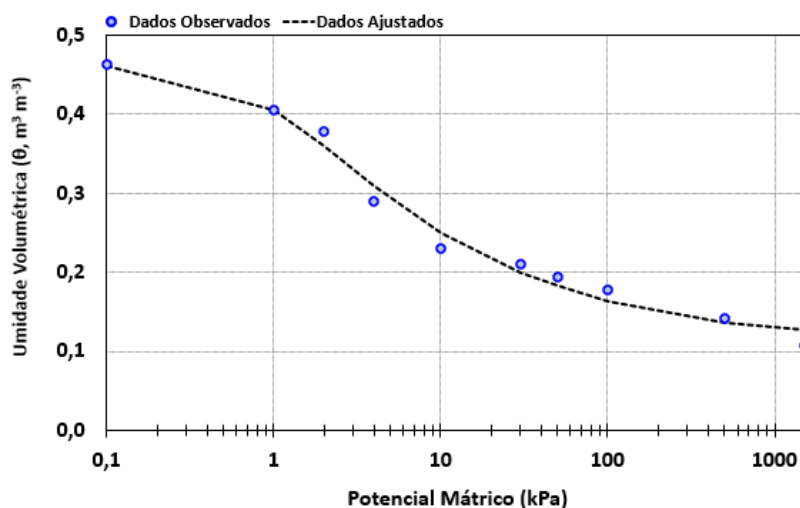


Figura 20. Curva de retenção de água no solo ajustada em função do potencial mátrico para o Latossolo Vermelho Amarelo.

5.2.4. Delineamento experimental

O ensaio foi distribuído seguindo o delineamento inteiramente casualizado (Figura 21), com quatro tratamentos e cinco repetições, totalizando 20 unidades experimentais. Os tratamentos consistiram nos modelos utilizados para o cálculo da ETo, os quais determinaram a quantidade de água evapotranspirada pela cultura que foi reposta via irrigação. Esse valor foi calculado a partir do método padrão de Penman-Monteith (PM) e dos algoritmos previamente treinados e testados (SVR, HGBR e MLPR).

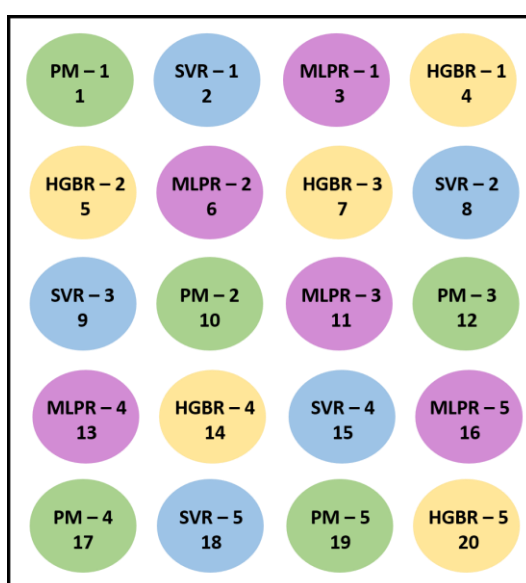


Figura 21. Croqui do experimento totalmente casualizado com seus respectivos tratamentos e repetições.

5.2.5. Condução do experimento

Para a instalação do experimento com a cultura de milho, utilizou-se o espaçamento de plantio de 0,35 m entre plantas, com população de 62.876 plantas por hectare. A semeadura foi realizada manualmente a 0,05 m no dia 26/05/2022, quando o solo foi umedecido à capacidade máxima de retenção de água, utilizando cerca de 50% a mais do número de sementes em razão das possíveis falhas de germinação, ocorrência de geada, ataque de pragas e outros fatores de redução de estande. Foi utilizado o híbrido FS710PWU da Forseed Sementes® (Figura 22), caracterizado por apresentar alto potencial produtivo, qualidade de colmo, tolerância ao estresse hídrico e rusticidade para condições adversas. O estande foi definido após o desbaste, quando as plantas se encontravam com três a quatro folhas totalmente expandidas, sendo mantidas duas plantas por vaso, totalizando 40.



Figura 22. Híbrido FS710PWU da Forseed Sementes®.

Os manejos de adubação foram baseados na análise do solo e na recomendação proposta por Fancelli e Dourado Neto (2000). A adubação convencional foi aplicada em três momentos: na fundação a 10 cm de profundidade (100 kg ha⁻¹ de superfosfato simples, 100 kg ha⁻¹ de cloreto de potássio e 100 kg ha⁻¹ de ureia), com 8 e 12 folhas desenvolvidas (40 kg ha⁻¹ de cloreto de potássio e 55 kg ha⁻¹ de ureia), constituindo a adubação de cobertura. Na fase R1, foi aplicado via fertirrigação o fertilizante YaraTera® REXOLIN CXK, com o objetivo de fornecer micronutrientes às plantas.

Na área experimental foi instalada uma estação meteorológica automática (Figura 23A) para a coleta de dados climáticos. Os dados meteorológicos coletados foram temperatura do ar ($^{\circ}\text{C}$), radiação solar (W m^{-2}), umidade relativa do ar (%) e velocidade do vento (m s^{-1}), permitindo, assim, a estimativa da evapotranspiração de referência diária. Para auxiliar no balanço hídrico, também foi instalado um pluviômetro (Figura 23B).



Figura 23. Estação meteorológica automática (A) e pluviômetro Ville de Paris (B) instalados na área experimental.

Para a estimativa da evapotranspiração da cultura (ET_c), foram identificadas as fases fenológicas do ciclo do milho, uma vez que a demanda hídrica da cultura é influenciada pela sua fase fenológica, e os coeficientes de cultivo (K_c) propostos no Manual FAO 56 (ALLEN et al., 1998). O valor da ET_o é multiplicado pelo K_c correspondente (Eq. 22), sendo, assim, obtido o consumo de água diário da planta.

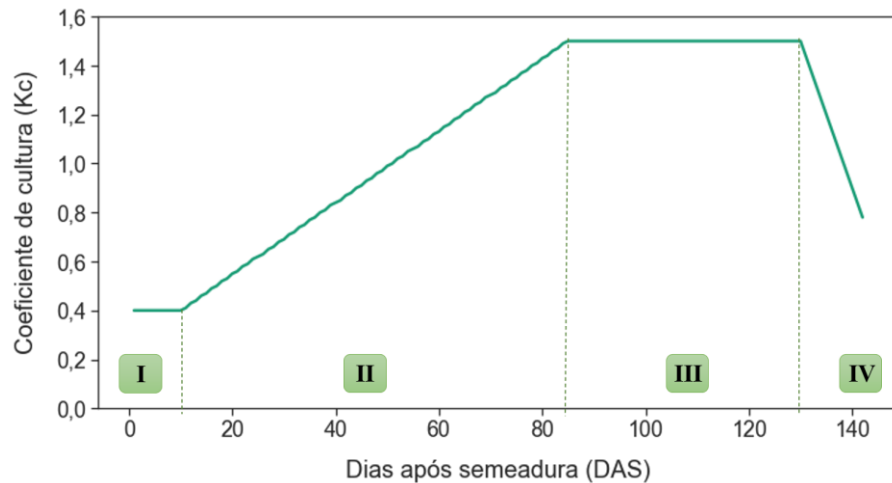
$$\text{ET}_c = \text{ET}_o \cdot \text{K}_c \quad (22)$$

Allen et al. (1998) adotaram quatro fases distintas para definir os valores de K_c : Fase I, período compreendido desde a semeadura até o aparecimento da plântula (germinação e emergência), sua duração depende da temperatura e umidade do solo germinação e emergência; Fase II, período compreendido entre a emissão da segunda folha e o início do florescimento (crescimento vegetativo); Fase III, caracterizada pelo início da polinização, fecundação até o enchimento dos grãos, podendo durar entre 40 e 60 dias (estádio reprodutivo); e Fase IV, período compreendido desde o começo da maturação até a senescência, indicativa do final do ciclo de vida da planta. A duração de cada etapa do ciclo é altamente influenciada pela variedade, condições edafoclimáticas e data de plantio. Com isso, os valores de K_c foram tabelados (Tabela 22).

Tabela 22. Coeficientes de cultivo da cultura do milho em função dos estádios de desenvolvimento.

Fases	I (inicial)	II	III (médio)	IV	V (final)
Kc	0,4	0,41 – 1,49	1,5	1,44 – 0,78	0,6

Nos estádios I e III, os valores de Kc permanecem constantes, enquanto que nas fases II e IV, os valores variam linearmente (Figura 24) entre as fases de valores constantes.

**Figura 24.** Curva do coeficiente de cultivo (Kc) da cultura do milho em função das fases fenológicas.

O manejo de irrigação consistiu na reposição da água evapotranspirada diariamente pela cultura (ETc), valor esse calculado pela metodologia padrão e pelos três algoritmos treinados e testados, com turno de rega fixo de dois dias. A diferenciação dos tratamentos ocorreu 15 DAS para garantir o estabelecimento da cultura.

Pela Figura 25 pode-se observar uma visão geral da estrutura do experimento em que as plantas iniciaram a fase reprodutiva.

**Figura 25.** Visão geral da estrutura do experimento com 40 plantas de milho em 20 vasos.

A fase vegetativa da cultura durou 75 dias, enquanto que o florescimento ocorreu na semana 12 e o ciclo de cultivo teve duração de 142 dias. A colheita ocorreu em 14/10/2022 com um total de 40 espigas de milho colhidas.

5.2.6. Colheita

A colheita foi realizada manualmente no dia 14 de outubro de 2022, aos 142 dias após a semeadura (DAS), momento em que os grãos iniciaram o processo de maturidade fisiológica. As espigas foram retiradas, armazenadas em sacos identificados e levadas à estufa para secagem à 60 °C durante sete dias. Após a secagem, deu-se início às medições dos componentes de produção da cultura.

5.2.7. Avaliações biométricas e componentes de produção

As plantas de todos os tratamentos foram cortadas no nível do solo e as espigas colhidas após 142 dias de ciclo, para serem avaliados os seguintes parâmetros biométricos: altura das plantas (ALP), diâmetro do colmo (DIP), altura de inserção da espiga (AIE), peso da espiga com palha (PEP), peso da espiga despalhada (PES), tamanho (TE) e diâmetro da espiga (DE) e número de fileiras de grãos por espiga (NFG).

Após as avaliações biométricas, foi realizada a debulha manual do milho, com objetivo de manter o controle e aproveitar ao máximo os grãos de cada tratamento, para a determinação da produtividade de milho. Foi realizada a contagem do número de grãos por espiga (Ngrãos), enquanto o peso dos grãos (PG) foi determinado utilizando uma balança de precisão de três dígitos.

A produtividade (Prod) foi obtida após pesagem dos grãos com uma balança de precisão de três dígitos. Em seguida, foi determinado o peso médio de cada tratamento e feita a extrapolação do resultado para kg ha^{-1} .

5.2.8. Análises estatísticas

Os resultados das análises biométricas foram submetidos aos testes de Shapiro-Wilk, Kolmogorov-Smirnov e Bartlett para verificação da normalidade e homocedasticidade das variâncias, através do software R (versão 4.3.0). Em seguida, foram realizadas as análises de variância (ANOVA) aplicando-se o teste F, com o objetivo de determinar se existiram diferenças estatísticas significativas entre os tratamentos ao nível de 5% de probabilidade.

5.3. Resultados e Discussão

5.3.1. Condições meteorológicas

As variações da temperatura do ar, umidade relativa do ar, velocidade do vento e radiação solar líquida ao longo do ciclo de cultivo do milho, aqui caracterizados como a média diária ao longo do período de observação, podem ser conferidas na Figura 26.

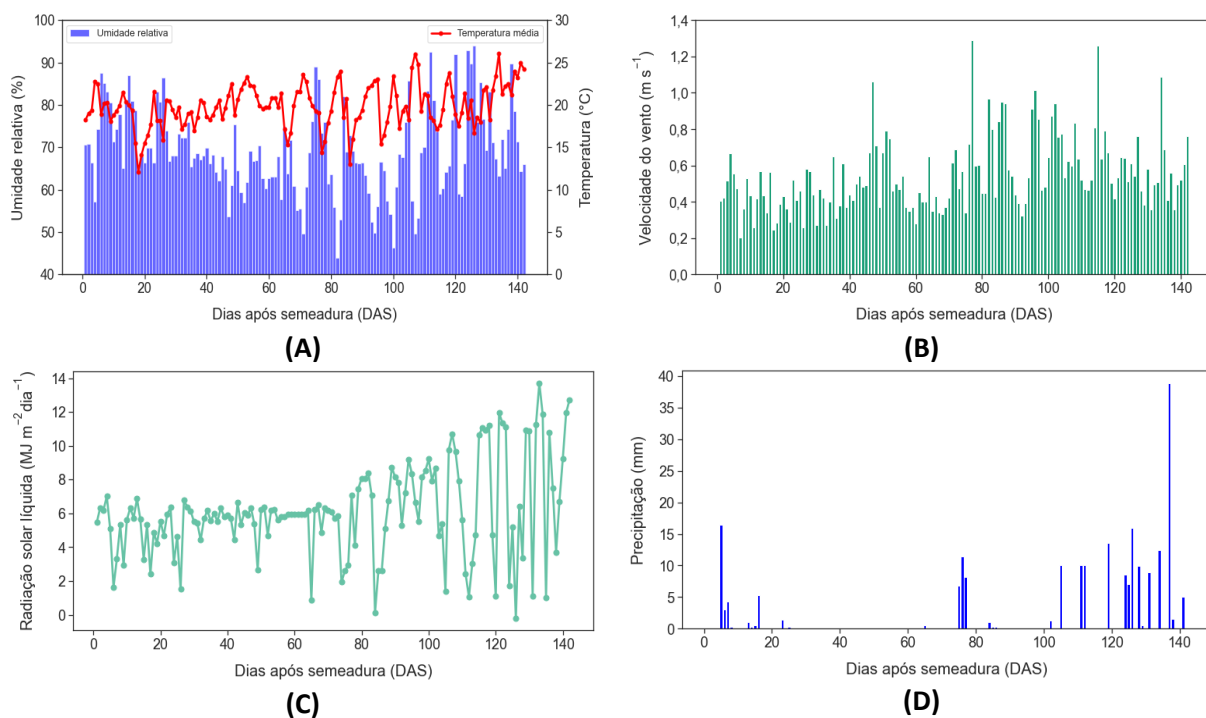


Figura 26. Valores diários da umidade relativa do ar e temperatura média (A), velocidade do vento (B), radiação solar líquida (C) e precipitação (D) referentes ao período de cultivo da cultura de milho (26/05/2022 – 14/10/2022).

A temperatura média do ar no decorrer do ciclo da cultura oscilou entre 12,07 e 26,07 °C, com média de 19,86 °C, enquanto que a umidade relativa do ar variou entre 43,91 e 93,92%, com média de 69,01%. A velocidade média do vento na área de cultivo foi de 0,54 m s⁻¹. Tanto a temperatura do ar quanto a umidade relativa apresentam elevada importância no processo germinativo do milho e são responsáveis por determinar a sua duração.

Um dos fatores de maior relevância na produção de milho é a temperatura do ar, como mostrado na Figura 13, sendo que o processo de germinação é desfavorecido em condições de temperatura do solo abaixo de 15 °C.

Segundo Fancelli (2015), para semeaduras em períodos mais frios, como ocorrido nesse experimento, é recomendada a adoção de medidas que favoreçam a germinação e emergência de plântulas, como a redução da profundidade de semeadura, o tratamento de sementes com

fungicidas e a utilização de sementes íntegras e vigorosas, garantindo um adequado estande de plantas. As temperaturas mais baixas, porém, provocam adaptações das plantas, como a redução da área foliar, o que, associado a baixos valores de radiação solar líquida, impacta diretamente na atividade fotossintética e conseqüentemente no crescimento e desenvolvimento da cultura.

De acordo com Souza e Barbosa (2015), a ocorrência de altas temperaturas contribui para o aumento da transpiração vegetal, enquanto que noites quentes promovem o aumento da respiração das plantas, aumentando o consumo de fotoassimilados e gastos com respiração, causando prejuízo ao crescimento da cultura.

De forma geral, pode-se inferir que, durante o experimento, a temperatura do ar foi adequada, visto que segundo Didonet et al. (2002) a cultura do milho produz satisfatoriamente com temperaturas médias entre 18 e 25 °C.

As condições favoráveis durante as fases R2-R5, correspondentes à grãos leitosos e grãos duros, respectivamente, são: temperatura entre 16 e 35 °C e umidade relativa superior a 65%, o que foi observado durante o ciclo. Caso contrário altas temperaturas e baixos valores de umidade podem reduzir a viabilidade dos grãos de pólen, reduzindo a fecundação e conseqüentemente a produtividade.

A variação da radiação solar líquida ao longo do ciclo esteve entre 0,12 e 12,72 MJ m⁻² dia⁻¹, com valor médio observado de 6,11 MJ m⁻² dia⁻¹. O milho apresenta maior sensibilidade à variação de luz no início da fase reprodutiva, ou seja, nos primeiros 15 dias após o pendoamento, onde a radiação solar líquida apresentou-se maior que a média do ciclo, favorecendo a produção. O aproveitamento efetivo da radiação solar no milho depende da estrutura da planta, especialmente da distribuição espacial das folhas, sendo que a redução de 30 a 40% na intensidade luminosa causa atraso na maturação dos grãos (FANCELLI, 2015).

Cabe ressaltar que existe uma forte relação entre o número de grãos, principal fator determinante da produtividade, e a taxa de crescimento da planta, a qual é basicamente impulsionada pela fotossíntese. Por isso, a baixa intensidade luminosa, como observada no experimento, especialmente durante o período vegetativo, influenciou diretamente no crescimento e produção do milho, sendo que o início da fase reprodutiva foi retardado, ocasionando em redução do número de grãos por planta, o qual é determinado por volta da época de espigamento.

A normal climatológica de 1992-2021 para a cidade de Piracicaba é marcada por chuvas anuais de aproximadamente 1257,8 mm, com distribuição heterogênea, onde o total precipitado para o mês mais chuvoso (janeiro) é de 225,7 mm e para o mês mais seco (agosto) de 26,3 mm. Nesse sentido, durante o experimento, foi observada uma precipitação total de 215,17 mm, valor considerado baixo diante da necessidade hídrica da planta, a qual varia de 500 a 800 mm, quando bem distribuída ao longo do ciclo (CRITCHLEY e SIEGERT, 1991).

Pode-se observar que até 100 DAS não houve ocorrência de chuvas significativas, o que representou um período crítico para as plantas, especialmente por ser a fase de estabelecimento da cultura, onde ocorre a definição do potencial produtivo da lavoura. Foi nesse período que a complementação de água por meio da irrigação tornou-se essencial para promover um desenvolvimento contínuo dos elementos de rendimento do milho, apesar do estresse hídrico apresentado pelas plantas.

Nos primeiros dias da floração, a qual iniciou-se 92 DAS, ainda pode-se observar um período de seca aliado com baixos valores de umidade relativa, o que pode ter contribuído para desempenho inferior da cultura, visto que essa fase representa alto consumo hídrico do milho.

As plantas tendem a apresentar extrema sensibilidade ao déficit hídrico especialmente durante a floração até o início da fase de enchimento dos grãos, por isso a deficiência de água nessas etapas provoca elevada redução da produtividade. Durante essas fases caracterizadas por requererem maiores quantidades de água, as quais ocorreram a partir de 92 DAS, observou-se início da ocorrência de chuvas, o que, em conjunto com a irrigação, contribuiu para manter o desempenho da cultura.

A variação da ETo durante o cultivo, determinada a partir do método padrão (PM), pode ser observada pela Figura 27. A demanda atmosférica, caracterizada pela ETo, apresentou valores médios de $1,88 \text{ mm dia}^{-1}$, em função dos menores índices de radiação solar e temperatura, com acumulado de $266,22 \text{ mm ciclo}^{-1}$.

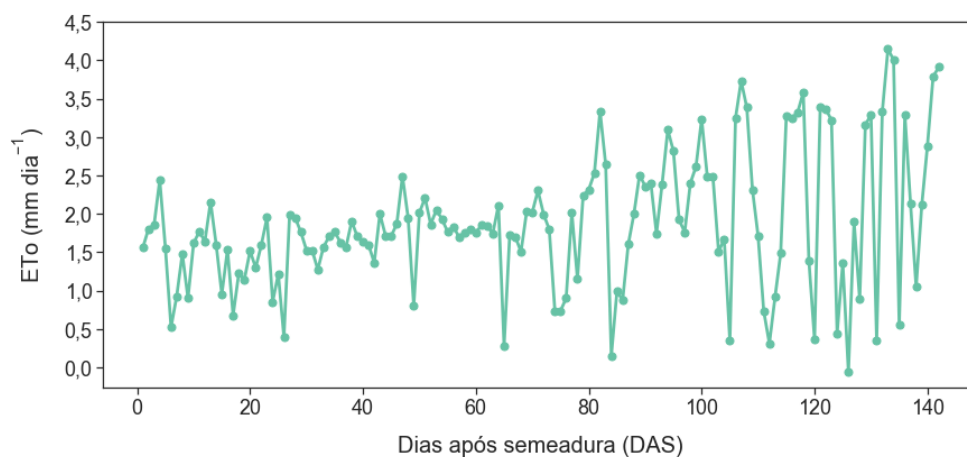


Figura 27. Evapotranspiração de referência (ETo) durante o período de cultivo da cultura de milho (26/05/2022 – 14/10/2022).

5.3.2. Determinação da ETo

Com o intuito de realizar o manejo da irrigação do milho, a ETo foi calculada diariamente durante o ciclo de cultivo a partir do método de PM e dos modelos de *machine learning* desenvolvidos para estimar a ETo com base em 10 anos de dados, conforme apresentados no Capítulo 1 da presente pesquisa.

A utilização de algoritmos de *machine learning* com quantidade reduzida de variáveis de entrada mostrou efeito positivo e significativo como uma alternativa para a estimativa da ETo em um experimento de campo (Tabela 23). De acordo com os índices estatísticos apresentados, nos quais os valores de ETo obtidos pelos algoritmos foram comparados com os do método padrão, é possível verificar que os três modelos obtiveram uma performance satisfatória, sendo que o SVR apresentou o melhor desempenho na predição da ETo para as condições do experimento, com R^2 de 0,986 e RMSE igual à 0,102 mm dia⁻¹.

Tabela 23. Performance comparativa dos algoritmos na estimativa da ETo durante o ciclo de cultivo.

Modelo	R^2	RMSE (mm dia ⁻¹)
SVR	0,986	0,102
HGBR	0,943	0,210
MLPR	0,921	0,247

A confiabilidade dos algoritmos de regressão em fazer previsões da ETo durante a condução do experimento foi garantida pelos altos e baixos valores de R^2 e RMSE, respectivamente. Os valores de R^2 indicaram que os modelos SVR, HGBR e MLPR foram capazes de fazer estimativas com uma precisão de aproximadamente 98%, 94% e 92%, respectivamente, por meio de uma combinação linear de variáveis relacionadas à temperatura, radiação solar, umidade e vento.

Os valores aqui obtidos corroboram com os resultados apresentados pela Tabela 11, onde os três algoritmos mostraram eficiência na predição da ETo, com o MLPR apresentando uma inferioridade na precisão, o que resultou em maiores RMSE. No que se refere à essa comparação, pode-se notar uma redução na acurácia dos modelos e aumento do RMSE quando esses foram aplicados com o objetivo de determinar a ETo diariamente durante a condução do cultivo de milho, isso provavelmente pode ser explicado por serem condições distintas de treino, onde o espaço de busca durante experimento de campo foi distinto em relação aos 10 anos de dados disponíveis para treinamento dos algoritmos.

Uma comparação mais abrangente entre os valores da ETo calculados a partir do método da FAO-PM e os valores estimados diariamente por cada algoritmo durante o experimento, pode ser conferida pela Figura 28.

Nela é possível analisar que os modelos geraram estimativas pouco dispersas, com elevado R^2 e seguiram o padrão estabelecido pela abordagem FAO-PM. Os algoritmos HGBR e MLPR produziram estimativas mais dispersas do que o SVR, sendo que este foi o modelo mais preciso para as condições do experimento. Os valores da ETo estimados oscilaram durante o ciclo de cultivo no mesmo intervalo observado para o método padrão, indicando coerência nas previsões ao longo dos 142 dias de experimento.

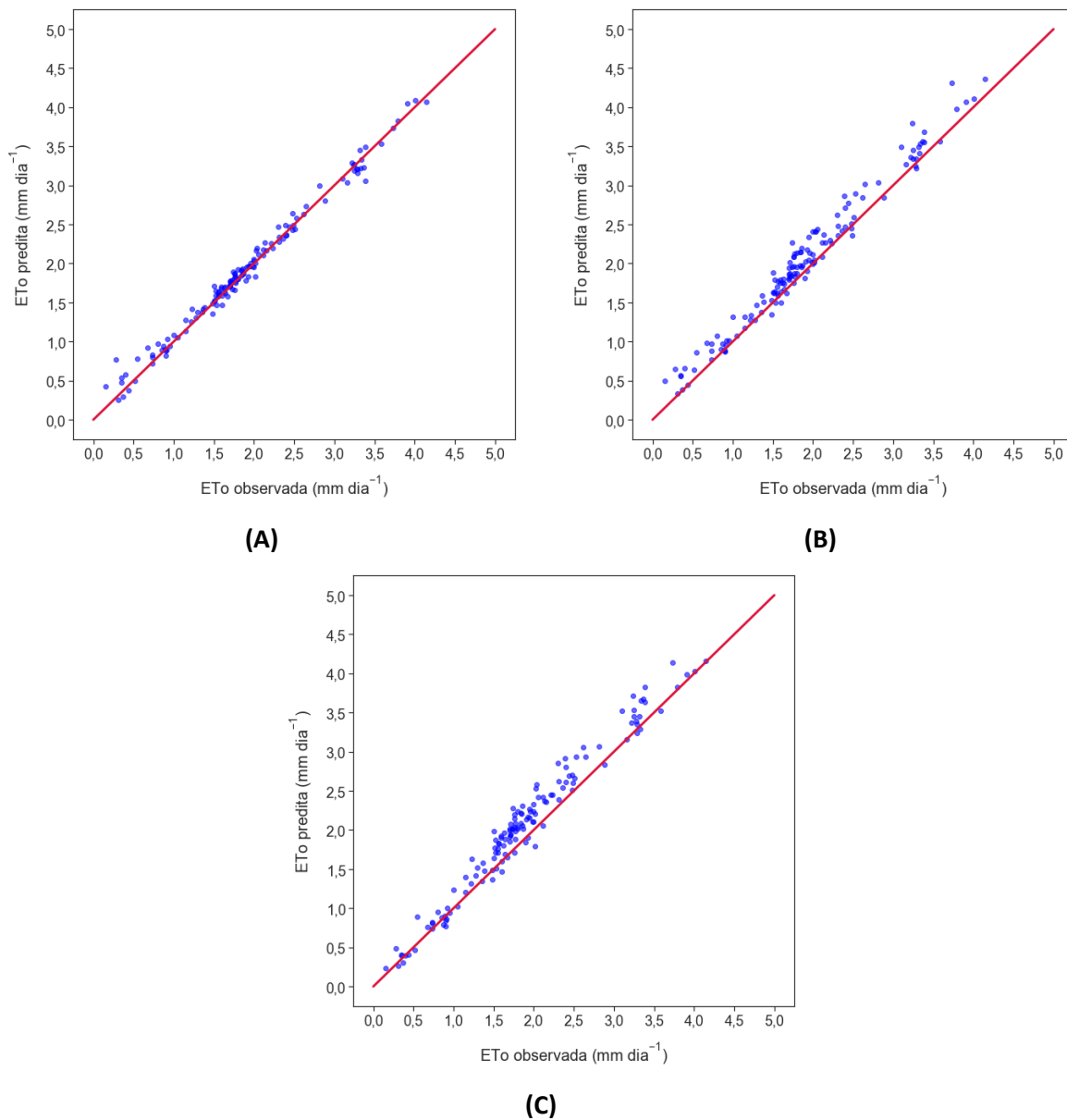


Figura 28. Comparação entre os valores de ETo observados durante o ciclo de cultivo usando o método da FAO-PM e os valores preditos pelos algoritmos com base em série de 10 dados: SVR (A), HGBR (B) e MLPR (C).

5.3.3. Manejo da irrigação

De acordo com os valores de K_c apresentados na Figura 24, a ET_c diária foi calculada para os quatro tratamentos, com o objetivo de determinar o volume de água repostado para a realização do manejo da irrigação. Pela Figura 29 pode-se conferir a variação da ET_c ao longo do ciclo de cultivo, na qual a ET_o determinada a partir do método padrão (PM) foi utilizada.

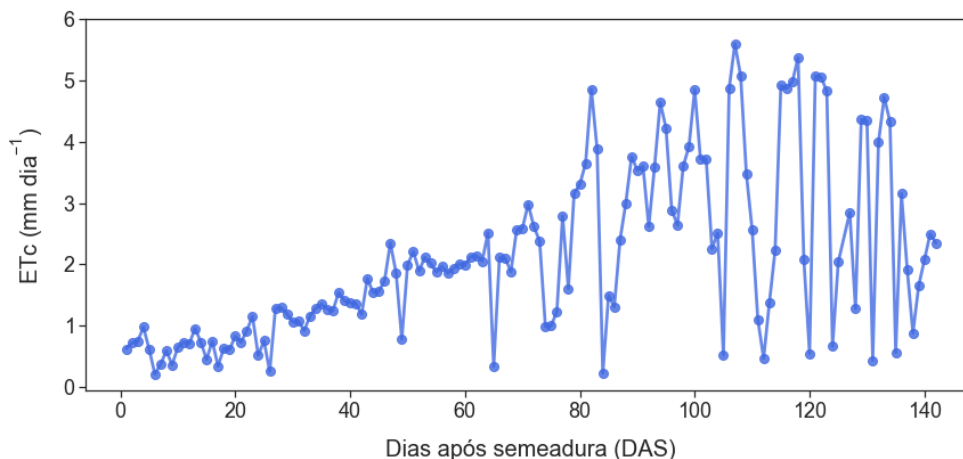


Figura 29. Evapotranspiração da cultura (ET_c) durante o ciclo de cultivo da cultura de milho (26/05/2022 – 14/10/2022).

No decorrer do experimento, a ET_c oscilou entre 0,21 e 5,60 mm dia^{-1} , com média de 2,12 mm dia^{-1} e acumulado de 298,87 mm ciclo^{-1} . A partir da Figura 29, pode ser observado os valores crescentes da ET_c conforme o ciclo de desenvolvimento, com os valores mais altos de evapotranspiração coincidindo ao período da floração.

Os valores diários de ET_c obtidos corroboram com os resultados semelhantes propostos por Fancelli (2015), o qual evidencia que durante o período compreendido entre o florescimento e os grãos farináceos, o consumo hídrico da planta pode aumentar entre 5,0 e 7,0 mm dia^{-1} . Além disso, segundo o autor, em clima quente e seco, o consumo de água raramente excede 3,0 mm dia^{-1} .

Como o milho foi semeado em um período caracterizado por baixos índices de radiação solar, bem como temperaturas menores do que a média, pode-se constatar que a demanda hídrica da cultura foi baixa durante o estágio vegetativo para as condições do experimento.

Em experimento realizado na cidade de Rio Largo, AL, com objetivo de avaliar o crescimento, desenvolvimento e produtividade da cultura do milho sob diferentes níveis de irrigação, Barbosa (2017) constatou um valor médio da ET_c de 4,0 mm dia^{-1} e uma ET_c total acumulada de 451,94 mm. Esses resultados, maiores do que os observados nessa pesquisa, se devem às condições ambientais em que as plantas foram submetidas, uma vez que no experimento em questão as temperaturas máxima e mínima do ar ao longo do ciclo variaram de 27 a 37,7 °C e 19,6 a 24 °C,

respectivamente, com umidade relativa do ar oscilando de 61,3 a 86,5%, o que resultou em maior consumo hídrico da cultura.

Pela Tabela 24 confere-se os valores da ETc média diária, ETc, lâminas via irrigação, precipitação e lâminas totais acumuladas em todo o ciclo para cada tratamento. É possível analisar que os tratamentos em que foram utilizados como base a estimativa da ETo pelos algoritmos HGBR e MLPR, apresentaram ETc mais distantes do valor obtido pelo método padrão, já que esses modelos apresentaram menor precisão em relação ao SVR, conforme mostrado na Tabela 23, com isso, para esses tratamentos as plantas consumiram 323,11 mm e 326,86 mm durante o ciclo, respectivamente, aproximadamente 25 mm a mais do que o observado para PM. Isso resultou em maiores lâminas d'água via irrigação para esses tratamentos, enquanto que as lâminas aplicadas para o PM e SVR foram similares.

Tabela 24. Valores da ETc média diária e ETc, lâminas via irrigação, precipitação e lâminas totais acumuladas referentes aos quatro tratamentos aplicados durante o ciclo de cultivo.

Tratamento	ETc	ETc acumulada	Lâmina via irrigação	Precipitação	Lâmina total
	mm dia ⁻¹			mm ciclo ⁻¹	
PM	2,12	298,87	338,22		553,39
SVR	2,15	302,80	343,65		558,82
HGBR	2,29	323,11	362,99	215,17	578,16
MLPR	2,32	326,86	367,18		582,35

Oliveira (2020) em experimento conduzido na cidade de Sete Lagoas (MG) com o objetivo de avaliar o efeito de diferentes lâminas de irrigação e da integração de duas cultivares de braquiária com o milho na produtividade de fitomassa para silagem e de grãos, constatou valor médio de ETc de 549,33 mm ao longo do ciclo com ETo média de 525 mm. Esses resultados foram superiores aos observados nessa pesquisa, uma vez que por ter sido semeado no mês de fevereiro, as plantas foram submetidas à valores mais altos de temperatura e radiação solar, o que resultou em altos valores de ETo, onde a máxima observada foi de 6,5 mm. De acordo com Alencar et al. (2011), a ETc reflete as variações da ETo, bem como das características (Kc) da cultura, o que pode ser observado nos experimentos.

No total, as plantas de milho receberam em média 568,18 mm de água ao longo do ciclo, valor esse considerado satisfatório de acordo com o intervalo de demanda da cultura relatado por Critchley e Siegert (1991). Devido ao baixo índice pluviométrico, a complementação de água por meio da irrigação foi essencial para promover um desenvolvimento contínuo dos elementos de rendimento do milho.

5.3.4. Respostas biométricas no desenvolvimento vegetativo do milho

Os valores médios referentes à altura das plantas (ALP), diâmetro do colmo (DIP) e altura de inserção da espiga (AIE) relativos a cada tratamento em estudo podem ser conferidos na Tabela 25.

Tabela 25. Valores médios referentes à altura da planta (ALP), diâmetro do colmo (DIP) e altura de inserção da espiga (AIE) de cada tratamento.

Tratamento	ALP	DIP	AIE
	cm	mm	cm
PM	106,2	22,48	58,00
SVR	108,0	22,68	58,30
HGBR	110,3	23,02	56,50
MLPR	108,9	24,88	59,45
Média	108,35	23,27	58,06
Desvio-padrão	5,49	1,96	3,36

A partir dos resultados obtidos, é possível analisar que o tratamento HGBR proporcionou maior altura das plantas, enquanto que PM apresentou menores médias de altura e diâmetro do colmo. Os maiores valores de diâmetro e altura de inserção da espiga foram observados no tratamento MLPR.

Em relação a essas variáveis, o desvio-padrão, medida estatística que indica o quanto os valores de um conjunto de dados estão dispersos em relação à média, foi baixo, isso indica que, aparentemente, esses componentes não variaram de forma significativa entre os quatro tratamentos.

O valor médio da altura das plantas considerando todo o experimento foi de 108,35 cm, valor considerado baixo, uma vez que a cultura de milho pode atingir 2 m de altura. Isso, provavelmente, se deve ao fato de que as plantas foram cultivadas em vasos de 50 L com altura de 38 cm, o que limitou o desenvolvimento adequado do sistema radicular e impactou diretamente no crescimento. Além disso, os fatores climáticos também podem ter influenciado, devido ao período de seca e baixos índices de radiação solar. De acordo com Taiz et al. (2013) e Santos (2016), plantas sob condições de baixa disponibilidade hídrica tendem a fechar seus estômatos, o que reduz a atividade fotossintética, o crescimento e a fixação de CO₂, portanto períodos de seca acarretam em decréscimo na altura das plantas.

Resultados semelhantes de diâmetro do colmo foram obtidos por Ben (2015), onde o híbrido de milho Dekalb 250 VT Pro 2 foi cultivado em Santa Maria, RS, com objetivo de avaliar o efeito de diferentes lâminas de irrigação, calculadas com base na ETc, e densidades de plantas sobre os componentes de crescimento, desenvolvimento e rendimento da cultura. Os valores obtidos pelo

autor para altura da planta e altura de inserção da espiga estão em desacordo com os resultados da presente pesquisa, uma vez que foram constatados valores médios de 2,38 e 1,15 m, respectivamente.

Em experimento conduzido na cidade de Botucatu, SP, por Silva (2018) com o objetivo de avaliar o efeito de diferentes lâminas de irrigação no desenvolvimento do milho, foram observados valores superiores de altura da planta, diâmetro do colmo e altura de inserção da espiga, com média de 2,01 m, 24,34 mm e 86,74 cm, respectivamente. Almeida (2016), em experimento conduzido em Piracicaba, SP, também encontrou resultados superiores, com média de 1,99 m para altura da planta e 86,17 cm para altura de inserção da espiga.

Conforme destacado por Fancelli e Dourado Neto (2000), o diâmetro do colmo desempenha papel fundamental na obtenção de rendimentos mais elevados de espiguetas, uma vez que quanto maior o seu diâmetro, maior a capacidade da planta de armazenar fotoassimilados essenciais para o crescimento e desenvolvimento de espiguetas. O colmo também atua como reservatório de sólidos solúveis, os quais são utilizados na formação dos grãos. Já a altura do dossel está intimamente relacionada à produtividade agrícola e garantia de um desenvolvimento saudável, de acordo com Silva (2006), devido ao fato de que plantas mais altas provavelmente passam por menores deficiências nutricionais ao longo do seu crescimento e acumulam maiores quantidade de reserva no colmo. Portanto, essas são variáveis que devem ser observadas para se obter uma estimativa do rendimento produtivo.

Após confirmação de que os dados seguem distribuição normal e apresentam homogeneidade de variâncias, os dados foram submetidos à ANOVA. A partir dos resultados da análise para as variáveis referentes ao desenvolvimento vegetativo das plantas de milho avaliadas na colheita, pode-se observar que, para as condições do experimento, não houve diferença significativa pelo teste F ao nível de 5% de probabilidade para todas as variáveis (ALP, DIP E AIE) em relação ao tratamento utilizado para o manejo da irrigação (Tabela 26).

Tabela 26. Resumo da análise de variância para as variáveis altura da planta (ALP), diâmetro do colmo (DIP) e altura de inserção da espiga (AIE) em plantas de milho expostas a diferentes manejos de irrigação.

Fator de variação	Quadrado médio			
	GL	ALP	DIP	AIE
Tratamentos	3	14,75 ^{ns}	6,05 ^{ns}	7,38 ^{ns}
Resíduos	16	33,05	3,42	12,02
CV (%)	---	5,31	7,95	5,97

Os resultados evidenciaram que as estimativas da ETo utilizando modelos de *machine learning* para o manejo da irrigação, em comparação com o método padrão de PM, não afetaram o desenvolvimento vegetativo de plantas de milho nas condições em que o experimento foi conduzido.

Portanto, em relação ao crescimento vegetativo do milho, a utilização de algoritmos com quantidade reduzida de variáveis de entrada mostrou efeito positivo como uma alternativa para a estimativa da ETo.

5.3.5. Respostas biométricas no desenvolvimento das espigas de milho

Os valores médios referentes à peso da espiga com palha (PEP), peso da espiga despalhada (PES), tamanho da espiga (TE) e diâmetro da espiga (DE) relativos a cada tratamento em estudo podem ser conferidos na Tabela 27.

Tabela 27. Valores médios referentes ao peso da espiga com palha (PEP), peso da espiga despalhada (PES), tamanho da espiga (TE) e diâmetro da espiga (DE) de cada tratamento.

Tratamento	PEP	PES	TE	DE
	g		cm	mm
PM	78,37	64,11	10,86	37,37
SVR	72,46	55,15	10,37	36,17
HGBR	75,01	58,59	9,97	38,19
MLPR	80,44	62,12	11,22	38,88
Média	76,57	59,99	10,61	37,65
Desvio-padrão	15,51	15,25	1,45	2,96

A partir dos resultados obtidos, é possível analisar que o tratamento SVR proporcionou menores valores de pesos da espiga com palha e despalhada, enquanto que MLPR apresentou maiores médias de peso da espiga com palha e tamanho e diâmetro da espiga.

Em relação à essas variáveis, o desvio-padrão, medida estatística que indica o quanto os valores de um conjunto de dados estão dispersos em relação à média, foi baixo, especialmente para tamanho e diâmetro da espiga, isso indica que, aparentemente, esses componentes não variaram de forma significativa entre os quatro tratamentos.

Valores semelhantes referentes ao tamanho da espiga foram observados por Almeida (2016), o qual realizou experimento em Piracicaba com o objetivo de avaliar o efeito de diferentes lâminas de irrigação (déficit e excesso) no decorrer do ciclo total e em fases fenológicas específicas da cultura do milho. O autor constatou tamanho médio de 12,84 cm, porém em relação ao diâmetro da espiga e pesos com e sem palha, os resultados obtidos foram maiores do que os da

presente pesquisa, com média de 45,25 mm, 182,22 g e 156,26 g, respectivamente, indicando que a fase de enchimento de grãos foi eficiente.

Segundo Pegorare et al. (2009), os baixos valores referentes à massa seca da espiga (PES) pode ser atribuído à diminuição da área foliar no período vegetativo, o que reduz a radiação absorvida pelas folhas, a fixação de CO₂, a produção de fotoassimilados, biomassa total, a massa e o número de grãos por espiga.

Ben et al. (2015), avaliando a resposta do rendimento de produção do milho safrinha à diferentes reposições de irrigação, observaram que o comprimento da espiga variou de 7,8 a 10,1 cm, corroborando com os resultados obtidos na presente pesquisa. Enquanto que, em experimento conduzido por Silva (2018) em Botucatu, SP, foram observados valores superiores de tamanho e diâmetro da espiga, com média de 17,06 cm e 51,89 mm, respectivamente.

De acordo com Kappes et al. (2009), a variável correspondente ao tamanho da espiga representa uma importante característica que pode influenciar diretamente no número de grãos por fileira e, conseqüentemente, na produtividade da cultura do milho, por isso a sua avaliação após a colheita é essencial.

Vale ressaltar que os resultados apresentados na Tabela 27, foram diretamente influenciados pelo desenvolvimento vegetativo do milho, uma vez que a atividade fotossintética e acúmulo de fotoassimilados que são deslocados para as espigas durante a fase reprodutiva são influenciados pela altura da planta, diâmetro do colmo e área foliar.

A Figura 30 mostra as espigas de milho com palha (A) e despalhadas (B) após a colheita referentes a cada um dos quatro tratamentos. Nela é possível confirmar que, de fato, o MLPR levou aos maiores valores de peso, tamanho e diâmetro da espiga, seguido do PM, enquanto que SVR apresentou espigas menos desenvolvidas.



Figura 30. Espigas com palha (A) e despalhadas (B) referentes aos tratamentos utilizados para o manejo da irrigação do milho.

Após confirmação de que os dados seguem distribuição normal e apresentam homogeneidade de variâncias, os dados foram submetidos à ANOVA. A partir dos resultados da análise para as variáveis referentes ao desenvolvimento das espigas de milho avaliadas na colheita, pode-se observar que, para as condições do experimento, não houve diferença significativa pelo teste F ao nível de 5% de probabilidade para todas as variáveis (PEP, PES, TE e DE) em relação ao tratamento utilizado para o manejo da irrigação (Tabela 28).

Tabela 28. Resumo da análise de variância para as variáveis peso da espiga com palha (PEP), peso da espiga despalhada (PES), tamanho da espiga (TE) e diâmetro da espiga (DE) em plantas de milho expostas a diferentes manejos de irrigação.

Fator de variação	Quadrado médio				
	GL	PEP	PES	TE	DE
Tratamentos	3	62,58 ^{ns}	78,12 ^{ns}	1,50 ^{ns}	6,77 ^{ns}
Resíduos	16	273,94	261,60	2,20	9,11
CV (%)	---	21,61	26,96	13,99	8,02

Os resultados evidenciaram que as estimativas da ETo utilizando modelos de *machine learning* para o manejo da irrigação, em comparação com o método padrão de PM, não afetaram o desenvolvimento das espigas de milho nas condições em que o experimento foi conduzido. Portanto, para as condições do experimento, a utilização desses algoritmos com quantidade reduzida de dados de entrada mostrou efeito positivo como uma alternativa para a estimativa da ETo.

5.3.6. Avaliação dos componentes de produção do milho

Os valores médios referentes ao número de grãos por espiga (Ngrãos), número de fileiras de grãos por espiga (NFG), peso de grãos (PG) e produtividade (Prod) relativos a cada tratamento em estudo podem ser conferidos na Tabela 29.

Tabela 29. Valores médios referentes ao número de grãos por espiga (Ngrãos), número de fileiras de grãos por espiga (NFG), peso de grãos (PG) e produtividade (Prod) de cada tratamento.

Tratamento	Ngrãos	NFG	PG	Prod
			g	kg ha ⁻¹
PM	289,90	15,00	47,53	2.988,74
SVR	226,60	13,70	38,16	2.399,16
HGBR	245,90	14,60	41,70	2.621,30
MLPR	258,00	15,60	43,52	2.736,09
Média	255,10	14,73	42,72	2.686,00
Desvio-padrão	73,41	1,53	12,31	773,83

A partir dos resultados obtidos, é possível analisar que o tratamento SVR proporcionou menores valores dos componentes de produção da cultura do milho, enquanto que PM apresentou maiores médias de número de grãos por espiga, número de fileiras de grãos por espiga, peso de grãos e produtividade.

As produtividades médias observadas na Tabela 29 estão de acordo com as simuladas pelo DSSAT (Tabela 16) considerando a mesma região, época de plantio e características do solo.

Em relação à essas variáveis, o desvio-padrão, medida estatística que indica o quanto os valores de um conjunto de dados estão dispersos em relação à média, foi baixo, especialmente para número de fileiras de grãos por espiga e peso de grãos, isso indica que, aparentemente, esses componentes não variaram de forma significativa entre os quatro tratamentos.

Em experimento conduzido por Silva (2018) em Botucatu, SP, para estudar o efeitos dos fatores lâminas de irrigação e doses de adubação nitrogenada e a sua interação sobre a produtividade da cultura do milho segunda safra, foram constatados valores superiores de número de grãos por espiga com média de 507 grãos, considerando que a região, época de cultivo e cultivares utilizadas foram distintas. Em relação ao número de fileiras de grãos, o autor obteve valores semelhantes ao da presente pesquisa. Silva et al. (2017) também obtiveram resultados superiores de números de grãos por espiga em experimento na cidade de Alegrete, RS, com valor médio de 528 grãos.

O rendimento da produção de milho varia de acordo com as condições climáticas, época de plantio e fatores relacionados ao solo, com isso altos valores de produtividades podem ser obtidos. Kisekka et al. (2017), trabalhando com diferentes lâminas de irrigação de pré-plantio e com lâmina diária de 5 mm, encontram produções que variaram de 11.083 a 11.450 kg ha⁻¹.

No experimento realizado em Piracicaba, SP, para avaliar o efeito de diferentes lâminas de irrigação no ciclo do milho, Almeida (2016) constatou produtividade média de 8.026,03 kg ha⁻¹, com valores que variaram de 3459,5 kg ha⁻¹ para tratamentos com restrição hídrico a 10.758,8 kg ha⁻¹. O autor obteve 368 como número médio de grãos por espiga, resultados superiores aos encontrados nessa pesquisa.

Ressalta-se que os processos envolvidos no crescimento das culturas possuem estreita relação com a disponibilidade de água. O déficit hídrico nas fases de maiores demandas pode inviabilizar a disponibilidade de alguns nutrientes essenciais ao desenvolvimento da cultura, já que esses elementos são absorvidos pelas raízes das plantas de forma mais eficiente estando em meio aquoso. O desenvolvimento e peso dos grãos podem ser afetados, também, pela redução da atividade fotossintética, já que a menor produção de carboidratos implica em menor quantidade de matéria seca nos grãos (PEGORARE et al., 2009).

Logo, pode ser observado que os valores dos componentes de rendimento do milho obtidos nesse experimento foram baixos comparados com resultados encontrados na literatura, indicando, novamente, que o tamanho do vaso foi um fator limitante no desenvolvimento e crescimento adequado da cultura, uma vez que o volume de solo explorado pelo sistema radicular foi pequeno, especialmente para absorção eficiente de água e nutrientes. Além disso, as condições climáticas e época de plantio também, provavelmente, não foram favoráveis ao pleno desenvolvimento.

Após confirmação de que os dados seguem distribuição normal e apresentam homogeneidade de variâncias, os dados foram submetidos à ANOVA. A partir dos resultados da análise para as variáveis referentes aos componentes de produção do milho avaliadas na colheita, pode-se observar que, para as condições do experimento, não houve diferença significativa pelo teste F ao nível de 5% de probabilidade para todas as variáveis (Ngrãos, NFG, PG e Prod) em relação ao tratamento utilizado para o manejo da irrigação (Tabela 30).

Tabela 30. Resumo da análise de variância para as variáveis número de grãos por espiga (Ngrãos), número de fileiras de grãos por espiga (NFG), peso de grãos (PG) e produtividade (Prod) em plantas de milho expostas a diferentes manejos de irrigação.

Fator de variação	Quadrado médio				
	GL	Ngrãos	NFG	PG	Prod
Tratamentos	3	3527,2 ^{ns}	3,18 ^{ns}	76,15 ^{ns}	301041 ^{ns}
Resíduos	16	5737,8	2,17	165,59	654644
CV (%)	---	29,69	10,00	30,12	30,12

Os resultados evidenciaram que as estimativas da ETo utilizando modelos de *machine learning* para o manejo da irrigação, em comparação com o método padrão de PM, não afetaram os componentes do rendimento da cultura de milho nas condições em que o experimento foi conduzido. Portanto, para as condições do experimento, a utilização desses algoritmos com quantidade reduzida de dados de entrada mostrou efeito positivo como uma alternativa para a estimativa da ETo.

5.3.7. Estimativa da produtividade com algoritmos de *machine learning*

Os modelos de *machine learning* (EXT e RF) desenvolvidos a partir das variáveis climáticas e dos dados simulados pelo DSSAT para predição da produtividade da cultura do milho foram alimentados com os dados obtidos no experimento de campo com o objetivo de avaliar o

desempenho desses algoritmos na estimativa do rendimento de grãos para o experimento conduzido.

Os modelos foram alimentados com os seguintes dados de entrada para todos os tratamentos: data = 146, Tmax = 28,63 °C, Tmin = 13,45 °C, Tmed = 19,86 °C, Uz = 47,06 km dia⁻¹, URmin = 40,90%, Rs = 14,75 MJ m⁻² dia⁻¹ e Precip = 215,17 mm. Em relação à lâmina total de irrigação e à ETo acumulada ao longo do ciclo de cultivo, os valores de entrada variaram de acordo com o tratamento aplicado (Tabela 31).

Tabela 31. Dados de entrada referentes à irrigação e ETo inseridos nos modelos de *machine learning* para estimativa da produtividade de milho no experimento de campo.

Tratamento	Irrig	ETo
	mm	
PM	338,22	266,22
SVR	343,65	270,74
HGBR	362,99	288,01
MLPR	367,18	291,88

Pela Tabela 32 pode-se conferir as produtividades estimadas pelos algoritmos EXT e RF em cada tratamento estudado em comparação com os valores reais observados no campo.

Tabela 32. Produtividades do cultivo de milho preditas pelos algoritmos de aprendizado de máquina para as condições observadas no experimento de campo.

Tratamento	Produtividades preditas		Produtividade observada
	kg ha ⁻¹		
	EXT	RF	Experimento
PM	2.789,83	2.367,26	2.988,74
SVR	2.859,71	2.368,55	2.399,16
HGBR	3.115,44	2.494,25	2.621,30
MLPR	3.183,64	2.688,97	2.736,09
Média	2.987,16	2.479,76	2.686,32

Pode ser observado que mesmo com altos valores de RMSE nas fases de teste e validação cruzada (Tabela 18), os modelos conseguiram prever a produtividade com desempenho satisfatório para as condições em que o experimento foi conduzido, onde o erro médio entre o resultado estimado em relação ao real foi de 400,29 kg ha⁻¹ para o algoritmo EXT e 206,57 kg ha⁻¹ para o RF.

Para as circunstâncias do experimento, o modelo RF obteve melhor performance na estimativa do rendimento de grãos de milho, enquanto que o EXT superestimou a produtividade, diferentemente do observado na Tabela 18, em que o EXT mostrou performance superior, isso pode ser justificado pela alta capacidade de predição do RF.

É possível analisar que os modelos relacionaram produtividades superiores com maiores valores de lâmina de irrigação e ETo, logo como no tratamento MLPR foi aplicado maior volume de água para repor as perdas da cultura por ET (Tabela 24), esse apresentou produtividade predita superior, seguido pelo HGBR.

Os resultados presentes na Tabela 31 confirmaram que a utilização de algoritmos de *machine learning* com variáveis climáticas como dados de entrada para a estimativa da produtividade de milho na região de Piracicaba apresentou efeito positivo e significativo em um experimento de campo, se estabelecendo como uma alternativa viável e prática ao modelo DSSAT, a qual pode contribuir nas tomadas de decisão em relação ao planejamento e gestão da safra.

5.4. Conclusões

Diante dos resultados obtidos e considerando as condições do presente estudos, pode-se concluir que:

- a. Os modelos aplicados (SVR, HGBR e MLPR) foram capazes de identificar relações lineares e não lineares entre ETo e dados meteorológicos, resultando em desempenho satisfatório da predição da ETo durante o ciclo de cultivo do milho;
- b. Não houve diferença significativa pelo teste F ao nível de 5% de probabilidade para todas as variáveis analisadas em relação ao tratamento utilizado para o manejo da irrigação;
- c. As estimativas da ETo utilizando modelos de *machine learning* para o manejo da irrigação, em comparação com o método padrão de PM, não afetou o crescimento, desenvolvimento e produção da cultura de milho; e
- d. Os algoritmos EXT e RF foram capazes de estimar os rendimentos produtivos de milho obtidos no experimento de campo de forma coerente.

Portanto, pode-se concluir que os algoritmos de aprendizado de máquina adotados, apresentaram-se como sendo ferramentas de alto potencial para predição da ETo, utilizada no manejo da irrigação, bem como, para a estimativa da produtividade de milho, para as condições climáticas da região de Piracicaba e, assim, se apresentam como alternativas na gestão eficiente e sustentável do sistema produtivo.

Referências

- ALENCAR, L. P.; SEDIYAMA, G. C.; MANTOVANI, E. C.; MARTINEZ, M. A. Tendência recentes nos elementos do clima e suas implicações na evapotranspiração da cultura do milho em Viçosa – MG. **Engenharia Agrícola**, v. 31, n. 4, p. 631–642, 2011.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. 1998. **Crop evapotranspiration: guidelines for computing crop water requirements**. Rome: FAO, 1998, 300 p. (FAO irrigation and drainage paper no. 56).
- ALMEIDA, B. M. **Déficit e excesso hídrico na cultura do milho (Zea mays L.) em ambiente protegido**. Tese. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 2016.
- BARBOSA, W. S. S. **Milho cultivado sob diferentes lâminas de irrigação e adubação nitrogenada**. Dissertação. Universidade Federal de Alagoas, Rio Largo, 2017.
- BASSO, B.; LIU, L. Chapter four - seasonal crop yield forecast: methods, applications, and accuracies. **Advances in Agronomy** 154, 201–255, 2019.
- BEN, L. H. B. **Influência de lâminas de irrigação e densidades de plantas sobre o cultivo do milho “safrinha”**. Dissertação. Universidade Federal de Santa Maria, Santa Maria, RS, 88p., 2015.
- BERNARDO, S.; MANTOVANI, E. C.; SILVA, D. D.; SOARES, A. A. **Manual de Irrigação**. 9. ed. Viçosa: UFV, 545 p., 2019.
- CHEN, Z.; ZHU, Z.; JIANG, H.; SUN, S. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. **Journal of Hydrology**, v. 591, p. 125–286, 2020.
- CHIA, M. Y.; HUANG, Y. F.; KOO, C. H. Support vector machine enhanced empirical reference evapotranspiration estimation with limited meteorological parameters. **Computers and Electronics in Agriculture**, v. 175, p. 1–12, 2020.
- CONAB. COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira de grãos, v. 10, Safra 2022/23 – Primeiro levantamento**, Brasília, p. 1–77, 2022.
- CRANE-DROESCH, A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. **Environmental Research Letters**, v. 13, n. 11, 114003, 2018.
- CRITCHLEY, W.; SIEGERT, C. **Water Harvesting Manual**. FAO Paper AGL/MISC/17/91, FAO, Rome, 1991.
- DIDONET, A. D.; RODRIGUES, O.; MARIO, J. L.; IDE, F. Efeito da radiação solar e temperatura na definição do número de grãos de milho. **Pesquisa Agropecuária Brasileira**, v. 37, p. 933–938, 2002.

- GRANATA, F.; GARGANO, R.; de MARINIS, G. Artificial intelligence based approaches to evaluate actual evapotranspiration in wetlands. **Science of the Total Environment**, v. 703, p. 1–15, 2020.
- EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Manual de métodos de análises de solo**. 2.ed. Rio de Janeiro: Embrapa Solos, 23 p., 2011.
- EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. **Sistema brasileiro de classificação de solos**. 2.ed. Rio de Janeiro: Embrapa Solos, Brasília: Sistema de Produção de Informação, 306 p., 2006.
- FANCELLI, A. L. Cultivo racional e sustentável requer maior conhecimento sobre planta de milho. **Revista Visão Agrícola**, v. 13, n. 13, p. 20–23, 2015.
- FANCELLI, A. L.; DOURADO NETO, D. Produção de milho. **Livraria e Editora Agropecuária**, Guaíba, 360 p., 2000.
- GREAVES, G. E.; WANG, Y. M. Effect of regulated deficit irrigation scheduling on water use of corn in southern Taiwan tropical environment. **Agricultural Water Management**, v. 188, p. 115–125, 2017.
- JING, W.; YASEEN, Z. M.; SHAHID, S.; SAGGI, M. K.; TAO, H.; KISI, O.; SALIH, S. Q.; AL-ANSARI, N.; CHAU, K. W. Implementation of evolutionary computing models for reference evapotranspiration modeling: short review, assessment and possible future research directions. **Engineering Applications of Computational Fluid Mechanics**, v. 13, p. 811–823, 2019.
- KAPPES, C.; CARVALHO, M. A. C.; YAMASHITA, O. M.; SILVA, J. A. N. Influência do nitrogênio no desempenho produtivo do milho cultivado na segunda safra em sucessão à soja. **Pesquisa Agropecuária Tropical**, v. 39, n. 3, p. 251–259, 2009.
- KIPKULEI, H. K.; BELLINGRATH-KIMURA, S.D.; LANA, M. et al. Modelling cropland expansion and its drivers in Trans Nzoia County, Kenya. **Modeling Earth Systems and Environment**, v. 8, p. 5761–5778, 2022.
- KISEKKA, I. et al. Optimizing preplant irrigation for maize under limited water in the High Plains. **Agricultural Water Management**, v. 187, p. 154–163, 2017.
- KLUTE, A. Water retention: laboratory methods. In: KLUTE, A. (Ed) **Methods of soil analysis**. 2. Ed. Madisson: American Society of Agronomy, Soil Science of America, p. 635–662, 1996. (Agronomy Monograph, 9).
- MIRANDA, R. A. Uma história de sucesso da civilização. Embrapa Milho e Sorgo. **A Granja**, v. 74, n. 829, p. 24-27, 2018.
- MISHRA, S.; MISHRA, D.; SANTRA, G. H. Applications of machine learning techniques in agricultural crop production: A review paper. **Indian Journal of Science and Technology**, v. 9(38), p. 1–14, 2016.

- MOHAMMADI, B.; MEHDIZADEH, S. Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. **Agricultural Water Management**, v. 237, p. 1–13, 2020.
- MUPANGWA, W.; CHIPINDU, L.; NYAGUMBO, I.; MKUHLANI, S.; SISITO, G. Evaluating machine learning algorithms for predicting maize yield under conservation agriculture in Eastern and Southern Africa. **SN Applied Sciences** 2, 952, 2020.
- OLIVEIRA, I. C. M. **Irrigação subótima na cultura do milho: efeitos na produção de silagem e grãos na segunda safra**. Dissertação. Universidade Federal de São João del-Rei, Sete Lagoas, AL, 92 p., 2020.
- OLIVEIRA, L. A.; MIRANDA, J. H.; COOKE, R. A. C. Water management for sugarcane and corn under future climate scenarios in Brazil. **Agricultural Water Management**, v. 201, p. 199–206, 2018.
- PALHARES, M. **Distribuição e população de plantas e produtividade de grãos de milho**. Dissertação. Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, SP, 90p., 2003.
- PEGORARE, A. B.; FEDATTO, E.; PEREIRA, S. B.; SOUZA, L. C. F.; FIETZ, C. R. Irrigação suplementar no ciclo do milho “safrinha” sob plantio direto. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 13, n. 3, p. 262–271, 2009.
- RAZA, A.; SHOAIB, M.; FAIZ, M. A.; BAIG, F.; KHAN, M. M.; ULLAH, M. K.; ZUBAIR, M. Comparative assessment of reference evapotranspiration estimation using conventional method and machine learning algorithms in four climatic regions. **Pure and Applied Geophysics**, v. 177, n. 9, p. 4479–4508, 2020.
- SANTOS, L. S. **Produtividade do milho em diferentes épocas de plantio, na região dos tabuleiros costeiros de alagoas**. Dissertação. Universidade Federal de Alagoas, Rio Largo, 2016.
- SILVA, D. A.; VITORINO, A. C. T.; SOUZA, L. C. F.; GONÇALVES, M. C.; ROSCOE, R. Culturas antecessoras e adubação nitrogenada na cultura do milho em sistema de plantio direto. **Revista Brasileira de Milho e Sorgo**, v. 5, p. 75–88, 2006.
- SILVA, G. U. D. A. et al. Manejo de irrigação via solo e clima na cultura do milho (*Zea mayz* L.) na região de Alegrete/RS. **Revista de Ciência e Inovação**, v. 2, n. 1, p. 101–112, 2017.
- SILVA, P. F. **Manejo da irrigação e da adubação nitrogenada no milho cultivado no outono/inverno em Botucatu-SP**. Tese. Universidade Estadual Paulista “Júlio de Mesquita Filho”, UNESP, Botucatu, SP, 91 p., 2018.
- SOUZA, G. M.; BARBOSA, A. M. Fatores de estresse no milho são diversos e exigem monitoramento constante. **Revista Visão Agrícola**, v. 13, n. 13, p. 30–34, 2015.
- TAIZ, L.; ZEIGER, E. Fisiologia Vegetal, 5ª Edição. **Artmed**, Porto Alegre, p. 954, 2013.

- van GENUCHTEN, M.T. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. **Soil Science Society of America Journal**, v.44, p. 892–898, 1980.
- WANG, S.; LIAN, J.; PENG, Y.; HU, B.; CHEN, H. Generalized reference evapotranspiration models with limited climatic data based on random forest and gene expression programming in Guangxi, China. **Agricultural Water Management**, v. 221, p. 220–230, 2019.
- YAMAÇ, S. S.; TODOROVIC, M. Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of available meteorological data. **Agricultural Water Management**. v. 228, 105875, 2020.

ANEXOS

ANEXO A. Variáveis climáticas obtidas no Posto Meteorológico da ESALQ/USP e dados simulados pelo modelo DSSAT/CERES-MAIZE utilizados para o desenvolvimento de modelos preditivos da produtividade de milho.

Ano	Data de semeadura	Tmed	Tmax	Tmin	Uz	URmin	Rs	Precip	Irrig	ETo	Prod
	dia juliano	°C			km dia ⁻¹	%	MJ m ⁻² dia ⁻¹		mm		kg ha ⁻¹
1992	32	22,30	28,40	16,20	174,31	54,04	13,60	458,80	122,40	431,00	4.280,00
1993	32	22,05	28,10	16,00	169,34	55,43	13,50	612,40	205,90	508,20	7.159,00
1994	32	22,90	29,00	16,80	181,36	55,77	14,90	552,30	181,20	453,30	4.708,00
1995	32	22,15	28,40	15,90	164,15	56,89	13,60	773,00	185,30	442,90	4.818,00
1996	32	22,65	29,00	16,30	176,81	53,88	14,90	458,10	238,50	486,40	5.589,00
1997	32	21,15	27,30	15,00	110,00	47,42	16,70	343,90	239,10	476,00	8.934,00
1998	32	21,70	27,40	16,00	98,20	55,88	15,50	673,30	158,40	428,10	10.764,00
1999	32	21,40	27,70	15,10	104,02	52,68	16,40	575,70	513,20	227,10	10.073,00
2000	32	21,70	28,10	15,30	110,76	50,30	16,40	292,00	277,40	443,80	7.820,00
2001	32	22,95	28,90	17,00	81,75	54,13	16,60	332,20	188,10	404,30	6.740,00
2002	32	23,25	29,30	17,20	85,86	53,25	17,20	558,80	164,70	425,40	8.522,00
2003	32	22,35	28,50	16,20	93,19	48,20	16,50	347,90	190,40	427,60	7.413,00
2004	32	21,10	27,00	15,20	112,66	52,84	15,30	518,90	140,30	478,30	10.112,00
2005	32	23,15	29,40	16,90	80,24	51,36	16,40	368,60	169,00	325,00	5.698,00
2006	32	21,45	27,70	15,20	88,94	54,72	16,60	366,70	263,50	477,00	10.103,00
2007	32	22,70	29,00	16,40	97,10	52,99	17,80	426,70	223,90	468,40	8.560,00
2008	32	21,50	27,10	15,90	96,53	62,74	16,00	509,90	160,60	467,20	9.389,00
2009	32	22,00	28,10	15,90	86,63	58,02	16,90	327,70	225,50	429,80	8.130,00
2010	32	22,30	28,50	16,10	70,22	51,65	17,50	354,40	183,30	422,90	9.506,00
2011	32	22,10	28,00	16,20	75,28	56,12	16,30	565,10	142,20	442,50	11.191,00
2012	32	23,55	29,60	17,50	80,46	54,77	20,00	492,70	177,70	455,50	10.975,00
2013	32	24,00	30,30	17,70	76,40	66,22	18,20	511,90	157,70	378,20	7.335,00
2014	32	23,65	30,20	17,10	89,28	56,59	20,60	252,80	199,00	408,40	8.670,00
2015	32	24,40	30,40	18,40	64,87	68,96	18,20	360,80	162,40	411,70	8.203,00
2016	32	24,35	30,40	18,30	111,84	68,42	17,40	512,10	182,30	418,60	7.872,00
2017	32	22,75	28,70	16,80	127,96	50,95	17,20	537,40	172,00	419,30	7.337,00
2018	32	23,15	29,60	16,70	127,49	45,89	17,70	328,60	251,50	440,10	8.477,00
2019	32	23,40	29,20	17,60	113,65	51,95	16,70	447,90	182,90	438,10	7.433,00
2020	32	21,95	28,60	15,30	135,26	46,67	17,10	531,40	304,90	514,00	8.468,00
2021	32	22,35	29,00	15,70	121,21	43,75	17,10	297,70	255,60	428,30	7.593,00
1992	146	19,85	26,60	13,10	207,75	48,43	12,90	452,60	102,60	255,30	542,00
1993	146	20,05	26,80	13,30	199,63	51,40	13,00	410,60	103,50	274,20	906,00
1994	146	20,70	28,70	12,70	221,37	41,04	14,80	122,00	149,80	221,70	482,00
1995	146	20,60	27,90	13,30	200,84	47,23	13,80	264,90	118,80	232,30	641,00
1996	146	19,85	27,10	12,60	200,52	46,42	14,40	432,70	103,00	277,40	1.047,00
1997	146	19,90	26,70	13,10	119,76	44,01	15,20	335,00	119,20	339,00	2.175,00
1998	146	19,65	25,90	13,40	122,65	51,70	14,40	358,20	211,50	455,00	5.088,00
1999	146	19,60	26,70	12,50	137,85	44,20	15,70	193,40	250,90	378,90	3.549,00
2000	146	20,10	27,20	13,00	121,20	43,87	15,20	309,20	131,30	301,60	2.261,00
2001	146	20,10	26,90	13,30	92,60	47,83	15,20	295,20	159,50	345,00	3.317,00
2002	146	20,53	27,62	13,43	111,67	45,36	14,80	141,80	142,60	250,40	1.852,00
2003	146	20,05	27,52	12,58	122,29	35,82	15,12	143,90	126,10	213,20	874,00
2004	146	19,53	26,28	12,78	125,16	47,27	14,54	324,10	144,20	323,10	2.428,00
2005	146	20,36	27,00	13,72	113,79	49,32	14,76	226,20	175,80	373,10	4.130,00
2006	146	20,07	27,43	12,70	109,06	44,93	16,46	190,40	136,80	317,00	3.231,00

2007	146	20,39	27,81	12,96	115,14	44,27	16,28	259,90	174,70	263,40	2.034,00
2008	146	20,07	26,88	13,26	113,27	50,59	16,10	304,90	161,90	370,60	3.576,00
2009	146	19,90	26,04	13,76	96,50	59,94	14,70	369,60	198,80	432,00	5.878,00
2010	146	20,27	27,63	12,91	102,73	42,21	16,79	247,70	208,10	357,00	3.858,00
2011	146	20,30	27,73	12,86	99,75	43,38	17,73	263,40	183,60	301,40	3.215,00
2012	146	20,93	28,25	13,60	90,43	49,42	18,59	247,80	310,50	451,60	6.959,00
2013	146	21,07	28,33	13,81	86,93	58,89	16,77	343,40	239,40	444,90	6.350,00
2014	146	21,08	28,77	13,39	102,69	47,20	18,75	126,50	218,90	345,00	3.373,00
2015	146	21,72	28,74	14,70	70,72	57,93	16,82	317,20	143,70	310,40	2.714,00
2016	146	20,15	27,37	12,92	139,64	41,42	15,86	368,90	125,90	286,90	1.948,00
2017	146	20,50	27,92	13,08	134,28	39,45	16,27	183,00	158,50	259,50	2.087,00
2018	146	20,69	27,83	13,55	133,98	42,65	15,17	332,80	110,90	261,40	1.412,00
2019	146	20,98	28,64	13,32	138,99	38,96	16,56	160,10	150,50	261,10	1.877,00
2020	146	21,46	29,39	13,53	146,51	36,06	15,92	147,90	134,70	212,70	1.067,00
2021	146	20,67	28,62	12,71	143,39	35,21	15,68	124,80	166,80	254,40	1.915,00
1992	274	23,50	29,80	17,20	212,64	50,66	17,20	826,70	112,80	469,10	6.376,00
1993	274	24,60	30,70	18,50	215,52	51,91	18,00	437,10	72,50	352,20	4.562,00
1994	274	25,45	31,60	19,30	226,37	49,59	19,00	775,90	117,70	474,60	6.989,00
1995	274	23,95	29,90	18,00	211,26	50,28	18,20	883,50	149,30	505,50	8.368,00
1996	274	24,00	29,50	18,50	175,19	56,22	17,40	947,80	80,30	464,10	9.013,00
1997	274	24,60	30,20	19,00	125,13	51,29	19,90	613,80	97,20	421,80	6.698,00
1998	274	23,50	29,00	18,00	120,14	55,27	19,10	823,40	105,10	429,50	8.280,00
1999	274	23,25	29,20	17,30	157,03	49,26	20,90	582,00	89,90	399,00	6.176,00
2000	274	24,50	30,40	18,60	113,43	52,09	20,30	718,70	73,60	457,60	9.218,00
2001	274	23,65	29,20	18,10	107,48	56,49	19,50	808,90	88,80	473,10	12.057,00
2002	274	24,90	30,70	19,10	133,60	54,51	19,40	663,50	96,20	402,80	8.055,00
2003	274	23,75	29,50	18,00	139,65	46,66	19,00	574,50	105,60	399,40	8.153,00
2004	274	23,25	28,50	18,00	113,91	58,02	17,40	619,90	116,60	436,30	10.009,00
2005	274	24,00	29,80	18,20	121,42	55,60	20,50	583,40	126,40	492,80	10.417,00
2006	274	24,15	29,50	18,80	121,76	60,63	20,20	783,90	80,50	427,00	9.682,00
2007	274	23,90	29,80	18,00	126,33	55,88	20,00	793,90	126,70	416,50	9.760,00
2008	274	23,75	29,50	18,00	122,81	57,72	20,20	475,80	97,20	389,90	7.181,00
2009	274	24,50	29,90	19,10	105,67	63,09	20,00	701,30	66,10	412,60	9.801,00
2010	274	24,25	30,10	18,40	106,63	52,75	21,20	750,60	121,60	423,70	9.110,00
2011	274	23,85	29,90	17,80	112,46	52,12	23,50	717,50	142,00	504,00	15.114,00
2012	274	25,50	31,90	19,10	105,35	58,00	23,60	546,00	106,40	448,20	11.983,00
2013	274	25,75	32,30	19,20	126,01	57,20	24,80	434,70	167,30	493,00	10.981,00
2014	274	26,30	33,10	19,50	116,36	54,12	24,90	469,20	123,80	400,20	5.689,00
2015	274	26,95	33,00	20,90	87,59	68,50	21,50	706,60	58,40	378,10	8.952,00
2016	274	24,35	30,30	18,40	148,38	48,38	21,00	807,20	66,90	533,50	10.972,00
2017	274	24,45	30,40	18,50	143,47	48,14	20,70	679,00	90,40	457,10	10.240,00
2018	274	24,80	30,80	18,80	146,88	48,74	21,40	603,50	160,30	488,80	9.861,00
2019	274	25,05	31,40	18,70	151,82	46,64	21,30	649,00	98,50	451,00	9.743,00
2020	274	25,05	31,50	18,60	148,54	42,94	21,30	516,10	80,10	423,30	8.973,00
2021	274	23,90	30,00	17,80	150,20	47,44	20,70	875,60	158,50	511,10	10.793,00

ANEXO B. Variáveis climáticas diárias referentes ao período de cultivo da cultura de milho em Piracicaba, SP (26/05/2022 – 14/10/2022).

Data	Dia de cultivo	Tmed	Tmax	Tmin	URmed	Uz	Rn	Precip	ETo (PM)
			°C		%	m s ⁻¹	MJ m ⁻² dia ⁻¹	mm	
26/05/2022	1	18,22	28,20	10,30	70,57	0,403	5,46	0	1,557
27/05/2022	2	19,02	29,20	11,30	70,73	0,423	6,31	0	1,802
28/05/2022	3	19,34	29,30	11,50	66,38	0,516	6,19	0	1,856
29/05/2022	4	22,79	31,00	17,20	57,19	0,667	7,01	0	2,442
30/05/2022	5	22,43	30,30	19,00	74,23	0,556	5,08	16,50	1,556
31/05/2022	6	18,89	22,00	16,40	87,46	0,473	1,62	3,00	0,522
01/06/2022	7	20,15	27,60	16,60	85,10	0,203	3,33	4,30	0,926
02/06/2022	8	20,25	27,80	16,40	83,02	0,361	5,32	0,30	1,479
03/06/2022	9	18,10	24,30	12,30	80,61	0,529	2,94	0	0,902
04/06/2022	10	18,76	28,10	12,50	71,35	0,434	5,64	0	1,617
05/06/2022	11	19,25	29,10	12,90	74,18	0,260	6,32	0	1,770
06/06/2022	12	19,94	28,50	13,30	77,82	0,417	5,73	0	1,641
07/06/2022	13	21,49	30,40	16,10	65,01	0,567	6,87	1,00	2,156
08/06/2022	14	20,36	27,70	16,30	81,04	0,433	5,65	0,30	1,597
09/06/2022	15	19,97	24,80	16,20	86,92	0,341	3,28	0,50	0,953
10/06/2022	16	19,31	25,90	14,20	80,92	0,564	5,34	5,30	1,534
11/06/2022	17	15,47	21,80	10,00	78,72	0,248	2,41	0	0,678
12/06/2022	18	12,07	22,20	5,00	65,49	0,284	4,87	0	1,224
13/06/2022	19	14,08	23,80	6,40	67,76	0,388	4,23	0	1,147
14/06/2022	20	15,45	25,70	8,10	66,27	0,430	5,55	0	1,516
15/06/2022	21	16,36	25,70	9,80	69,88	0,360	4,67	0	1,298
16/06/2022	22	17,65	30,30	9,50	69,81	0,290	5,94	0	1,596
17/06/2022	23	21,55	30,90	17,10	66,40	0,520	6,36	1,44	1,953
18/06/2022	24	18,12	25,80	12,30	83,05	0,410	3,08	0	0,853
19/06/2022	25	18,11	27,00	13,60	80,74	0,460	4,63	0,30	1,216
20/06/2022	26	15,89	22,40	11,70	86,45	0,260	1,51	0	0,397
21/06/2022	27	20,57	29,60	14,00	73,86	0,580	6,78	0	1,986
22/06/2022	28	20,45	30,80	12,10	66,75	0,570	6,38	0	1,948
23/06/2022	29	19,41	31,00	11,00	68,09	0,440	6,14	0	1,762
24/06/2022	30	18,52	31,20	10,10	67,99	0,270	5,51	0	1,526
25/06/2022	31	19,74	30,20	13,80	73,14	0,470	5,41	0	1,524
26/06/2022	32	17,17	25,70	10,10	72,32	0,420	4,42	0	1,273
27/06/2022	33	17,72	27,80	10,60	72,17	0,270	5,69	0	1,566
28/06/2022	34	18,94	28,70	12,30	75,86	0,400	6,16	0	1,705
29/06/2022	35	19,12	28,00	11,50	65,40	0,650	5,55	0	1,766
30/06/2022	36	16,95	28,60	8,80	67,44	0,310	5,99	0	1,629
01/07/2022	37	18,51	29,50	11,30	68,60	0,380	5,54	0	1,566
02/07/2022	38	20,54	31,20	13,10	67,09	0,610	6,30	0	1,903
03/07/2022	39	20,25	31,80	11,00	68,00	0,370	5,81	0	1,705
04/07/2022	40	18,62	30,50	9,80	69,90	0,440	5,88	0	1,638
05/07/2022	41	18,25	30,70	9,70	66,09	0,410	5,69	0	1,595
06/07/2022	42	18,81	27,80	12,60	68,11	0,500	4,45	0	1,362
07/07/2022	43	19,68	30,40	11,30	64,17	0,540	6,63	0	2,000
08/07/2022	44	20,51	30,60	11,50	62,17	0,480	5,32	0	1,715
09/07/2022	45	18,29	29,50	10,30	67,82	0,490	6,04	0	1,710
10/07/2022	46	19,64	28,90	12,50	64,83	0,670	5,91	0	1,871

11/07/2022	47	21,11	28,90	15,10	53,60	1,060	6,33	0	2,491
12/07/2022	48	22,50	29,80	17,80	60,97	0,710	5,37	0	1,949
13/07/2022	49	18,82	27,40	12,30	75,32	0,370	2,66	0	0,803
14/07/2022	50	20,61	29,40	13,90	64,51	0,670	6,24	0	2,015
15/07/2022	51	21,80	31,00	15,50	59,31	0,790	6,36	0	2,208
16/07/2022	52	22,58	30,80	15,80	57,03	0,750	4,70	0	1,862
17/07/2022	53	23,28	32,90	14,90	61,80	0,460	6,17	0	2,051
18/07/2022	54	22,31	32,30	14,90	69,04	0,500	6,23	0	1,930
19/07/2022	55	22,23	32,40	14,70	66,73	0,470	5,60	0	1,768
20/07/2022	56	21,05	30,50	13,20	66,89	0,540	5,81	0	1,834
21/07/2022	57	19,89	29,90	12,00	70,40	0,370	5,81	0	1,701
22/07/2022	58	19,56	31,10	11,10	62,74	0,350	5,93	0	1,755
23/07/2022	59	19,75	31,20	10,70	60,30	0,370	5,93	0	1,798
24/07/2022	60	19,74	31,00	10,80	62,60	0,280	5,93	0	1,755
25/07/2022	61	20,82	31,50	11,80	63,01	0,450	5,93	0	1,853
26/07/2022	62	20,82	31,30	11,90	62,97	0,400	5,93	0	1,843
27/07/2022	63	19,71	30,50	12,00	67,87	0,400	5,97	0	1,743
28/07/2022	64	21,35	30,90	14,30	57,66	0,650	6,19	0	2,111
29/07/2022	65	17,13	27,10	10,10	74,89	0,350	0,87	0,58	0,278
30/07/2022	66	15,31	25,20	6,30	63,75	0,430	6,21	0	1,728
31/07/2022	67	16,63	28,40	9,80	71,66	0,340	6,50	0	1,698
01/08/2022	68	19,91	30,80	11,50	60,83	0,330	4,88	0	1,506
02/08/2022	69	21,53	32,80	12,10	55,10	0,370	6,33	0	2,031
03/08/2022	70	21,54	32,40	12,70	55,53	0,420	6,20	0	2,022
04/08/2022	71	23,55	33,10	14,90	49,66	0,613	6,09	0	2,303
05/08/2022	72	22,71	32,80	14,10	60,71	0,690	5,69	0	1,995
06/08/2022	73	20,86	30,00	14,00	68,74	0,475	5,86	0	1,798
07/08/2022	74	19,86	24,60	16,30	76,13	0,570	1,93	0	0,736
08/08/2022	75	19,29	23,30	17,00	89,08	0,340	2,60	6,78	0,736
09/08/2022	76	19,09	24,40	13,90	86,07	0,719	2,95	11,43	0,903
10/08/2022	77	14,36	19,50	9,90	73,41	1,290	7,08	8,13	2,017
11/08/2022	78	15,64	22,70	11,00	75,89	0,600	4,11	0	1,149
12/08/2022	79	17,86	27,50	9,30	61,33	0,603	7,46	0	2,235
13/08/2022	80	19,29	31,40	10,70	63,53	0,447	8,08	0	2,316
14/08/2022	81	21,46	32,60	13,10	55,96	0,446	8,06	0	2,530
15/08/2022	82	23,29	31,20	17,20	43,91	0,967	8,39	0	3,337
16/08/2022	83	23,93	33,50	16,40	52,89	0,799	7,06	0	2,642
17/08/2022	84	18,50	23,00	14,60	81,91	0,428	0,12	1,02	0,153
18/08/2022	85	20,76	30,20	13,90	68,89	0,842	2,63	0,25	0,996
19/08/2022	86	13,02	18,30	10,00	72,76	0,952	2,60	0,25	0,872
20/08/2022	87	15,98	21,20	13,10	69,21	0,943	5,09	0	1,601
21/08/2022	88	18,21	26,20	13,00	66,42	0,578	6,74	0	1,998
22/08/2022	89	18,55	28,50	10,20	66,15	0,541	8,73	0	2,506
23/08/2022	90	19,36	29,50	11,40	66,32	0,440	8,14	0	2,360
24/08/2022	91	21,01	30,00	13,60	63,37	0,394	7,84	0	2,399
25/08/2022	92	21,97	29,20	16,10	59,21	0,322	5,29	0	1,744
26/08/2022	93	22,26	31,70	13,50	52,38	0,392	7,23	0	2,390
27/08/2022	94	22,88	33,10	12,40	49,84	0,534	9,20	0	3,096
28/08/2022	95	23,01	34,40	16,50	56,02	0,913	8,34	0	2,817
29/08/2022	96	15,38	23,20	10,80	66,48	1,016	6,66	0	1,923
30/08/2022	97	16,41	23,60	11,00	64,56	0,854	5,54	0	1,761
31/08/2022	98	17,93	28,50	8,50	57,26	0,467	8,14	0	2,401

01/09/2022	99	19,79	30,90	10,90	54,16	0,481	8,51	0	2,615
02/09/2022	100	23,39	33,70	14,80	46,39	0,646	9,22	0	3,234
03/09/2022	101	21,05	33,10	12,60	60,65	0,872	7,94	0	2,482
04/09/2022	102	17,19	24,30	13,10	68,31	0,941	8,67	1,27	2,479
05/09/2022	103	19,29	27,40	14,90	67,61	0,756	4,68	0	1,506
06/09/2022	104	19,84	25,60	15,70	75,84	0,775	5,37	0	1,673
07/09/2022	105	18,21	25,40	15,00	85,60	0,532	1,39	10,11	0,348
08/09/2022	106	24,37	33,40	18,60	57,28	0,622	9,76	0	3,248
09/09/2022	107	26,01	36,20	17,50	49,69	0,600	10,71	0	3,731
10/09/2022	108	24,80	36,30	14,70	53,24	0,834	9,65	0	3,385
11/09/2022	109	19,24	28,20	12,70	68,72	0,636	7,91	0	2,314
12/09/2022	110	21,28	30,10	16,30	70,11	0,519	5,60	0	1,712
13/09/2022	111	21,08	27,00	18,10	83,30	0,470	2,43	10,00	0,731
14/09/2022	112	18,50	21,60	16,40	92,56	0,465	1,06	10,00	0,311
15/09/2022	113	18,17	22,60	15,60	81,09	0,521	3,03	0	0,915
16/09/2022	114	17,17	23,70	10,50	73,74	0,809	4,72	0	1,485
17/09/2022	115	17,57	25,20	11,80	58,96	1,257	10,63	0	3,277
18/09/2022	116	19,45	30,40	10,20	60,26	0,636	11,06	0	3,245
19/09/2022	117	22,50	33,40	15,70	64,20	0,790	10,95	0	3,320
20/09/2022	118	23,77	31,00	17,80	65,56	0,669	11,20	0	3,581
21/09/2022	119	21,04	28,20	18,50	76,43	0,505	4,71	13,50	1,389
22/09/2022	120	18,91	22,70	14,00	91,95	0,419	1,10	0	0,365
23/09/2022	121	17,49	27,00	8,80	59,06	0,535	11,96	0	3,384
24/09/2022	122	19,03	28,50	10,70	58,44	0,646	11,34	0	3,367
25/09/2022	123	21,33	32,20	16,00	66,21	0,641	11,12	0	3,218
26/09/2022	124	18,40	22,20	16,60	92,91	0,510	1,70	8,60	0,443
27/09/2022	125	20,51	27,60	16,50	89,72	0,611	5,21	7,10	1,360
28/09/2022	126	16,70	20,30	14,60	93,92	0,544	0	16,00	0
29/09/2022	127	18,49	23,40	15,50	75,99	0,760	6,43	0	1,898
30/09/2022	128	17,96	24,40	15,20	85,26	0,461	3,37	9,90	0,888
01/10/2022	129	21,75	30,30	16,10	76,75	0,385	10,95	0,50	3,161
02/10/2022	130	22,12	30,70	14,60	69,37	0,580	10,88	0	3,288
03/10/2022	131	18,24	23,80	15,30	83,13	0,357	1,11	8,90	0,345
04/10/2022	132	21,77	30,40	14,90	71,21	0,494	11,26	0	3,326
05/10/2022	133	23,43	33,00	16,10	67,25	0,507	13,70	0	4,145
06/10/2022	134	26,05	35,10	19,40	63,32	1,089	11,88	12,41	4,007
07/10/2022	135	21,28	29,10	14,80	71,82	0,689	1,02	0	0,550
08/10/2022	136	22,21	30,40	15,80	65,09	0,408	10,80	0	3,284
09/10/2022	137	22,48	33,00	18,60	74,20	0,560	7,52	38,90	2,136
10/10/2022	138	21,15	26,30	18,10	89,79	0,359	3,71	1,50	1,049
11/10/2022	139	23,92	29,90	18,70	78,42	0,497	6,72	0	2,116
12/10/2022	140	23,25	29,40	18,50	71,40	0,521	9,23	0	2,885
13/10/2022	141	24,91	35,10	16,60	64,33	0,605	11,95	5,10	3,785
14/10/2022	142	24,25	34,40	18,20	65,98	0,763	12,72	0	3,912

ANEXO C. Valores referentes as leituras da tensão da água no solo utilizando-se um tensímetro, ao potencial mátrico do solo e à umidade volumétrica do solo em função do potencial mátrico observados durante o período de cultivo da cultura de milho em Piracicaba, SP.

Data	Leitura	Potencial mátrico	Umidade volumétrica do solo
	bar	cm	cm ³ cm ⁻³
26/05/2022	-0,080	127,72	0,239
30/05/2022	-0,122	170,55	0,224
01/06/2022	-0,068	115,49	0,243
03/06/2022	-0,084	132,31	0,236
06/06/2022	-0,105	153,73	0,229
08/06/2022	-0,126	172,83	0,223
10/06/2022	-0,130	176,91	0,222
14/06/2022	-0,117	163,66	0,225
20/06/2022	-0,165	210,85	0,214
22/06/2022	-0,152	202,56	0,216
24/06/2022	-0,173	223,36	0,212
27/06/2022	-0,248	300,25	0,201
30/06/2022	-0,373	427,87	0,191
01/07/2022	-0,418	473,15	0,188
04/07/2022	-0,572	630,44	0,179
06/07/2022	-0,700	760,25	0,172
08/07/2022	-0,506	562,62	0,182
11/07/2022	-0,589	647,37	0,176
13/07/2022	-0,725	786,28	0,171
15/07/2022	-0,686	746,64	0,174
18/07/2022	-0,733	794,21	0,171
20/07/2022	-0,131	180,58	0,225
22/07/2022	-0,557	615,25	0,177
25/07/2022	-0,782	844,86	0,168
27/07/2022	-0,832	895,63	0,167
29/07/2022	-0,865	929,38	0,166
01/08/2022	-0,859	923,11	0,166
03/08/2022	-0,877	941,56	0,165
05/08/2022	-0,860	924,53	0,166
08/08/2022	-0,757	818,64	0,169
12/08/2022	-0,109	157,84	0,227
15/08/2022	-0,640	699,44	0,174
17/08/2022	-0,823	886,54	0,167
19/08/2022	-0,808	871,19	0,167
22/08/2022	-0,848	912,41	0,166
24/08/2022	-0,885	949,31	0,165
26/08/2022	-0,879	943,40	0,165
31/08/2022	-0,451	506,65	0,186
02/09/2022	-0,547	605,21	0,180
05/09/2022	-0,816	879,27	0,167
09/09/2022	-0,572	630,24	0,178
12/09/2022	-0,783	845,30	0,168
16/09/2022	-0,075	123,50	0,239
19/09/2022	-0,237	288,93	0,204
23/09/2022	-0,130	180,22	0,224