

CENTRO UNIVERSITÁRIO FEI
BRENNO TONDATO DE FARIA

**EVAPOTRANSPIRAÇÃO DE REFERÊNCIA MODELADA USANDO IOT E
ALGORITMOS DE APRENDIZAGEM DE MÁQUINA**

São Bernardo do Campo

2021

BRENNO TONDATO DE FARIA

**EVAPOTRANSPIRAÇÃO DE REFERÊNCIA MODELADA USANDO IOT E
ALGORITMOS DE APRENDIZAGEM DE MÁQUINA**

Dissertação de Mestrado apresentada ao Centro
Universitário da FEI para obtenção do título de
Mestre em Engenharia Elétrica com ênfase em
Inteligência Artificial Aplicada à Automação e
Robótica, orientado pelo Prof. Dr. Plínio Tho-
maz Aquino Júnior.

São Bernardo do Campo

2021

Tondato de Faria , Brenno.

Evapotranspiração de Referência Modelada Usando IoT e Algoritmos de Aprendizagem de Máquina / Brenno Tondato de Faria . São Bernardo do Campo, 2020.

109 p.

Dissertação - Centro Universitário FEI.

Orientador: Prof. Dr. Plinio Thomaz Aquino Junior.

1. Evapotranspiração de referência. 2. Agricultura de precisão. 3. Aprendizagem de máquina. 4. Satélite. I. Thomaz Aquino Junior, Plinio, orient. II. Título.

Aluno: Brenno Tondato de Faria

Matrícula: 118302-9

Título do Trabalho: Evapotranspiração de Referência Modelada Usando IoT e Algoritmos de Aprendizagem de Máquina.

Área de Concentração: Inteligência Artificial Aplicada à Automação e Robótica

Orientador: Prof. Dr. Plinio Thomaz Aquino Junior

Data da realização da defesa: 15/12/2020

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

A banca foi realizada no dia 15 de dezembro às 15:00 horas, e se iniciou com a apresentação do aluno, que foi muito boa, e seguiu para a arguição, onde o aluno respondeu a todas as questões de forma adequada demonstrando conhecimento pleno do tema. A aprovação foi por unanimidade.

São Bernardo do Campo, 15 / 12 / 2021.

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. Plinio Thomaz Aquino Junior

Ass.: _____

Prof. Dr. Reinaldo Augusto da Costa Bianchi

Ass.: _____

Prof. Dr. André Torre Neto

Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO

REPROVADO

VERSÃO FINAL DA DISSERTAÇÃO

APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA EXAMINADORA

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

Ao planeta terra que em sua vasta fatura
nos proporciona os desafios que possibilitam o
avanço ao conhecimento.

“Quanto melhor suas ações expressam as paixões que as impulsionam, mais uma figura é admirável”

Leon Battista Alberti

AGRADECIMENTOS

Agradeço ao professor doutor Plínio T. Aquino Jr. pela orientação e todo o auxílio em todas as etapas desta pesquisa.

Agradeço ao professor doutor Rodrigo Filev Maia pelo convite e aceitação da minha participação no programa de mestrado e o projeto SWAMP.

Agradeço ao professor doutor Reinaldo Bianchi pela coordenação e auxílio prestado durante o projeto SWAMP.

Agradeço a Fundação Educacional Inaciana pela oportunidade e bolsa de estudos concedida para a realização desta pesquisa.

Agradeço a Comissão Europeia e à Rede Nacional de Pesquisa - RNP, pela oportunidade da bolsa auxílio concedida para a realização desta pesquisa.

Agradeço aos meus pais, que com muita paciência conseguiram lidar com a minha ansiedade, aconselharam e guiaram para completar mais esta etapa da vida. Faço aqui um agradecimento especial a minha Mãe que sempre se esforçou muito para garantir meus estudos.

Agradeço a minha Tia Marly Stuardo que com muita ternura esteve me auxiliando em todo o desenvolvimento textual desta dissertação.

À minha família, pelo apoio, colaboração e empatia em todos os momentos deste projeto.

A equipe da secretária de pós graduação strict sensu, que sempre me apoiaram e orientaram nas questões administrativas do programa de mestrado.

Aos meus colegas de pesquisa Sergio Monteleone, Rafael Gomes Alves e Gilberto Souza pelas discussões que foram muito produtivas e certamente auxiliaram no desenvolvimento desta pesquisa.

Agradeço por fim todos aqueles que, de maneira direta ou indireta, contribuíram para a realização desta pesquisa.

RESUMO

A Organização das Nações Unidas para Alimentação e Agricultura (FAO) estima um crescimento populacional que chega a 11.2 bilhões até ano de 2100, o que sem dúvida contribuirá para o crescimento da demanda por produtos agrícolas alimentares e não alimentares, tornando a otimização de recursos hídricos indispensável. Neste contexto, o parâmetro de evapotranspiração de referência determinado pelo método FAO-PM destaca-se. Contudo, uma medida precisa necessita de diversos parâmetros climáticos, que podem não estar disponíveis em algumas regiões rurais, neste sentido, uma solução promissora são abordagens que se utilizam da menor quantidade de parâmetros climáticos, que podem ser medidos por satélites e estações meteorológicas da região e modelos de aprendizagem de máquina. Nesta pesquisa os modelos MLP (*Multilayer perceptron*) e SVM (*Support Vector Machines*) foram utilizados para modelar a evapotranspiração de referência a partir de dados de satélites e estações meteorológicas sob duas abordagens: a abordagem local, onde os modelos foram treinados e testados em um local de treinamento, e a abordagem regional, onde os modelos treinados no local de treinamento foram aplicados em um local de teste, em dois experimentos: um em uma região de clima temperado e outra em uma região de clima tropical. Os resultados indicam que o modelo MLP se sobressaiu diante ao modelo SVM em todas as simulações realizadas, no qual os modelos treinados com os parâmetros relativos a temperatura e radiação obtiveram as métricas R^2 de 0.6568, $RMSE$ de 0.1103 e MAE de 0.0882 para o experimento da região de clima temperado e métricas R^2 de 0.7391, $RMSE$ 0.1266 e MAE de 0.1063 para a região de clima tropical na abordagem local, o que demonstra o potencial de uso de apenas estes parâmetros para a modelagem da evapotranspiração de referência. Já na abordagem regional o modelo MLP pode ser aplicado com sucesso, o qual obteve as métricas R^2 de 0.7158, $RMSE$ de 0.1592, e MAE de 0.1428, contudo, no experimento de clima temperado, os resultados da aplicação foram insatisfatórios, demonstrando que para as condições daquele local os modelos não puderam ser aplicados.

Palavras-chave: Evapotranspiração de referência, Agricultura de precisão, Aprendizagem de máquina, Satélites.

ABSTRACT

The Food and Agriculture Organization of the United Nations (FAO) estimates that population growth will reach 11.2 billion by the year 2100, which will contribute to food and agricultural product demand, making irrigation optimization essential. In this context is highlighted the evapotranspiration parameter determined by the FAOPM method. However, a precise measurement of this parameter requires several climatic parameters that can not be available in rural areas, in which, a promise solution belongs in approaches that use just a few climatic parameters which can be obtained by satellites and weather stations combined with machine learning models. In this research, the MLP (Multilayer perceptron) e SVM (Support Vector Machines) models were used to model the reference evapotranspiration with satellite and weather stations data under two approaches: the local approach where the models are trained and tested on the training location, and the regional approach where the models trained on the training location were applied on a test location. These approaches were applied in two experiments: the first on a temperate climate zone, and the second on a tropical climate zone. The results indicate that the MLP model stood out when compared with the SVM model in all tests realized, in which, the models trained with the climatic parameters of temperature and radiation obtained the metrics of R^2 of 0.6568, $RMSE$ of 0.1103, and MAE de 0.0882 for the temperate climatic zone experiment and metrics R^2 of 0.7391, $RMSE$ 0.1266, and MAE of 0.1063 for the tropical climate zone experiment on the first approach which demonstrate the potential of using only these parameters to model de evapotranspiration. For the second approach the MLP model could be applied on the tropical climate zone in which the metrics R^2 of 0.7158, $RMSE$ of 0.1592, and MAE of 0.1428 were obtained. Yet the result obtained by the models applied on the temperate climate zone was inconclusive which indicates that for the conditions of this location the models can't be applied with the second approach.

Keywords: Reference evapotranspiration, Precision agriculture, Machine learning, Satellites

LISTA DE ILUSTRAÇÕES

Figura 1	–	Interações da radiação solar	25
Figura 2	–	Mapa global da Classificação <i>Köppen</i>	27
Figura 3	–	<i>Framework</i> conceitual da Internet das Coisas (IoT)	29
Figura 4	–	Ideia Geral do Modelo <i>Support Vector Machines</i> (SVM)	33
Figura 5	–	Modelo de Neurônio Biológico e Modelo Perceptron	35
Figura 6	–	Problemas linearmente separáveis e problemas não linearmente separáveis	35
Figura 7	–	Modelo de <i>Multlayer Perceptron</i>	36
Figura 8	–	Diagrama de processos executados para a execução da metodologia empregada	53
Figura 9	–	Mapa de Classificações Climáticas	54
Figura 10	–	Diagrama da avaliação dos parâmetros climáticos	55
Figura 11	–	Diagrama do estudo, treinamento e validação dos modelos MLP e SVM no local de treinamento	57
Figura 12	–	Diagrama de teste dos modelos MLP e SVM no local de teste	58
Figura 13	–	Localização do experimento de Brotas	62
Figura 14	–	Localização do experimento de MATOPIBA	63
Figura 15	–	Avaliação do local de teste e treinamento para o experimento de Brotas	65
Figura 16	–	Avaliação do local de teste e treinamento para o experimento de MATOPIBA	66
Figura 17	–	Gráficos de avaliação dos parâmetros climáticos INMET e a ETo calculada no experimento Brotas	67
Figura 18	–	Gráficos de avaliação dos parâmetros climáticos NASA/POWER e INMET no experimento Brotas	68
Figura 19	–	Gráficos de avaliação dos parâmetros climáticos INMET e a ETo calculada no experimento MATOPIBA	70
Figura 20	–	Gráficos de avaliação parâmetros climáticos NASA/POWER e INMET no experimento MATOPIBA	72
Figura 21	–	Avaliações dos modelos que utilizaram-se do conjunto <i>C1</i> no experimento de Brotas	75
Figura 22	–	Avaliações dos modelos que se utilizaram do conjunto <i>C2</i> no experimento de Brotas	76

Figura 23 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de Brotas	77
Figura 24 – Avaliações dos modelos que se utilizaram do conjunto $C1$ no experimento de MATOPIBA	79
Figura 25 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de MATOPIBA	80
Figura 26 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de MATOPIBA	81
Figura 27 – Gráficos de avaliação de parâmetros climáticos NASA/POWER e estação local e ET_o do local de treinamento e teste do experimento Brotas	83
Figura 28 – Avaliações dos modelos que utilizaram-se do conjunto $C1$ no experimento de Brotas	84
Figura 29 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de Brotas	86
Figura 30 – Avaliações dos modelos que utilizaram-se do conjunto $C3$ no experimento de Brotas	87
Figura 31 – Gráficos de avaliação de parâmetros climáticos NASA/POWER e estação local e ET_o do local de treinamento e teste do experimento MATOPIBA	88
Figura 32 – Avaliações dos modelos que se utilizaram do conjunto $C1$ no experimento de MATOPIBA	89
Figura 33 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de MATOPIBA	90
Figura 34 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de MATOPIBA	91
Figura 35 – Coeficiente de determinação do experimento de Brotas e MATOPIBA	94
Figura 36 – Métricas de todos os experimentos de Brotas	95
Figura 37 – Métricas de todos os experimentos de MATOPIBA	96

LISTA DE TABELAS

Tabela 1	– Tabela de codificação e descrição de perfis climáticos	28
Tabela 2	– Modelos de aprendizagem de máquina utilizados na literatura pesquisada .	52
Tabela 3	– Resultados do coeficiente de determinação entre a ET_o calculada e os parâmetros climáticos INMET para o experimento de Brotas	68
Tabela 4	– Resultados do coeficiente de determinação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER para o experimento de Brotas	69
Tabela 5	– Resultados do coeficiente de determinação entre a ET_o calculada e os parâmetros climáticos INMET para o experimento de MATOPIBA	71
Tabela 6	– Resultados do coeficiente de determinação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER para o experimento de MATOPIBA	73
Tabela 7	– Tabela de conjuntos de parâmetros utilizados	73
Tabela 8	– Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C1$	76
Tabela 9	– Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C2$	77
Tabela 10	– Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C3$	78
Tabela 11	– Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C1$	79
Tabela 12	– Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C2$	81
Tabela 13	– Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C3$	82
Tabela 14	– Avaliações de relação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER e a evapotranspiração de referência do local de teste e treinamento no experimento de Brotas	84
Tabela 15	– Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C1$	85

Tabela 16 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C2$	86
Tabela 17 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C3$	87
Tabela 18 – Resultados da avaliação de relação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER e a ET_o do local de treinamento e teste do experimento de MATOPIBA	89
Tabela 19 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C1$	90
Tabela 20 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C2$	91
Tabela 21 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C3$	92

LISTA DE ABREVIATURAS

ANA	Agência Nacional de Águas
ANFIS	<i>Adaptive Neuro-Fuzzy Inference System</i>
ANFISGP	<i>Adaptive neuro-fuzzy inference systems with grid partitioning</i>
ANFISSC	<i>Adaptive neuro-fuzzy inference systems with subtractive clustering</i>
CNN	<i>Convolutional neural network</i>
DENFIS	<i>dynamic evolving neural-fuzzy inference system</i>
DENFISOF	<i>off-line dynamic evolving neural-fuzzy inference system</i>
DENFISON	<i>on-line dynamic evolving neural- fuzzy inference system</i>
DESI	<i>Differential Evolution and System Identification</i>
ELM	<i>extreme learning machine</i>
ER	<i>evolutionary regression</i>
FAO	<i>Food and Agriculture Organization</i>
FAOPM	Método FAO Penman-Monteith
FFNN	<i>Feed Forward Neural Network</i>
GBM	<i>Gradient-Boosting Machine</i>
GEP	<i>gene expression programming</i>
GFF	<i>Generalized Feedforward</i>
GLM	<i>Generalized Linear Model</i>
GRNN	<i>generalized regression neural network</i>
IoT	Internet das Coisas
KSOM	<i>Kohonen self-organizing map</i>
LR	<i>Linear Regression</i>
LSSVM	<i>Least square support vector machine</i>
LSSVR	<i>Least Squares Support Vector Regression</i>
LSSVRGSA	<i>least-square support vector regression with a gravitational search algorithm</i>
M5RT	<i>M5 model tree</i>
MARS	<i>multivariate adaptive regression splines</i>
MGA	<i>Modified Genetic Algorithm</i>
ML	<i>Machine Learning</i>
MLNR	<i>multi nonlinear regression</i>
MLP	<i>Multilayer Perceptron</i>

MLR	<i>Multilinear Regression</i>
MVRVM	<i>Multivariate Relevance Vector Machine</i>
PNN	<i>Probabilistic Neural Network</i>
RBNN	<i>radial basis neural networks</i>
RF	<i>Random forest</i>
RFID	Indetificação por Radio Frequência
SVM	<i>Support Vector Machines</i>
XGBoost	<i>Extreme gradient boosting</i>

SUMÁRIO

1	Introdução	18
1.1	Motivação	19
1.2	Objetivos	19
1.3	Justificativa	20
1.4	Organização do texto	20
2	Conceitos Fundamentais	22
2.1	Agricultura	22
2.2	A Irrigação	22
2.3	A Evapotranspiração de Referência	23
2.3.1	Radiação solar	24
2.3.2	Temperatura	25
2.3.3	Umidade do Ar	26
2.3.4	Velocidade do vento	26
2.4	Classificação <i>Köppen</i>	26
2.5	Internet das Coisas	28
2.6	Agricultura de Precisão	29
2.6.1	Sensoriamento remoto a partir de satélites	30
2.7	Inteligência Artificial	30
2.8	Modelos de aprendizagem de máquina	31
2.8.1	Aprendizado Supervisionado	32
2.8.2	O Modelo Máquina de Vetores de Suporte	32
2.8.3	Redes Neurais Artificiais	33
2.8.4	<i>Perceptron</i>	34
2.8.5	<i>Perceptron de Múltiplas Camadas</i>	36
3	Revisão Bibliográfica Sistemática	37
3.1	Trabalhos Relacionados	38
4	Metodologia	53
4.1	Métricas de Avaliação	59
4.2	Parâmetros Climáticas e Bases de Dados	59
4.3	Experimentos	61

5	Resultados	64
5.1	Definição dos Locais de Treinamento	64
5.1.1	Experimento de Brotas	64
5.1.2	Experimento de MATOPIBA	65
5.2	Conjunto de Parâmetros	66
5.2.1	Experimento de Brotas	66
5.2.2	Experimento de MATOPIBA	69
5.2.3	Definição dos conjuntos de parâmetros climáticos	73
5.3	Arquitetura dos Algoritmos	73
5.3.1	Arquitetura SVM	74
5.3.2	Arquitetura MLP	74
5.4	Treinamento e validação dos Modelos no Local de Treinamento	74
5.4.1	Experimento de Brotas	74
5.4.2	Experimento de MATOPIBA	78
5.5	Aplicação dos modelos no local de teste	82
5.5.1	Experimento de Brotas	82
5.5.2	Experimento de MATOPIBA	88
6	Discussão	93
7	Conclusão	98
	REFERÊNCIAS	100
	ÍNDICE	110

1 Introdução

Segundo o relatório elaborado pela *Food and Agriculture Organization* (FAO), as projeções de crescimento populacional chegam à 9.7 bilhões de habitantes em 2050, 10.8 bilhões em 2080 e 11.2 bilhões em 2100, e que o crescimento populacional vem acompanhado do crescimento de renda e urbanizações, que por sua vez contribuem para crescimento da demanda por alimentos e produtos agrícolas não alimentares (FAO, 2018). O pesquisador Perera et al. (2014) elabora que o crescimento da demanda agrícola impacta diretamente nos recursos naturais do planeta, e que, hoje, a agricultura é responsável pelo uso de 70% da água disponível.

No Brasil, a agricultura é responsável por 67% do consumo de água, sendo 66.1% consumido pela irrigação, conforme o atlas da irrigação elaborado pela ANA (2017). Este fato agrava ainda mais as projeções realizadas pela Agência Nacional de Águas (ANA) no qual se projeta um crescimento potencial de 45% na área irrigada do país até 2030 (ANA, 2019). Sendo assim, novas técnicas são necessárias para se utilizar a irrigação com eficiência, a fim de atender as demandas populacionais por alimentos e produtos agrícolas não alimentares.

Uma irrigação sustentável depende de alguns fatores, dentre os quais, é destacável o parâmetro da evapotranspiração de referência (SCHERER-WARREN; RODRIGUES, 2013). Este parâmetro pode ser encontrado tanto no planejamento de recursos hídricos (KISI; DEMIR, 2016), quanto no processo de balanceamento hídrico dos cultivares (LIU et al., 2017), além de abranger todos os processos de transformação de água líquida em vapor (SCHERER-WARREN; RODRIGUES, 2013).

Conforme elaborado por Subedi e Chávez (2015), a evapotranspiração de referência pode ser obtida por diversos métodos, contudo, o método elaborado por Allen et al. (1998), denominado de Método FAO Penman-Monteith (FAOPM), recomendado pela FAO, é o mais popular entre todos. A pesquisa elaborada por Pereira et al. (2015) evidencia que esse método se destaca devido a sua estabilidade em diferentes localizações geográficas e a baixa necessidade de calibração de parâmetros.

No entanto, a evapotranspiração de referência determinada pelo método de FAOPM é um dos métodos que necessita da maior quantidade de parâmetros climáticos, os quais muitas vezes são medidos por estações meteorológicas, ou ferramentas semelhantes, como torres de fluxo de calor (GRANATA, 2019).

No contexto brasileiro, verifica-se a dificuldade de se obter a evapotranspiração de referência, a qual entre outros fatores, é afetada pela falta de ferramentas, o que muitas vezes

cabe ao agricultor confiar em estações meteorológicas geralmente distantes de sua região (DO-ORENBOS; PRUITT, 1977), ou ainda, em outras ocasiões, cabe ao agricultor determinar a medida, confiando apenas em sua experiência, o que pode gerar desperdícios indesejados.

Apesar disso novos paradigmas tecnológicos, como a IoT, têm se mostrado promissores, onde uma alternativa para o uso de estações meteorológicas é o uso de imagens de satélites, contudo mostram-se limitadas, restringindo a utilização plena apenas em grandes áreas e em grandes espaços de tempo.

Ainda assim, a utilização de novas abordagens de modelos que estimam parâmetros climáticos a partir de imagens de satélites possibilitam a aquisição de dados meteorológicos em qualquer lugar do globo a uma escala espacial mais precisa, porém o uso destes dados possuem defasagens quanto a alguns parâmetros climáticos.

Apesar disso o uso de modelos de aprendizagem de máquina pode ajudar a resolver este problema, uma vez que, ao serem treinados, estes modelos podem modelar a evapotranspiração de referência com uma precisão semelhante ao método FAOPM, o que pode oferecer uma solução para agricultores em regiões agrícolas que possuem baixo acesso às ferramentas mais robustas, como estações meteorológicas e torres de fluxo de calor, ou até mesmo servir como alternativa a locais que possuem estas ferramentas, porém não possuem um método para determinar a evapotranspiração.

1.1 Motivação

A evapotranspiração de referência é um importante parâmetro que contribui para a irrigação sustentável. Para se obter a evapotranspiração de referência, o método FAOPM criado por Allen é recomendado, o qual, para uma medida precisa, é necessário uma série de parâmetros climáticos obtidos a partir de estações meteorológicas, que muitas vezes não estão disponíveis. Com isso, abordagens que se utilizam de parâmetros climáticos medidos a partir de satélites e o uso de algoritmos de aprendizagem de máquina podem representar uma alternativa em regiões de baixo acesso à ferramentas de medição.

1.2 Objetivos

Neste cenário, esta pesquisa visa estudar os modelos de aprendizagem de máquina Máquina Vetores de Suporte e Perceptron Múltiplas Camadas, para modelar a evapotranspiração

de referência a partir de parâmetros climáticos medidos por satélites. Como sub-objetivos destacam-se:

- a) Avaliação da acurácia dos modelos de aprendizagem de máquina em diferentes locais que possuam perfis climáticos distintos.
- b) Comparação das acurácias dos modelos entre si, e também entre o método FA-OPM.
- c) Avaliação dos dados meteorológicos obtidos por satélites, comprando-os com os dados obtidos por estações meteorológicas.
- d) Avaliação dos modelos de aprendizagem de máquina com diferentes combinações de parâmetros climáticos.
- e) Avaliação a aplicabilidade de um modelo treinado em uma localidade, em outra localidade da mesma região climática.

1.3 Justificativa

O desafio em atender as demandas populacionais por alimentos e produtos agrícolas torna o *status quo* da irrigação global incerto, sendo necessária a adoção de novas tecnologias, onde a evapotranspiração de referência é um importante parâmetro para uma irrigação sustentável, na medida em que ele é utilizado nos diferentes processos de irrigação. Neste sentido, a literatura apresenta diferentes métodos para se obter um resultado satisfatório, dentre os quais o método FAOPM é o mais adequado para tal aplicação, posto que é o método que oferece os resultados mais consistentes, apesar de necessitar de diversos parâmetros climáticos que não estão disponíveis com facilidade. Ainda assim, o uso de parâmetros climáticos obtidos a partir de tecnologias de sensoriamento IoT baseadas em satélites junto uso de modelos de aprendizagem de máquina para modelar a evapotranspiração de referência pode mostrar-se uma solução viável.

1.4 Organização do texto

A presente pesquisa está organizada da seguinte maneira: o capítulo dois apresenta a fundamentação teórica dos temas abordados nesta pesquisa, no qual é elaborado um breve contexto histórico da agricultura e da irrigação. Também são abordados os temas relacionados com a evapotranspiração de referência e os parâmetros climáticos necessários para obtê-la.

Ainda neste capítulo, é apresentado um breve histórico do paradigma de IoT, bem como a descrição de agricultura de precisão e o sensoriamento remoto a partir de satélites, após estas apresentações é explicado o conceito de classificações climáticas de *Köppen*. Por último, são apresentados os conceitos de inteligência artificial e aprendizagem de máquina, finalizando-o com os fundamentos dos modelos utilizados.

O capítulo três apresenta a revisão bibliográfica sistemática realizada sobre o uso de algoritmos de aprendizagem de máquina e parâmetros climáticos medidos por satélites, para a modelagem da evapotranspiração de referência.

O capítulo quatro apresenta os trabalhos relacionados obtidos a partir da revisão bibliográfica realizada.

O capítulo cinco apresenta a metodologia para se obter os modelos de aprendizagem de máquina capazes de modelar a evapotranspiração de referência a partir de parâmetros climáticos medidos por satélites. Também nesse capítulo são elaboradas as bases de dados utilizadas, bem como as métricas de avaliação dos modelos e os experimentos selecionados.

O capítulo seis elabora os resultados obtidos a partir da metodologia realizada.

O capítulo sete discute os resultados em comparação com a literatura.

O capítulo oito elabora uma conclusão dos resultados obtidos e aponta trabalhos futuros para a extensão desta pesquisa.

Por fim, o capítulo nove encerra a pesquisa, com as referências bibliográficas.

2 Conceitos Fundamentais

Neste Capítulo são apresentados os conceitos fundamentais para a compreensão dos temas abordados nesta pesquisa, bem como a fundamentação teórica que serve de base para o desenvolvimento.

2.1 Agricultura

Quando os nômades, há mais de dez mil anos, deixaram progressivamente a caça e a coleta de alimentos para então cultivar seus próprios grãos, originou-se a prática da agricultura. Durante toda a Antiguidade, Idade Média e Renascença, esta prática significou um desafio, que apesar de alguns povos conseguirem garantir sua segurança alimentar, muitos ainda sofriam com a fome devido as práticas precárias e rudimentares (EHLERS, 2009).

No século XVIII, na Europa Ocidental, deu-se o início das práticas da agricultura moderna, o que possibilitou ao homem começar a produção de alimentos em larga escala. Do século XIX até os dias atuais, tais práticas foram ganhando força através das inovações tecnológicas, adubos químicos e sementes geneticamente melhoradas, levando a uma produção de crescimento exponencial. Contudo, a produção de grandes safras cedeu lugar a outras preocupações, relacionadas a problemas sociais, econômicos e ambientais (EHLERS, 2009).

Daker (1973) elabora que uma produção econômica é dada por um conjunto de fatores, sendo estes a fertilidade do solo, o calor e a água. A fertilidade do solo pode ser facilmente controlada pelo agricultor, a partir da escolha de um solo que permita diversas culturas e pelo controle da adubação. O calor pode ser controlado escolhendo-se a melhor época para o plantio e, a água, que apesar de ser dada pela escolha da época certa, pode ser suprida pela prática da irrigação.

2.2 A Irrigação

Assim como a agricultura a irrigação é intrínseca à história da própria humanidade, tendo suas primeiras aparições já nas grandes civilizações de mais de 4.000 anos atrás, como os egípcios e os mesopotâmios (DAKER, 1973).

Esta atividade milenar pode ser compreendida como um sistema de controle da deficiência de água, no qual muitas vezes acaba por substituir grande parte, se não toda, a água oriunda

de precipitações. Por sua vez este controle oferece algumas vantagens como: produção garantida, maior produção por unidade de área, obtenção de mais de uma colheita por ano, obtenção de colheitas fora da época normal e, controle de ervas daninhas (DAKER, 1973).

Dentre os diversos fatores que contribuem para a irrigação, o planejamento hídrico se destaca. Isto porque este planejamento permite avaliar a adequação de fontes de água, necessidade de armazenamento e a determinação da capacidade dos sistemas de irrigação. Entretanto, a estimativa incorreta da quantidade de água pode levar a desperdícios, controle inadequado da umidade na zona raiz, inundações, salinidade e lixiviação de nutrientes do solo (SAVVA; FRENKEN, 2002).

O parâmetro da quantidade de água para a irrigação possui como primeiro passo de seu planejamento a determinação da quantidade de água consumida pela cultura. Este parâmetro por sua vez é representado pela evapotranspiração da cultura (SAVVA; FRENKEN, 2002).

2.3 A Evapotranspiração de Referência

Ainda que seja de difícil separação, dada a sua natureza simultânea, o processo de transpiração e evaporação pode ser visto em um único processo chamado de evaporação (ALLEN et al., 1998).

Diversos fatores afetam a transpiração e a evaporação, dentre eles encontram-se os parâmetros climáticos, características da cultura, aspectos ambientais e gerenciais do local (ALLEN et al., 1998). Ademais, os parâmetros climáticos são considerados os principais fatores que contribuem para a evapotranspiração, sendo eles a temperatura e umidade ar, a velocidade do vento e a radiação solar (ALLEN et al., 1998).

A evapotranspiração de uma cultura pode ser definida em dois processos distintos, sendo estes a determinação da evapotranspiração de referência (ET_o) e a determinação do coeficiente da cultura (k_c). A evapotranspiração de referência, por sua vez, pode ser definida como a taxa de evapotranspiração de uma superfície extensa, completamente coberta por grama viva a uma altura de 8 a 15 centímetros, crescendo continuamente e sem a falta de água (DOORENBOS; PRUITT, 1977).

Diversos métodos empíricos foram desenvolvidos para estimar a evapotranspiração de referência. No documento de irrigação e drenagem número 24 estabelecido pela *Food and Agriculture Organization* (FAO) em 1977 é possível encontrar quatro métodos, sendo eles: *Blaney-Criddle*; Radiação; evaporação em tanque Classe A e *Penman*, dentre os quais, o método

Penman mostrou melhores resultados para a cultura de grama viva, adotada como referência (DOORENBOS; PRUITT, 1977).

Posteriormente, Monteith adaptou o método de *Penman* para obter a evapotranspiração de uma determinada cultura, dando origem ao método *Penman-Monteith* (MONTEITH, 1965). Este método se destacou devido a sua consistência em diferentes regiões climáticas tornando-se um método de referência (ALLEN et al., 1998). Contudo, alguns de seus parâmetros são difíceis de serem obtidos, além de ser necessário defini-lo a cada estágio de plantio (ALLEN et al., 1998).

Após o congresso realizado pela FAO em 1990, o método *Penman-Monteith* combinado a cultura de grama viva, utilizada como referência, deu origem ao método chamado FAOPM desenvolvido por (ALLEN et al., 1998), denotado pela equação 1, a qual junta a consistência do método *Penman-Monteith* com a baixa necessidade de parâmetros. Nesta pesquisa o método FAOPM será utilizado para se obter a evapotranspiração de referência.

$$ET_o = \frac{0.408\Delta (R_n - G) + \gamma \frac{900}{T + 273} u_2 (e_s - e_a)}{\Delta + \gamma (1 + 0.34u_2)} \quad (1)$$

No qual, ET_o representa a evapotranspiração de referência em (mm/dia), R_n o saldo de radiação diário em (MJ/m^2), G o fluxo de calor no solo em ($MJ/m^2/dia^{-1}$), γ a constante psicrométrica em ($kPa/^\circ C$), T a temperatura média em ($^\circ C$) e u_2 a velocidade do vento em (m/s) a uma altura de 2 metros, e_s a pressão de saturação de vapor de água em (kPa), e_a a pressão do vapor atual em (kPa) e por fim Δ a declividade da curva de pressão de vapor em relação a temperatura ($kPa/^\circ C$).

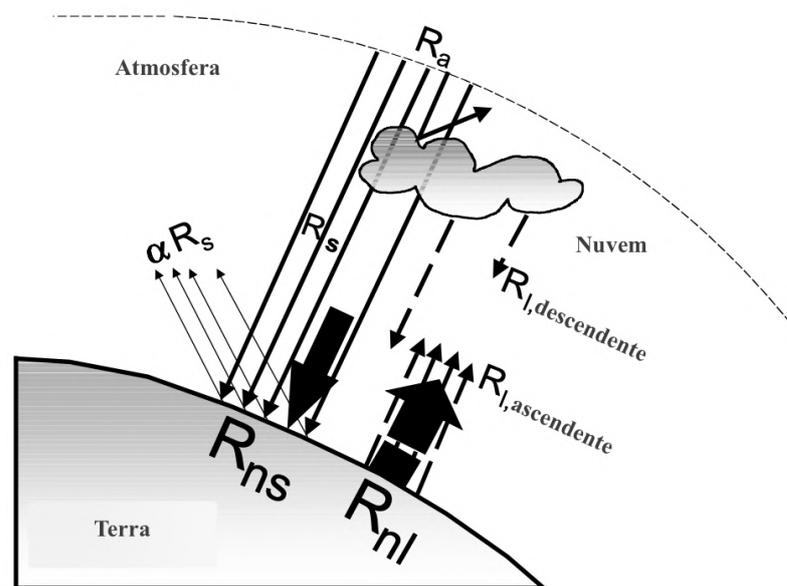
Outra vantagem deste método considera que a evapotranspiração de referência pode ser modelada a partir de parâmetros climáticos, como temperatura do ar, radiação solar, umidade relativa e velocidade do vento, medidos por estações meteorológicas (ALLEN et al., 1998). O relacionamento da evapotranspiração de referência com os parâmetros climáticos citados acima é elaborado nas subseções a seguir.

2.3.1 Radiação solar

A evapotranspiração é um processo determinado pela quantidade de energia disponível para vaporizar a água. A radiação solar é a maior fonte de energia, além de possuir a capacidade de vaporizar a água em grandes volumes. A quantidade de radiação solar que pode atingir uma

superfície de evaporação é determinada pela localização e hora do dia. Isto porque ao longo do ano as diferentes posições do sol em relação a terra mudam, alterando o potencial de radiação em diferentes localidades e estações. Além do mais, o potencial de radiação solar que chega a atingir uma superfície de evaporação ainda depende de fatores como, a turbidez da atmosfera e a presença de nuvens, nas quais a maior parte da radiação será refletida e absorvida (ALLEN et al., 1998). A figura 1 ilustra as diferentes interações da radiação solar, as quais impactam na evapotranspiração.

Figura 1 – Interações da radiação solar



Fonte: Adaptado de Allen et al., 1998

2.3.2 Temperatura

A radiação solar absorvida pela atmosfera e o calor emitido pela superfície terrestre são fatores que contribuem para o aumento da temperatura do ar. O calor sensível, gerado a partir da temperatura, fornece energia para a cultura e funciona como um controlador do fluxo de evapotranspiração, por exemplo, em um dia quente e ensolarado a taxa de evapotranspiração é maior do que em um dia frio e nublado (ALLEN et al., 1998).

2.3.3 Umidade do Ar

Assim como a energia vinda da radiação solar e o ar ao redor da cultura representam os principais fatores que contribuem para a evaporação da água, a diferença entre a pressão do vapor e o ar do local são os principais fatores que contribuem para a remoção do vapor (ALLEN et al., 1998).

Um exemplo que pode ser observado consiste em plantios bem irrigados de regiões áridas cujo cultivos consomem grandes quantidades de água, uma vez que há uma quantidade abundante de energia unido ao alto poder da atmosfera. Já em locais de clima tropical, a alta entrada de energia e a alta umidade do ar leva à redução da demanda de evapotranspiração, pois, uma vez que o ar esteja próximo da saturação, o fluxo de água tende a diminuir (ALLEN et al., 1998).

2.3.4 Velocidade do vento

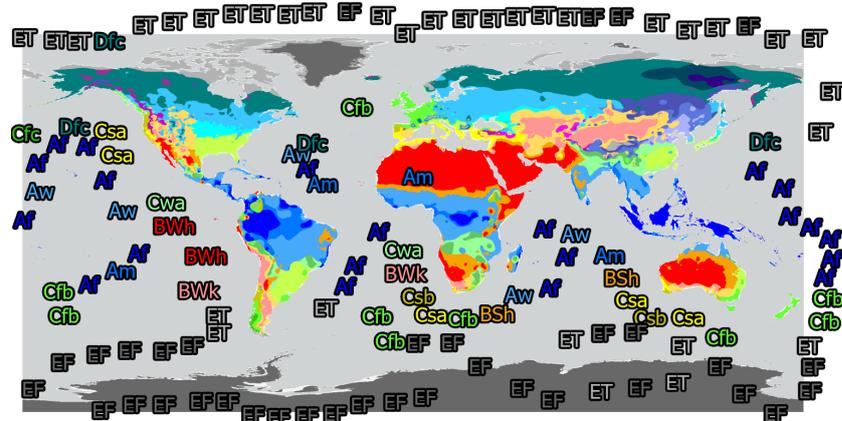
O processo de remoção do vapor d'água depende majoritariamente do vento e da turbulência de ar, no qual grandes porções de ar são transferidas para a superfície de evaporação. Se o ar ao redor da cultura não for substituído de forma contínua, por ar mais seco, tanto a força necessária para a remoção do vapor d'água quanto a evapotranspiração irão diminuir (ALLEN et al., 1998).

2.4 Classificação Köppen

O mapa de classificações climáticas proposto por Köppen e Geiger (1930) é uma referência para se determinar a região climática de uma determinada localidade. Contudo a falta de acesso a ferramentas que possibilitariam uma determinação precisa faz com que este mapa continue a receber melhorias e atualizações (PEEL; FINLAYSON; MCMAHON, 2007).

Diversas pesquisas utilizam-se deste mapa para que possam ser feitas análises com base no perfil climático das regiões de estudo, que são definidos a partir de parâmetros climáticos como temperatura e precipitação obtidos a partir de estações meteorológicas espalhadas no mundo todo e análises em escalas mensais e anuais (PEEL; FINLAYSON; MCMAHON, 2007).

Figura 2 – Mapa global da Classificação Köppen



Fonte: Peel, Finlayson e McMahon, 2007

As fronteiras de cada perfil climático é definido a partir de critérios, nos quais cada um pode ser subdividido em critérios de segunda e terceira ordem, onde são caracterizadas as regiões sub-climáticas (PEEL; FINLAYSON; MCMAHON, 2007). A Tabela 1 exhibe as codificações de cada critério.

A partir de alguns critérios como média anual de precipitação, média anual de temperatura, temperatura do mês mais quente, temperatura do mês mais frio, incidência de precipitação no mês mais seco (PEEL; FINLAYSON; MCMAHON, 2007), é possível definir fronteiras de regiões e sub-regiões climáticas, dando origem a mapas como o exibido na Figura 2.

Tabela 1 – Tabela de codificação e descrição de perfis climáticos

1o	2o	3o	Descrição
A			Tropical
	f		- Floresta Tropical
	m		- Monção
	w		- Savana
B			Arido
	W		- Deserto
	S		- Estepe
		h	- Quente
		k	- Fria
C			Temperado
	s		- verão seco
	w		- inverno seco
	f		- sem estação seca
		a	- Verão quente
		b	- Verão morno
		c	- Verão frio
D			Frio
	s		- Verão seco
	w		- Inverno seco
	f		- sem estação seca
		a	- Verão quente
		b	- Verão morno
		c	- Verão frio
		d	- Inverno muito frio
E			Polar
	T		- Tundra
	F		- Geada

(Do Autor, 2020)

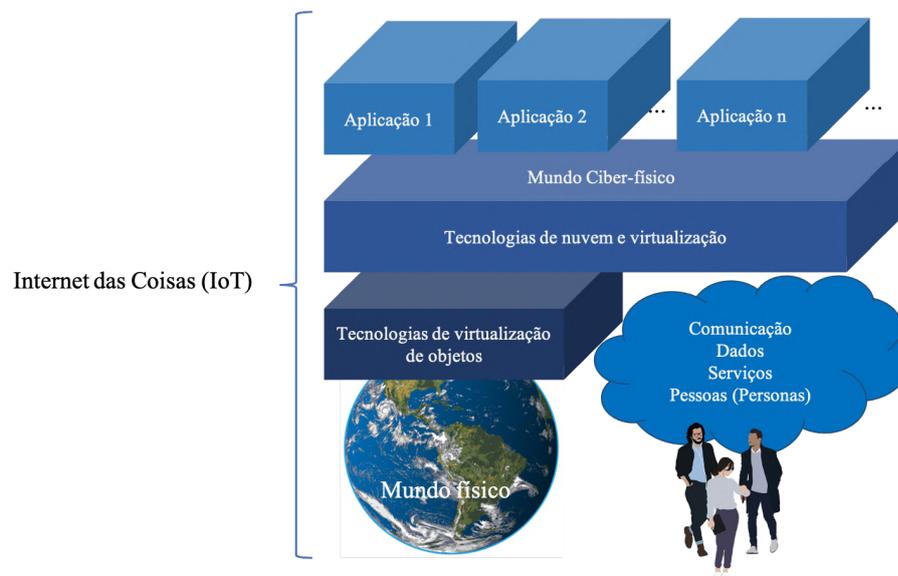
2.5 Internet das Coisas

A Identificação por Radio Frequência (RFID) (*Radio Frequency Identification*) é a importante tecnologia que deu origem à IoT. Esta tecnologia teve seus primeiros usos em meados da segunda guerra mundial, quando ondas de rádio eram utilizadas para identificar aviões inimigos. Desde então diversas aplicações foram elaboradas utilizando-se deste tipo de sinal. Até que em meados dos anos setenta cunhou-se o termo RFID para se referir a etiquetas que emitem ondas de rádio. Foi apenas no século XXI, com a ideia de se utilizar o RFID como base para dispositivos conectados à internet, que o termo IoT foi criado (INITIATIVE et al., 2015).

Desde então a definição de Internet das Coisas tem sofrido diversas transformações. Em 2010 a IoT estava sendo definida como um conceito de diversos dispositivos, como etiquetas RFID, sensores e atuadores conectados de forma pervasiva ao nosso redor (ATZORI; IERA; MORABITO, 2010).

Posteriormente Atzori, Iera e Morabito (2010) definiu que a Internet das Coisas poderia ser vista em diferentes formas ao longo de suas gerações. A atual geração da Internet das Coisas pode ser vista como uma plataforma que tenha elementos de conexão em rede, sensores, atuadores, autonomia, heterogeneidade e serviços, formando-se um *framework* conceitual que se utiliza destes elementos para fornecer uma base para aplicações que coloquem em um mesmo nível virtual, pessoas e objetos, como ilustra a Figura 3.

Figura 3 – *Framework* conceitual da IoT



Fonte: Adaptado de Atzori, Iera e Morabito, 2017

Ainda assim a definição de Internet das Coisas continuará a receber alterações dado às novas tecnologias que vão sendo desenvolvidas e associadas à ela ao longo do tempo (ATZORI; IERA; MORABITO, 2017).

2.6 Agricultura de Precisão

A Internet das Coisas é a tecnologia que contribuiu para o crescimento da agricultura de precisão (ANTONY et al., 2020). Este por sua vez pode ser definido como "uma estratégia de gestão que reúne, processa e analisa dados temporais, individuais e espaciais e os combina

com outras informações para apoiar as decisões de gestão de acordo com a variabilidade estimada para melhorar a eficiência no uso de recursos, produtividade, qualidade, rentabilidade e sustentabilidade da produção agropecuária"(ISPA, 2020).

O uso destas técnicas possibilita que o agricultor analise a variabilidade de informações espaço-temporais sítio-específicas de diversos fatores que afetam a saúde do plantio e a produtividade da cultura (CARDIM FERREIRA LIMA et al., 2020). Para alcançar esta precisão, diversas tecnologias de sensoriamento IoT podem ser utilizadas, dentre as quais encontram-se sensores de solo, estações meteorológica, *Drones* e satélites (MONTELEONE et al., 2020).

2.6.1 Sensoriamento remoto a partir de satélites

Sensoriamento remoto tem sido largamente utilizado nas últimas duas décadas (SHAFI et al., 2019a), o qual pode ser definido como a ciência ou a arte de obter informações de determinada área, ou fenômeno, por meio da coleta de dados que não entram em contato como alvo (MOLIN; AMARAL; COLAÇO, 2015).

Dentre os fenômenos observados encontra-se as condições físicas da terra que são observadas remotamente a partir de cálculos realizados com a radiação refletida a uma certa distância (SHAFI et al., 2019a). Estas observações podem ser realizadas através de diversas plataformas, como por exemplos *drones*, satélites e missões aéreas.

Dentre as plataformas utilizadas para realizar as observações, o satélite possui a maior estabilidade (SHAFI et al., 2019a). As vantagens que podem ser observadas ao utiliza-se desta plataforma consiste na alta resolução espacial, o que leva a uma grande quantidade de informações que podem ser processadas e o baixo ruído das imagens resultando em uma precisão maior das informações registradas (SHAFI et al., 2019a). Porém, esta plataforma apresenta desvantagens quanto a resolução temporal, que é dada pelo tempo de orbita e também pela alta sensibilidade a condições climáticas (SHAFI et al., 2019a).

2.7 Inteligência Artificial

A inteligência artificial é um campo de estudo recente que vem sendo estudado desde a década de 40, quando Warren McCulloch e Walter Pitts (1943) propuseram o modelo de neurônio artificial. A visão mais influente deste campo de estudo foi elaborada por Alan Turing (1950), cujo artigo "*Computing machinery and intelligence*"(TURING (1950) introduziu temas

relacionados ao Teste de Turing, aprendizagem de máquina, algoritmos genéticos e aprendizado por reforço (RUSSELL; NORVIG, 2002).

Apesar da inteligência artificial ser um campo recente, muitas outras áreas colaboram para os seus fundamentos, como por exemplos, a filosofia, matemática, economia, neurociência, engenharia da computação, teoria do controle, cibernética e linguística (RUSSELL; NORVIG, 2002).

Uma vez que tal área é fundamentada por tantas outras é normal que diferentes abordagens sejam elaboradas sobre o tema. Dentre as abordagens elaboradas, encontram-se sistemas que agem como humanos, ou seja, tenta-se criar sistemas que se passem por humanos em determinadas tarefas, como denota o teste de Turing. Nesta abordagem são utilizados algoritmos como processamento de linguagem natural, representação do conhecimento, raciocínio automatizado e aprendizagem de máquina (RUSSELL; NORVIG, 2002).

Outra abordagem diz respeito às máquinas que pensam como humanos, como tomada de decisões e soluções de problemas. A abordagem de sistemas que pensam racionalmente visa a elaboração do conhecimento a partir de notações lógicas, que representam todos os tipos de objetos e seus relacionamentos, contudo, esta abordagem possui uma deficiência na representação do conhecimento informal, além de exigir um grande poder computacional para resolver problemas que possuem muitas proposições lógicas (RUSSELL; NORVIG, 2002).

A abordagem de sistemas que agem racionalmente representa sistemas que, dentre diversas soluções possíveis, ao menos uma solução viável é encontrada. Estes sistemas também utilizam-se de algoritmos de raciocínio, processamento de linguagem natural, raciocínio automatizado e aprendizagem de máquina (RUSSELL; NORVIG, 2002).

2.8 Modelos de aprendizagem de máquina

Como uma subárea da inteligência artificial, os modelos de aprendizagem de máquina (*Machine Learning* (ML)) surgiram com o intuito de fazer com que máquinas aprendessem a partir de suas experiências (MITCHELL et al., 1997).

Como elabora Alpaydin (2020), modelos de aprendizagem de máquina utilizam-se de teorias estatísticas na construção de modelos matemáticos, uma vez que uma das principais tarefas deste tipo de modelo são as inferências realizadas a partir das amostras.

Na visão de Alpaydin (2020), outra área importante para construção de modelos de aprendizagem de máquina é a ciência da computação a qual contribui em dois aspectos princi-

país, sendo o primeiro, o desenvolvimento de modelos eficientes (em espaço e complexidade computacional), armazenamento e processamento de grandes quantidades de dados e, o segundo, diz respeito à eficiência (em espaço e complexidade computacional) de um algoritmo, uma vez que um modelo tenha aprendido a partir dos dados.

Tais modelos são utilizados em diversas aplicações, dentre as quais, pode-se definir os grupos de modelos de aprendizagem supervisionada, não supervisionada e de aprendizagem por reforço (ALPAYDIN, 2020). Nesta pesquisa são abordados apenas modelos de aprendizagem supervisionados, pois, uma vez que se tenha os exemplos da evapotranspiração de referência calculados pelo método FAOPM, é possível fornecer-los como objetivo dos modelos estudados.

2.8.1 Aprendizado Supervisionado

Pode-se definir o conceito de aprendizado supervisionado como o problema de procurar a hipótese que melhor se encaixa nos exemplos de treinamento dentro de um universo possível de hipóteses (MITCHELL et al., 1997).

Dois casos de aprendizagem supervisionada destacam-se. O primeiro é o caso de classificação, no qual, a partir de exemplos positivos e negativos, o modelo é capaz de encontrar a melhor descrição que define os exemplos positivos, para então classificá-los na categoria desejada. Diferente do primeiro caso, no qual a saída do modelo consiste em um valor binário (pertence, não pertence a uma classe), o segundo fornece uma saída numérica, obtida por uma função resultante do treinamento a partir dos exemplos fornecidos (ALPAYDIN, 2020).

Para a realização desta pesquisa serão utilizados os modelos Máquina de Vetores de Suporte e o Perceptron de Múltiplas Camadas, que consistem em modelos de aprendizagem supervisionada para o caso da regressão. Tais modelos foram selecionados pois dentre a literatura estudada para a modelagem da evapotranspiração de referência a partir de modelos de aprendizagem de máquina, estes se destacam.

2.8.2 O Modelo Máquina de Vetores de Suporte

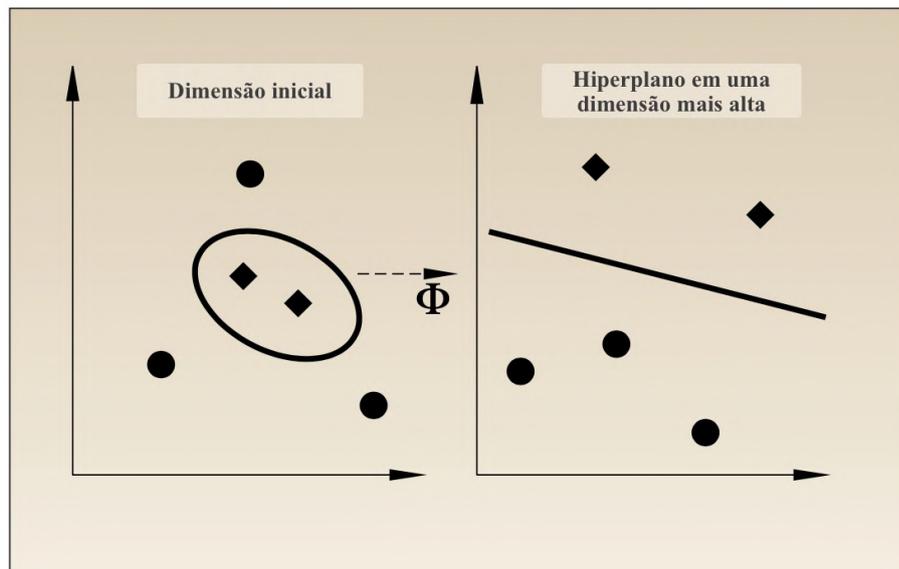
Desenvolvido por Vapnik (1995), este modelo de aprendizagem de máquina se destaca devido à sua sólida fundamentação teórica e ao alto desempenho em aplicações práticas (SANTOS et al., 2002). Conforme elaborado por Steinwart e Christmann (2008), o modelo Máquina de Vetores de Suporte (*Support Vector Machines* - SVM) se apoia em algoritmos de otimização

matemática, bem como, na teoria de aprendizado estatística, o que possibilita a identificação dos fatores que levam a uma aprendizagem de sucesso (HEARST et al., 1998).

Inicialmente o modelo SVM possibilitava apenas a aproximação de funções linearmente separáveis, o que mais tarde foi melhorado com a utilização de funções *Kernel* (SANTOS et al., 2002). Tal melhoria possibilitou que o modelo SVM lide com problemas complexos como reconhecimento de padrões, regressões e extração de características de populações de dados (HEARST et al., 1998).

A ideia geral do treinamento do modelo SVM em problemas não lineares é obter um hiperplano de separação máxima, obtido através de espaços geométricos de dimensões mais altas. Para se obter tais hiperplanos pode-se utilizar funções *kernel* como as funções de base radial e as funções polinomiais (HEARST et al., 1998). A figura 4 ilustra a ideia básica do treinamento do modelo SVM, no qual a população de dados passa de uma representação impossível de se separar linearmente, para uma representação linearmente separável, através do uso da função *kernel* denotada por Φ .

Figura 4 – Ideia Geral do Modelo SVM



Fonte: Adaptado de Hearst et al., 1998

2.8.3 Redes Neurais Artificiais

Redes neurais artificiais são modelos de aprendizagem de máquina que se baseia nos neurônios dos humanos e dos animais (MITCHELL et al., 1997).

Um neurônio é uma célula nervosa que possui um conjunto de dendritos que funcionam como receptores de estímulos. Estes estímulos são conhecidos como sinapses que são processados pelo axônio. Uma vez processada a informação o axônio envia o estímulo resultante para os dendritos de outros neurônios (ROSA, 2011). Tais conexões entre estas células formam uma grande rede, que ao ser modelada matematicamente, dá-se origem às redes neurais artificiais (ROSA, 2011).

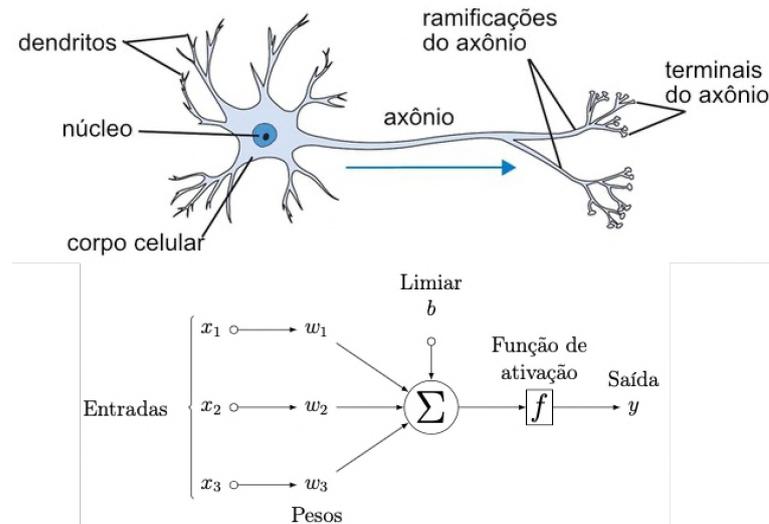
Modelos de aprendizagem baseados em redes neurais artificiais fornecem um abordagem robusta para aproximar valores-reais, valores-discretos, e valores-vetoriais, conforme elabora Mitchell et al. (1997). Outra motivação para o uso destes modelos advém da sua capacidade de processamento paralelo, baseado em suas representações distribuídas (MITCHELL et al., 1997), o que contribuiu para modelos mais eficientes (ALPAYDIN, 2020).

O trabalho pioneiro de McCulloch e Pitts (1943) deu origem ao modelo matemático de um neurônio artificial, no qual, mais tarde, foi adaptado por Rosenblatt (1958), dando-se origem ao modelo do *Perceptron*.

2.8.4 *Perceptron*

O modelo *Perceptron* consiste em um conjunto de entradas $(x_1, x_2, x_3, \dots, x_n \in R)$, ponderadas pelos pesos $w_1, w_2, w_3, \dots, w_n \in R$, respectivamente. Tais entradas ponderadas são combinadas através de uma função limiar, que gera o resultado 1, se a combinação é maior que um limiar e -1 caso contrário (MITCHELL et al., 1997). A Figura 5 ilustra o modelo do *Perceptron*.

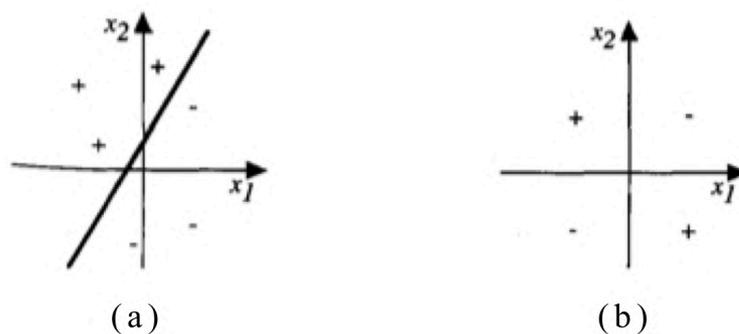
Figura 5 – Modelo de Neurônio Biológico e Modelo Perceptron



Fonte: Bezerra, 2016

Quando utilizado de forma unitária, o *Perceptron* possui a capacidade de representar diversas operações booleanas, como as operações AND, OR, NAND e NOR. Contudo, algumas operações booleanas não são possíveis de serem representadas pelo *Perceptron*, como por exemplo, a operação XOR, ilustrada na Figura 6 (b). Isto porque tal operação não é linearmente separável, ao contrário das outras operações que configuram resultados linearmente separáveis como ilustrado na Figura 6 (a) (MITCHELL et al., 1997).

Figura 6 – Problemas linearmente separáveis e problemas não linearmente separáveis



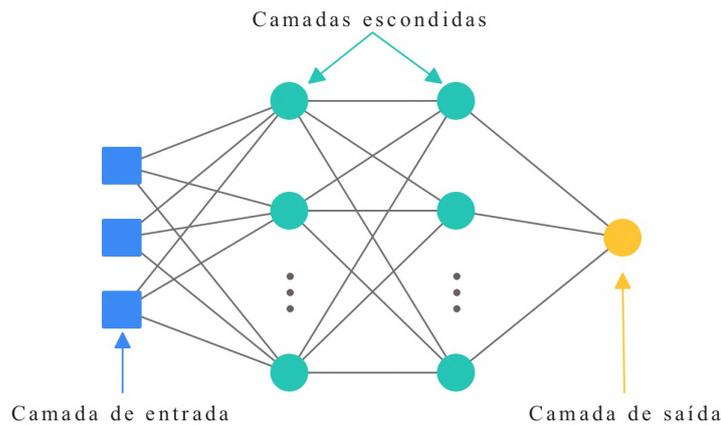
Fonte: Mitchell et al., 1997

Contudo, um modelo que se constitui de múltiplos *Perceptrons* interconectados entre si, possibilita resolver o problema da separação linear, uma vez que, criando-se uma nova camada, entre as entradas e a saída, é possível aproximar funções não lineares (HAYKIN, 2008).

2.8.5 *Perceptron de Múltiplas Camadas*

O modelo do *Perceptron* de Múltiplas Camadas (*Multlayer Perceptron* (MLP)) consiste de um conjunto de perceptrons conectados entre si que possui ao menos uma camada escondida, com (n) perceptrons (ALPAYDIN, 2020), conforme ilustrado na Figura 7.

Figura 7 – Modelo de *Multlayer Perceptron*



Fonte: Autor, 2020

Ao contrário do *Perceptron*, no qual a função de ativação era representada por uma função linear, o *Perceptron* de múltiplas camadas pode ser utilizado com diversas funções, como por exemplos, a função linear, logística e tangente, as quais possibilitam a regressão de uma função não linear (MITCHELL et al., 1997), (ALPAYDIN, 2020).

Para uma rede neural artificial aprender uma função não linear é necessário que o modelo seja treinado. No caso do *Perceptron* de múltiplas camadas, diversos algoritmos de treinamento podem ser utilizados, como o algoritmo de retro-propagação (ROSA, 2011) e o algoritmo *Adaptive moment estimation* (Adam) (KINGMA; BA, 2014), entre outros.

3 Revisão Bibliográfica Sistemática

Com o objetivo de verificar o estado da arte dos algoritmos de aprendizagem de máquina utilizados para modelar a evapotranspiração de referência a partir de parâmetros climáticos medidos por satélites, foi conduzida uma Revisão Bibliográfica Sistemática (RBS) (KITCHENHAM, 2004) com as seguintes questões a serem respondidas: Q1) Quais são os algoritmos de aprendizagem de máquina mais utilizados para determinar a evapotranspiração de referência? Q2) Qual é o método empírico mais utilizado como referência dos algoritmos? Q3) Quais são os principais parâmetros climáticos utilizados para o treinamento dos algoritmos? Q4) Quais são as bases de dados utilizadas para se obter parâmetros climáticos medidos por satélites?

Os critérios de inclusão (I) e exclusão (E) definidos foram: (I) Ter como objetivo o estudo da evapotranspiração a partir de algoritmos de *machine learning*; (I) estudos que se utilizavam de estações meteorológicas; (I) estudos que se utilizavam da base de dados de satélites que fornecessem parâmetros climáticos; (I) Estudos que determinam evapotranspiração obtida por parâmetros climáticos; (E) Estudos que não se utilizavam da evapotranspiração; (E) Estudos que se utilizavam de base de dados de satélites, mas que não fornecessem parâmetros climáticos; (E) Estudos que não utilizavam de estações meteorológicas ou satélites;

As buscas foram realizadas nas seguintes bases de dados eletrônicas: Springer, IEEE *Digital Library*, Elsevier *Digital Library*, MDPI e Scielo. O agregador de bases de pesquisa selecionado foi o Mendley. As ferramentas selecionadas para a análise dos resultados das buscas foram o VOSviewer (VOS) e o *software* Power BI. Também foi escolhido o software Notion para a realização de algumas etapas de filtragem e ordenação dos dados. A revisão foi realizada no período de janeiro/2020 a abril/2020, no qual as seguintes *strings* de busca foram utilizadas: a) ("*evapotranspiration*"AND "*machine learning*"); b) ("*evapotranspiration*"and "*limited climatic data*"); c) ("*evapotranspiration*"AND "*precision agriculture*"); d) ("*evapotranspiration*"AND "*forecasting*"); e) ("*evapotranspiration*"AND "*gridded weather data*"); f) ("*evapotranspiration*"AND "*reanalysis*").

Nas pesquisas foram encontrados 1048 artigos das máquinas de buscas (base), dos quais permaneceram 140, após a primeira fase de leitura do título. Após a segunda fase de leitura dos *abstracts*, permaneceram 47 artigos. Na terceira fase 26 artigos foram selecionados, como sendo os mais pertinentes ao escopo da pesquisa.

Após a extração dos dados da revisão sistemática, os algoritmos mais utilizados se concentraram no: *Multlayer Perceptron* (MLP) e *support vector machine* (SVM). Quanto aos mé-

todos empíricos utilizados como de referência, o que mais se destacou foi o método FAOPM. Os parâmetros climáticos mais significativos foram os parâmetros relacionados à temperatura, à radiação solar, à umidade relativa do ar e à velocidade do vento.

3.1 Trabalhos Relacionados

O uso de modelos de aprendizagem de máquina (ML) para determinar a evapotranspiração de referência (ET_o) vem sendo utilizados sob diferentes abordagens, por diversos pesquisadores. A principal motivação para o uso destas técnicas advém da possibilidade de obter-se este importante parâmetro, com uma precisão semelhante aos métodos empíricos utilizados como referência, porém com a menor quantidade de parâmetros climáticos possíveis.

No estudo realizado por Torres, Walker e McKee (2011) o objetivo era prever a evapotranspiração de referência, utilizando-se do método *Hargreaves* (HARGREAVES; ALLEN, 2003) como referência dos modelos. Para isso, os pesquisadores utilizaram os modelos Máquina Multivariada de Vetor de Relevância (*Multivariate Relevance Vector Machine* (MVRVM)) (THAYANANTHAN et al., 2008) e Perceptron de Múltiplas Camadas (MLP), sob duas abordagens. A primeira consistia em determinar o parâmetro ET_o utilizando-se dos modelos de ML, treinando-os com os parâmetros climáticos de temperatura mínima e máxima do ar. A segunda abordagem consistia em treinar os modelos de ML para prever os parâmetros climáticos e então utilizá-los no método *Hargreaves*.

Diferentes janelas de previsão foram testadas por Torres, Walker e McKee (2011), dentre as quais a janela de quatro dias mostrou os melhores resultados. Quanto as abordagens, a que se utiliza da previsão dos parâmetros de temperatura mínima e máxima gerou resultados mais precisos dentro da janela de quatro dias, representando uma melhor abordagem de treinamento. Entre os modelos utilizados o MVRVM provou ser o mais preciso dentre as diferentes janelas de previsão testadas e abordagens de treinamento utilizadas.

A pesquisa desenvolvida por Goyal et al. (2014) utilizou-se do método de evaporação em tanque Classe A (E_p) (ALLEN et al., 1998) como referência para o estudo dos modelos MLP, Mínimos Quadrados-Máquina de Vetores de Suporte (*Least Squares Support Vector Regression* (LSSVR)) (SUYKENS; VANDEWALLE, 1999), Lógica Nebulosa (ZADEH, 1996) e o Sistema de Inferência Neuro-Nebulosa Adaptativa (*Adaptive Neuro-Fuzzy Inference System* (ANFIS)) (JANG, 1993). O pesquisador também realizou um estudo comparativo entre os parâmetros de temperatura mínima e máxima do ar, umidade relativa do ar mínima e máxima, horas de

insolação e precipitação, utilizando-se da técnica teste *gamma* (KONCAR, 1997), para então obter a melhor combinação de parâmetros climáticos para o treinamento de cada modelo.

Dentre as diferentes combinações de parâmetros de entrada estudadas por Goyal et al. (2014), aquela que em todos os parâmetros climáticos são abordados mostrou melhores resultados. Contudo, com o uso do método do teste *gamma* foi possível perceber que os parâmetros relacionados a temperatura mostraram ser mais significativos para de terminar a Ep . Quanto aos modelos estudados, o LSSVR e a lógica nebulosa mostraram melhores resultados para de terminar a Ep , treinados com todos os parâmetros climáticos.

Apesar de diversos trabalhos serem desenvolvidos utilizando-se de métodos empíricos como o tanque Classe A e o método *Hargreaves*, como referência para os modelos de ML, muitos pesquisadores realizam seus estudos utilizando-se do método padrão recomendado pela FAO, sendo este o FAO Penman-Monteith (FAOPM).

A pesquisa de Pandey, Nyori e Pandey (2017) utiliza-se do método FAOPM como referência para o treinamento dos modelos Regressão Evolutiva (*evolutionary regression* (ER)) (GIUSTOLISI; SAVIC, 2006), Perceptron de Múltiplas Camadas (MLP), Regressão Multi Não Linear (*multi nonlinear regression* (MLNR)) (PANDEY; NYORI; PANDEY, 2017) e Máquinas de Vetores de Suporte (SVM). Neste caso o pesquisador utiliza-se de duas estações meteorológicas, que fornecem os parâmetros climáticos de temperatura mínima e máxima, radiação extraterrestre, umidade relativa mínima e máxima, horas de insolação e velocidade do vento. Semelhante à pesquisa de Goyal et al. (2014), a técnica teste *gamma* também é utilizada para investigar a melhor combinação de parâmetros climáticos para o treinamento dos modelos.

Os resultados de Pandey, Nyori e Pandey (2017) mostram que os modelos utilizados são técnicas promissoras para determinar a ET_o . Dentre os modelos estudados, o MLP e o ER obtiveram melhores resultados com a combinação de todos os parâmetros climáticos, porém o teste *gamma* revelou que os parâmetros de temperatura mínima e máxima, radiação extraterrestre, horas de insolação e velocidade do vento são mais significativos para determinar a evapotranspiração da região estudada.

A pesquisa de Nourani, Elkiran e Abdullahi (2019) utiliza estações meteorológicas de diversas regiões climáticas, das quais são obtidos os parâmetros climáticos de temperatura média, mínima e máxima, velocidade do vento mínima e máxima, radiação solar, precipitação, pressão atmosférica, umidade relativa e evaporação em tanque Classe A. Tais parâmetros são utilizados para treinar os modelos Rede Neural Feed Forward (*Feed Forward Neural Network* (FFNN)) (HORNIK; STINCHCOMBE; WHITE, 1989), Sistema de Inferência Neuro-Nebulosa

Adaptativa (ANFIS), Regressão de Vetores de Suporte (SVR) e Regressão Multilinear (*Multilinear Regression* (MLR)) (ÖZBAYOĞLU; ÖZBAYOĞLU, 2006) e também obter a evapotranspiração pelos métodos empíricos de *Hargreaves and Samani* (HS) (HARGREAVES; SAMANI, 1985), *Hargreaves and Samani* Modificado (MHS) (HU et al., 2011), *Makkink* (MK) (MAKKINK, 1957) e *Ritchie* (RT) (JONES, 1990).

Os autores Nourani, Elkiran e Abdullahi (2019) realizaram uma análise dos parâmetros climáticos obtidos que mais influenciam a evapotranspiração de referência de cada local estudado, para então obter um conjunto de parâmetros que serviram de entrada para os modelos de ML. Uma vez obtidos os conjuntos de parâmetros, duas abordagens são consideradas: a primeira consiste em determinar a ET_o por todos os métodos empíricos e modelos, para então comparar seus resultados entre si, e, a segunda consiste em combinar a ET_o resultante de cada método e cada modelo, utilizando-se de três modelos de combinação, sendo estes a média geral, média ponderada e um modelo baseado em rede neural artificial.

Os resultados obtidos por Nourani, Elkiran e Abdullahi (2019) mostraram que os parâmetros de temperatura média, mínima e máxima e radiação solar, são os parâmetros mais significativos em grande parte das estações. Outras estações são melhores representadas pela radiação solar, evaporação e precipitação. Por fim apenas uma estação é melhor representada pelos parâmetros de temperatura média, mínima e máxima. Quanto aos modelos estudados o ANFIS obteve melhores resultados na maior parte das estações, segundo o pesquisador este resultado pode ser dado devido à habilidade de aprendizado baseado em redes neurais e ao conceito de lógica Nebulosa, que são combinados em um único modelo. Os resultados também mostram que a evapotranspiração de referência, obtida pela abordagem de combinação dos modelos de ML mostram estimativas mais precisas, cujo método de combinação que forneceu as melhores estimativas, foi o baseado em redes neurais artificiais.

Diferentemente dos trabalhos anteriores, a pesquisa desenvolvida por Adeloje, Rustum e Kariyama (2011) utilizou-se do método de referência FAOPM, cujo método foi utilizado para treinar o modelo de busca não supervisionada Mapas Auto Organizáveis de Kohonen (*Kohonen self-organizing map* (KSOM)) (KOHONEN et al., 1996). Os parâmetros climáticos considerados pelo pesquisador foram temperatura mínima e máxima, radiação solar extraterrestre, radiação solar, horas de insolação máxima e umidade relativa mínima e máxima, com as quais diferentes combinações destes parâmetros foram testadas. Tais parâmetros foram obtidos a partir de duas estações meteorológicas em diferentes regiões climáticas, sendo uma em uma região

de clima temperado e a outra de clima árido. Isto por que os pesquisadores desejavam testar a adequação do modelo em regiões que variassem seus padrões climáticos.

Dentre os treinamentos realizados com as diferentes variações de parâmetros climáticos, o modelo treinado com todos os parâmetros obteve melhores resultados, tanto para a estação da região temperada, quanto para a região de clima árido. Contudo, os testes realizados considerando os parâmetros de temperatura mínima e máxima apenas, mostraram resultados satisfatórios, deixando claro o potencial de uso destes parâmetros, quando considerado cenários de clima temperado e árido.

A pesquisa de Muhammad Adnan et al. (2017), propõe uma abordagem diferente a qual ao invés de testar diferentes combinações entre os parâmetros climáticos, o modelo Análise das Componentes Principais (PCA) (AGARWAL et al., 2007) é utilizado para treinar um modelo MLP com suas componentes resultantes. Os parâmetros climáticos utilizados consistem na temperatura média, máxima e mínima do ar, velocidade do vento, precipitação, radiação solar e umidade relativa. Tais parâmetros foram recuperados através de uma estação meteorológica situada em uma região de clima semi-árido.

As análises realizadas na pesquisa de Muhammad Adnan et al. (2017) mostraram que a temperatura máxima do ar possui uma contribuição maior para a componente mais significativa, denotando que este parâmetro possui um grande impacto em determinar a evapotranspiração de referência da região estudada. As análises também mostram que utilizar o modelo PCA para então treinar o modelo MLP resultam em uma boa acurácia, mostrando que aquele é um método promissor.

Os pesquisadores Samsuri, Ahmad e Zakaria (2018) utilizam-se do método FAOPM como referência para o treinamento dos modelos Evolução Diferencial e Identificação do Sistema (*Differential Evolution and System Identification* (DESI)) (SAMSURI; AHMAD; ZAKARIA, 2018) e Algoritmo Genético Modificado (*Modified Genetic Algorithm* (MGA)) (JAMALUDDIN et al., 2007). Os parâmetros de temperatura mínima e máxima, radiação solar, umidade relativa e velocidade do vento, foram utilizados para treinar os modelos em uma região de clima tropical. Os resultados obtidos mostraram que tanto o DESI, quanto o MGA possuem resultados muito próximos do método de referência, porém, uma vez que tais modelos permitem explicar o relacionamento entre os parâmetros, foi possível mostrar que a radiação solar e umidade relativa são os parâmetros mais significativos para determinar a evapotranspiração de referência da região tropical estudada.

Outro exemplo que utiliza o método FAOPM como referência do treinamento dos modelos encontra-se na pesquisa realizada por Saggi e Jain (2019). Nesta pesquisa os modelos Aprendizagem Profunda (DL) (IVAKHNENKO; LAPA, 1965), Modelo Linear Generalizado (*Generalized Linear Model (GLM)*) (MCCULLAGH, 2018), Floresta Randômica (*Random forest (RF)*) (BREIMAN, 2001) e Máquina de Aumento de Gradiente (*Gradient-Boosting Machine (GBM)*) (FRIEDMAN, 2001), foram utilizados para prever a evapotranspiração de referência de duas estações meteorológicas localizadas em regiões de clima árido e semi-árido. As estações meteorológicas forneceram os parâmetros de temperatura mínima e máxima, umidade relativa, velocidade do vento, radiação solar e horas de insolação. Estes parâmetros, por sua vez, são pré-processados através modelo *MissForest* (SAHRI; YUSOF; WATADA, 2014), com o intuito de completar as lacunas de dados encontradas na base de dados. Os resultados obtidos mostraram que o modelo DL obteve melhor acurácia, quando comparado com outros modelos abordados.

Semelhantemente à pesquisa realizada pelos pesquisadores Adeloye, Rustum e Kariyama (2011) e Goyal et al. (2014), a pesquisa de Heddham et al. (2018) testa diferentes combinações dos parâmetros climáticos de temperatura mínima e máxima, velocidade do vento, umidade relativa e horas de insolação, obtidas a partir de duas estações meteorológicas. Tais combinações foram utilizadas para treinar os modelos Sistema de Inferência Neuro-Nebuloso de Evolução Dinâmica Offline (*off-line dynamic evolving neural-fuzzy inference system (DENFISOF)*), Sistema de Inferência Neuro-Nebuloso de Evolução Dinâmica Online (*on-line dynamic evolving neural-fuzzy inference system (DENFISON)*) (HEDDAM et al., 2018) e Rede Neuro-Nebulosa Evolutiva (EFuNN) (KASABOV, 2001), utilizando-se do método FAOPM como referência.

Os resultados obtidos por Heddham et al. (2018) mostram que os modelos treinados com todos os parâmetros possuem os melhores resultados, porém os modelos que se utilizam apenas dos parâmetros de temperatura mínima e máxima, velocidade do vento e horas de insolação, para a primeira estação, obtiveram resultados semelhantes aos modelos que utilizaram todos os parâmetros, assim como os modelos que utilizaram apenas os parâmetros de temperatura mínima e máxima, umidade relativa e horas de insolação, para a segunda estação. De todo modo o modelo DENFISOF mostrou os melhores resultados entre todos os agrupamentos testados.

Na pesquisa desenvolvida por Traore, Luo e Fipps (2016), foram utilizados quatro modelos de aprendizagem de máquina, com abordagens de redes neurais artificiais, sendo estes Rede *Feed Forward* Generalizado (*Generalized Feedforward (GFF)*) (TRAORE; LUO; FIPPS,

2016), Regressão Linear (*Linear Regression* (LR)) (TRAORE; LUO; FIPPS, 2016), Perceptron de Múltiplas Camadas (MLP) e a Rede Neural Probabilística (*Probabilistic Neural Network* (PNN)) (SPECHT, 1990). Tais modelos foram empregados para prever a evapotranspiração de referência obtida pelo método FAOPM sob diferentes janelas de previsão. Os parâmetros climáticos de temperatura mínima e máxima, radiação solar e radiação solar extraterrestre, obtidos por uma estação meteorológica de um local de clima úmido subtropical foram utilizados para treinar os modelos, dentre os quais diferentes combinações dos parâmetros foram testadas.

Os resultados obtidos por Traore, Luo e Fipps (2016) mostram que os modelos que se utilizaram do conjunto de todos os parâmetros climáticos considerados obtiveram melhores resultados, sendo o modelo MLP aquele que mostrou melhor acurácia. A janela de previsão de um dia foi a que obteve melhores resultados, porém a janela de 15 dias, utilizada com modelo MLP com todos os parâmetros climáticos fornece resultados promissores.

Já a pesquisa de Kisi e Demir (2016) utilizou apenas o modelo de aprendizagem de máquina MLP, porém foram testados seis algoritmos de treinamento diferentes, sendo estes o Quasi-Newton, Gradiente Conjugado, Levenberg-Marquardt, Secante de Um Passo, Retro-Propagação Resiliente e Gradiente Conjugado em Escala para obter a evapotranspiração pelo método FAOPM. Os parâmetros climáticos de radiação solar, temperatura média, umidade relativa e velocidade do vento foram obtidos a partir de duas estações meteorológicas em uma região de clima mediterrâneo. Dentre todos os modelos de treinamento utilizados, o Levenberg-Marquardt gerou os melhores resultados.

A pesquisa realizada por Rana Muhammad Adnan et al. (2020) analisa a adequação dos modelos de aprendizagem de máquina Mínimos Quadrados-Máquina de Vetores de Suporte com Algoritmo de Busca Gravitacional (*least-square support vector regression with a gravitational search algorithm* (LSSVRGSA)) (ADNAN, R. M. et al., 2020), Inferência Neuro-Nebuloso de Evolução Dinâmica (*dynamic evolving neural-fuzzy inference system* (DENFIS)) (KASABOV; SONG, 2002) e Árvore Modelo M5 (*M5 model tree* (M5RT)) (QUINLAN et al., 1992) para determinar a evapotranspiração de referência obtida pelo método FAOPM. Para isso o pesquisador utiliza-se de 3 estações meteorológicas de uma região de clima úmido-temperado, das quais, são recuperados os parâmetros de radiação solar extraterrestre, temperatura do ar e evapotranspiração de referência, cujos parâmetros de temperatura e a radiação são utilizados para gerar diferentes combinações de treinamento dos modelos.

Os resultados obtidos por Rana Muhammad Adnan et al. (2020) mostram que o LS-SVRGSA mostrou melhores resultados para os treinamentos realizados com o parâmetro de

temperatura ou radiação para todas as três estações meteorológicas. Porém o modelo DENFIS obteve melhores resultados quando treinado com os parâmetros climáticos de temperatura e radiação na terceira estação meteorológica. Contudo, os resultados revelam que os modelos treinados com apenas a radiação solar extraterrestre fornecem bons resultados.

Na literatura também é possível encontrar pesquisas que se utilizam de combinações de métodos empíricos para o cálculo da evapotranspiração de referência. Um exemplo é a pesquisa realizada por Kişi e Öztürk (2007) que se utilizou dos modelos Sistema de Inferência Neuro-Nebulosa Adaptativo (ANFIS), Perceptron de Múltiplas Camadas (MLP) e Regressão Multilinear (MLR) para modelar a evapotranspiração de referência, obtida pelo método da FAOPM e então comparou seus resultados com os métodos *Penman* (PENMAN, 1948), *Hargreaves e Ritchie*. Neste estudo foram utilizados duas estações meteorológicas na mesma região climática, das quais são recuperados os parâmetros de temperatura do ar, radiação solar, velocidade do vento e umidade do ar, as quais são combinadas em dois conjuntos para treinar os modelos. Os resultados obtidos por Kişi e Öztürk (2007) mostram que o modelo ANFIS, em todas as combinações de parâmetros climáticos testadas, obteve melhores resultados do que os métodos empíricos e os modelos MLP, treinado com os mesmos parâmetros.

A pesquisa realizada por Antonopoulos e Antonopoulos (2017), a qual utilizou o método FAOPM como referência do modelo Perceptron de Múltiplas Camadas (MLP), também utilizou-se dos métodos *Priestley-Taylor* (PRIESTLEY; TAYLOR, 1972), *Makkink*, *Hargreaves* e *Transferência de Massa* (SINGH; XU, 1997) para que o modelo fosse avaliado. Nesta pesquisa foi utilizada apenas uma estação meteorológica, onde foram recuperados os dados de temperatura média, radiação solar, umidade relativa e velocidade do vento. Os pesquisadores utilizaram-se do coeficiente de correlação para elencar os parâmetros climáticos mais significativos da região estudada, tais parâmetros foram recuperados de um período de 5 anos, no qual a base de dados treinamento foi configurada de forma cruzada, gerando uma divisão de 4 anos para o treinamento e 1 ano de teste. Tal abordagem é realizada para todos os anos.

Os resultados obtidos por Antonopoulos e Antonopoulos (2017) mostram que o modelo MLP obteve melhores resultados quando utilizado com os 4 parâmetros climáticos, porém os modelos que possuíam, ao menos, os parâmetros de temperatura e radiação obtiveram resultados melhores do que aqueles que não utilizaram-se destes parâmetros. Por fim a precisão do MLP depende mais dos parâmetros escolhidos para o seu treinamento do que a quantidade de anos de dados.

Utilizando-se de apenas um modelo de aprendizagem de máquina, Gocic et al. (2016) realiza uma abordagem diferente. Nesta pesquisa o modelo Aprendizagem de Máquina Extrema (*extreme learning machine* (ELM)) (HUANG; ZHU; SIEW, 2004) é utilizado para obter a evapotranspiração dos métodos *Hargreaves* Ajustado, *PriestleyTaylor* e *Turc* (TURC, 1961), para então ser comparado com o método de referência FAOPM. Para isso o pesquisador utiliza-se de duas estações meteorológicas, em uma mesma região climática, das quais são coletados os parâmetros de temperatura mínima e máxima, pressão de vapor atual, velocidade do vento e horas de insolação.

Os resultados obtidos por Gocic et al. (2016) mostram que o algoritmo ELM que utilizou o método *Hargreaves* ajustado como referência obteve melhores resultados, seguido do modelo que utilizou o método *PriestleyTaylor*.

A pesquisa desenvolvida por Shiri (2017) utilizou o modelo Programação de Expressão Gênica (*gene expression programming* (GEP)) (FERREIRA, 2001) para obter a evapotranspiração de referência de diferentes métodos empíricos, sendo estes o *Hargreaves* (HG), *Priestley-Taylor* (PT) (WRIGHT; JENSEN, 1972), *Turc* (Tr), *Kimberly-Penman* (KP) e FAOPM. Para isso o pesquisador testa diferentes combinações do modelo utilizando-se de parâmetros climáticos obtidos por dez estações meteorológicas de diferentes locais de clima desértico quente, tais parâmetros são: temperatura média, mínima e máxima, umidade relativa, velocidade do vento e radiação solar. Ademais, a fim de avaliar o impacto na precisão dos métodos e do modelo, também foram utilizados os parâmetros de radiação solar, velocidade do vento e umidade relativa, estimados por equações empíricas. Os resultados mostram que o modelo GEP superou todos os modelos empíricos de referência, contudo, sua precisão não pode ser estendida para outros locais.

Já a pesquisa realizada por Tangune e Escobedo (2018) utiliza-se dos modelos MLP e SVM para determinar a evapotranspiração de referência, obtida a partir do método FAOPM. Estes modelos foram avaliados utilizando-se de dois conjuntos de métodos empíricos, sendo o primeiro baseado em temperatura, representado pelos métodos *Benevides-Lopez* (BenL) (GARCIA BENEVIDES; LOPEZ, 1970), *Hamon* (Ham) (HAMON, 1960), *Blaney-Criddle Original* (BCO) (BLANEY et al., 1952) e *Hargreaves-Samani* (HS), e o segundo conjunto representado pelos métodos *Abtew* (AB) (ABTEW, 1996), *Jensen-Haise* (JensH) (JENSEN; HAISE, 1963), *Makkink* (MK) e *Irmak* (IRMAK et al., 2003). Os modelos foram treinados utilizando-se de diferentes combinações dos parâmetros climáticos, sendo estes a temperatura média, mínima e máxima, umidade relativa, velocidade do vento e radiação solar e extraterrestre. Tais parâ-

metros foram obtidos através de 22 estações meteorológicas, localizadas em diferentes regiões climáticas do estado de São Paulo.

Dentre as avaliações realizadas por Tangune e Escobedo (2018) com os métodos empíricos dos conjuntos baseados em radiação e o conjunto baseado em temperatura, os métodos AB e HS mostraram os melhores resultados. Quando comparados com os modelos abordados, o método HS obteve os piores resultados e o AB ficou à frente dos algoritmos que se utilizaram do parâmetro de temperatura apenas. Quanto aos modelos, aqueles que se utilizaram dos parâmetros de temperatura média, mínima e máxima e radiação solar e extraterrestre mostraram os melhores resultados para todas as estações meteorológicas.

Na pesquisa realizada por Ferreira et al. (2019b), a utilização da abordagem de diferentes métodos empíricos também foi empregada, cujos métodos *Hargreaves-Samani* (HS), *Oudin* (OUDIN et al., 2005), *Makkink* (MK), *Jensen and Haise* (JH), *Romanenko* (RK) (MEHDIZADEH; BEHMANESH; KHALILI, 2017), *Valiantzas* (VZ) (VALIANTZAS, 2013) e FAOPM foram utilizados para obter a evapotranspiração de referência, assim como o modelo Spline de Regressão Adaptativa Multivariada (*multivariate adaptive regression splines* (MARS)) (FRIEDMAN, 1991). Os parâmetros climáticos utilizadas foram a temperatura mínima e máxima do ar, radiação solar e extraterrestre, umidade relativa e velocidade do vento, os quais foram recuperados por diversas estações meteorológicas, localizadas em diferentes regiões climáticas do Brasil. Com o intuito de avaliar o modelo MARS comparando-o com os métodos empíricos, o pesquisador estuda diferentes combinações dos parâmetros climáticos para o treinamento do modelo, utilizando-se da ET_o determinada pelo método FAOPM como referência de treinamento.

Os resultados obtidos por Ferreira et al. (2019b) mostram que o FAOPM foi, de fato, o método empírico que mostra os melhores resultados, entre os métodos utilizados. O modelo MARS obteve os melhores resultados em todas as combinações de parâmetros climáticos testadas, quando comparado com os métodos empíricos utilizados. Contudo, os modelos testados com os parâmetros de temperatura mínima e máxima do ar, radiação solar e extraterrestre, forneceu os resultados mais precisos.

Posteriormente Ferreira et al. (2019a) continuou a pesquisa anterior, porém desta vez o pesquisador utilizou os modelos MLP e SVM para obter a evapotranspiração de referência, determinada pelo método FAOPM, que também foram comparados com métodos empíricos, sendo estes os métodos *Hargreaves-Samani* (HS), *Oudin*, *Hamon*, *Valiantzas* (VZ), *Romanenko* (RK) e *Schendel* (DJAMAN et al., 2015). Duas abordagens de treinamento dos modelos com

parâmetros climáticos foram utilizadas, a primeira abordagem consiste em utilizar o modelo *K-Means* (LIKAS; VLASSIS; VERBEEK, 2003) para determinar o melhor agrupamento de estações meteorológicas, e então treinar os modelos MLP e SVM com os parâmetros climáticos dos agrupamentos obtidos; a segunda abordagem consiste em utilizar os parâmetros climáticos de dias anteriores como variáveis de entrada dos modelos já treinados. Os parâmetros climáticos de temperatura mínima e máxima, umidade relativa, horas de insolação, velocidade do vento, utilizados pelos métodos empíricos e os modelos de aprendizagem de máquina foram recuperados a partir de 203 estações meteorológicas do Brasil.

Os resultados obtidos por Ferreira et al. (2019a) mostram que entre todos os métodos empíricos, originais e calibrados, o método FAOPM utilizado com o parâmetro de temperatura e o FAOPM utilizado com o parâmetro de temperatura e umidade, obtiveram melhores resultados comparado com todos os outros métodos empíricos. Dentre os modelos, o MLP obteve melhores resultados, quando comparado com o SVM, tanto para o treinamento realizado para cada estação, quanto para as abordagens realizadas. Por fim o pesquisador concluiu que utilizar a umidade relativa junto com os parâmetros de temperatura, para o treinamento dos modelos que utilizam-se de parâmetros climáticos limitados, pode melhorar os resultados quando utilizados em regiões climáticas semelhantes as estudadas.

A pesquisa elaborada por Feng et al. (2017) utiliza os modelos Aprendizado de Máquina Extremo ELM e Rede Neural de Regressão Generalizada (*generalized regression neural network* (GRNN)) (SPECHT et al., 1991) para determinar a evapotranspiração de referência obtida pelo método FAOPM, assim como o método empírico HS. Junto aos modelos o pesquisador também testa diferentes abordagens para a utilização das estações meteorológicas, cuja primeira abordagem consiste em treinar os modelos com os parâmetros de temperatura, recuperado por cada estação, para então comparar seus resultados com o método HS; e a segunda abordagem utiliza-se dos parâmetros de temperatura agrupados de todas as estações para então treinar os modelos, e por fim testar o resultado individualmente para cada estação utilizada. Para isso o pesquisador utiliza 6 estações, dentro de uma mesma região climática, no qual são recuperados os parâmetros de temperatura mínima e máxima, umidade relativa, velocidade do vento e horas de insolação. Neste sentido o pesquisador atesta que há lacunas nas séries temporais, e, para resolver este problema, é utilizado um método de interpolação linear.

Os resultados obtidos por Feng et al. (2017) mostram que os modelos GRNN e ELM elaboram resultados melhores que o método HS para as duas abordagens utilizadas. O modelo ELM mostrou melhores resultados para a abordagem de treinamento e teste em estações locais,

já o algoritmo GRNN mostrou resultados melhores para a abordagem de treinamento com estações agrupadas e teste em estações locais.

A pesquisa de Kisi (2015) utilizou-se dos modelos Mínimos Quadrados-Máquina de Vetores de Suporte (*Least square support vector machine* (LSSVM)), Spline de Regressão Adaptativa Multivariada (MARS) e Árvore Modelo M5 para determinar a evaporação em tanque Classe A. Nesta pesquisa duas estações meteorológicas foram utilizadas, onde são recuperados os parâmetros de temperatura, radiação solar, velocidade do vento e umidade relativa. Dois cenários foram testados, o primeiro realizou o treinamento e teste dos modelos, em cada estação, já o segundo cenário realizou o treinamento e o teste de maneira cruzada, cujos dados de uma estação são utilizados para o treinamento e os dados da outra para o teste.

Os resultados de Kisi (2015) mostram que o modelo LSSVM mostrou resultados mais precisos no cenário local, contudo para o cenário cruzado o modelo MARS mostrou melhores resultados.

Já a pesquisa de realizada por Reis et al. (2019) os modelos MLP, Regressão Multilinear (MLR) e Máquina de Aprendizado Extremo (ELM) são utilizados para determinar a evapotranspiração de referência pelo método FAOPM, junto aos métodos empíricos HS e HSA. Para realizar esta pesquisa, 5 estações meteorológicas de uma região semi-árida do estado de Minas Gerais foram utilizadas. Foram coletados os parâmetros de temperatura mínima e máxima, umidade relativa média, velocidade do vento e horas de insolação, porém algumas lacunas de dados foram encontradas, e para resolver este problema, o pesquisador utiliza o método de interpolação linear. O pesquisador utiliza duas abordagens para o treinamento dos modelos com os dados das estações, a primeira abordagem consiste em treinar e testar os modelos com os dados de cada estação para então testá-los em suas respectivas localidades; a segunda abordagem consiste em agrupar os dados de todas as estações, para então treinar os modelos e depois testá-los nas localidades de cada estação.

Os resultados obtidos por Reis et al. (2019) indicam que os modelos estudados mostraram melhor acurácia do que os métodos empíricos, nas duas abordagens consideradas. Segundo Reis et al. (2019) todos os três modelos obtiveram acurácia similares, contudo o MLR é aconselhado ante os outros, pois através dele é possível obter equações algébricas. Dentre as abordagens consideradas, os melhores resultados foram obtidos no cenário local.

Semelhante a outros trabalhos, Sanikhani et al. (2019) utiliza o método FAOPM como referência dos modelos (MLP), (GRNN), Rede Neural de Base Radial (*radial basis neural networks* (RBNN)) (BROOMHEAD; LOWE, 1988), Sistema de Inferência Neuro-Nebuloso

Integrado com Partição em Grade e Agrupamento Subtrativo (*Adaptive neuro-fuzzy inference systems with grid partitioning* (ANFISGP) and *Adaptive neuro-fuzzy inference systems with subtractive clustering* (ANFISSC)) e Programação de Expressão Gênica (GEP), cujos resultados destes modelos são comparados com os métodos empíricos HS e HSA. Nesta pesquisa também são utilizadas 2 estações meteorológicas, das quais são recuperados os parâmetros temperatura mínima e máxima, radiação solar, umidade relativa e velocidade do vento. Tais parâmetros são combinados de diferentes maneiras para serem realizados os treinamentos dos modelos, não obstante, o pesquisador também utiliza-se de dois cenários com as estações meteorológicas, o primeiro cenário foi treinados e testados os modelos com os dados de cada estação, no segundo cenário, os dados de uma estação são utilizados para o treinamento e teste dos modelos, para então ser utilizado no outro local.

Os resultados obtidos por Sanikhani et al. (2019) mostram que no cenário local, os modelos GEP e GRNN forneceram os melhores resultados para a primeira estação, já na segunda estação, os resultados mais precisos foram obtidos pelos modelos GRNN e ANFISSC. Quanto ao cenário cruzado, todos os modelos forneceram resultados melhores do que os métodos HS e HSA.

A pesquisa de Ferreira e Cunha (2020) utiliza os modelos MLP, Aumento de Gradiente Extremo (*Extreme gradient boosting* (XGBoost)) (CHEN et al., 2016), Floresta Randômica (RF) e Rede Neural Convolucional (*Convolutional neural network* (CNN)) (GOODFELLOW et al., 2016), para determinar a evapotranspiração de referência, pelo método FAO-PM. Nesta pesquisa foram utilizadas cinquenta e três estações meteorológicas do estado de Minas Gerais, das quais foram recuperados os parâmetros de temperatura mínima e máxima, umidade mínima e máxima, radiação solar e velocidade do vento. O pesquisador define diferentes conjuntos com tais parâmetros, onde, não somente foram coletados os parâmetros em escala diária, como também foram coletados os parâmetros em escala horária, com os quais o pesquisador testa quatro abordagens: na primeira, são utilizados os parâmetros relativos à temperatura, à umidade e à radiação solar, em escala diária, para o treinamento e testes dos modelos, na segunda são utilizados os mesmos parâmetros, porém em escala horária, na terceira a abordagem utiliza-se dos parâmetros de temperatura e umidade em escala horária e o parâmetro de radiação em escala diária, por fim na quarta, semelhante à segunda abordagem, utiliza-se dos parâmetros em escala diária, porém desta vez é obtida a evapotranspiração de referência em escala horária, para então ser determinada a ET_o diária.

Não obstante o pesquisador também utiliza-se de dois cenários com as estações meteorológicas, o primeiro, denominado de cenário local, onde Ferreira e Cunha (2020) define dez estações mais representativas de suas respectivas regiões climáticas, para então testar as abordagens em cada uma delas individualmente; já no segundo cenário, o pesquisador utiliza-se de quarenta e três estações para o treinamento dos modelos, sob as abordagens elaboradas, para então testá-los nas 10 estações representativas.

Os resultados de Ferreira e Cunha (2020) mostram que em ambos os cenários, local e regional, os melhores resultados foram obtidos pelos modelos testados com os parâmetros de temperatura mínima e máxima, umidade mínima e máxima e radiação solar. Dentre os modelos o MLP mostrou resultados melhores, quando treinado com os parâmetros em escala diária, contudo o modelo CNN superou-o quando treinado com os parâmetros em escala horária.

Apesar dos bons resultados, os pesquisadores acima citados restringem-se a locais que possuem estações meteorológicas para a realização de suas pesquisas, contudo outro cenário pode ser avaliado, no qual são utilizadas estações meteorológicas disponíveis publicamente e parâmetros climáticos obtidos por sensoriamento remoto a partir de satélites. Um exemplo deste tipo de cenário pode ser encontrado na pesquisa de Srivastava et al. (2020) que avaliou uma série de bases de dados de parâmetros climáticos obtidos a partir de satélites, para então determinar a evapotranspiração de referência obtida pelo método FAO-PM e o método *Hamon*, comparando-os com a ET_o , obtida por parâmetros climáticos de estações meteorológicas. As bases de dados testados por Srivastava et al. (2020) consistem no repositório de dados NASA/POWER e a base *National Center for Environmental Prediction* (NCEP), com os quais foram recuperados parâmetros climáticos de uma região do clima subtropical úmido.

Os resultados encontrados por Srivastava et al. (2020) elaboram que a evapotranspiração de referência obtida pelas bases de dados utilizados mostram conformidade com a ET_o determinada com os dados das estações meteorológicas, contudo a base de dados NASA/POWER mostrou uma fornecer dados de melhor precisão.

A pesquisa de Júnior, Valeriano e Souza Rolim (2019) visou desenvolver uma aplicação para dispositivos móveis, na qual a evapotranspiração de referência, obtida pelo método FAO-PM, pudesse ser obtida em qualquer geolocalização. Para isso o pesquisador avaliou a adequação da base de dados NASA/POWER. Esta por sua vez mostrou ser uma fonte de dados meteorológicos adequados para diversas regiões do Brasil.

Já a pesquisa realizada por Monteiro, Sentelhas e Pedra (2018) avaliou a adequação da base de dados NASA/POWER em todo o território brasileiro. Os resultados obtidos corroboram

os resultados de outros pesquisadores, porém os parâmetros climáticos de velocidade do vento e umidade possuem defasagens, podendo fornecer resultados suspeitos, sendo aconselhável o uso de métodos que se utilizam de parâmetros climáticos relacionados à temperatura e à radiação solar.

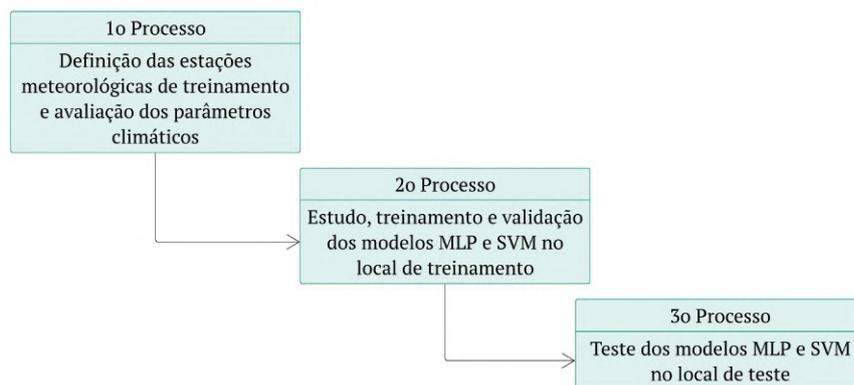
Dentre a literatura pesquisada diversos pesquisadores exploram o uso de modelos de aprendizagem de máquina para determinar a evapotranspiração de referência utilizando-se de métodos empíricos como referências de seus modelos. Contudo o método FAO-PM se destaca, uma vez que este é considerado a referência para se determinar a ET_o . Há também uma tentativa sistemática de se obter a ET_o com os parâmetros mais significativos, que normalmente se concentram em parâmetros climáticos relacionados à temperatura e à radiação solar na maior parte das regiões climáticas estudadas. Há também um movimento na literatura que se utiliza não somente de estações meteorológicas, mas também de parâmetros climáticos obtidos por satélites, para se determinar a ET_o o que reforça o objetivo desta pesquisa. Dentre os modelos de aprendizagem de máquina o modelo MLP e o modelo SVM se destacam em diversas pesquisas. A Tabela 2 ilustra os modelos de aprendizagem de máquina mais utilizados na literatura pesquisada.

4 Metodologia

Conforme elaborado no seção 3.1 os modelos SVM e MLP destacam-se em diversas pesquisas da modelagem da evapotranspiração de referência (ET_o), onde os parâmetros climáticos obtidos por satélites mostram-se promissores como alternativa ao uso de estações meteorológicas. No entanto para se avaliar o uso destes modelos, treinados com estes parâmetros climáticos, é necessário a avaliação junto a parâmetros medidos por estações meteorológicas, para que os resultados dos modelos possam ser comparados com alguma referência.

Sendo assim o diagrama da Figura 8 ilustra os processos realizados para a execução da metodologia proposta, utilizando-se a literatura avaliada, bem como algumas ferramentas utilizadas.

Figura 8 – Diagrama de processos executados para a execução da metodologia empregada



Fonte: Autor, 2020

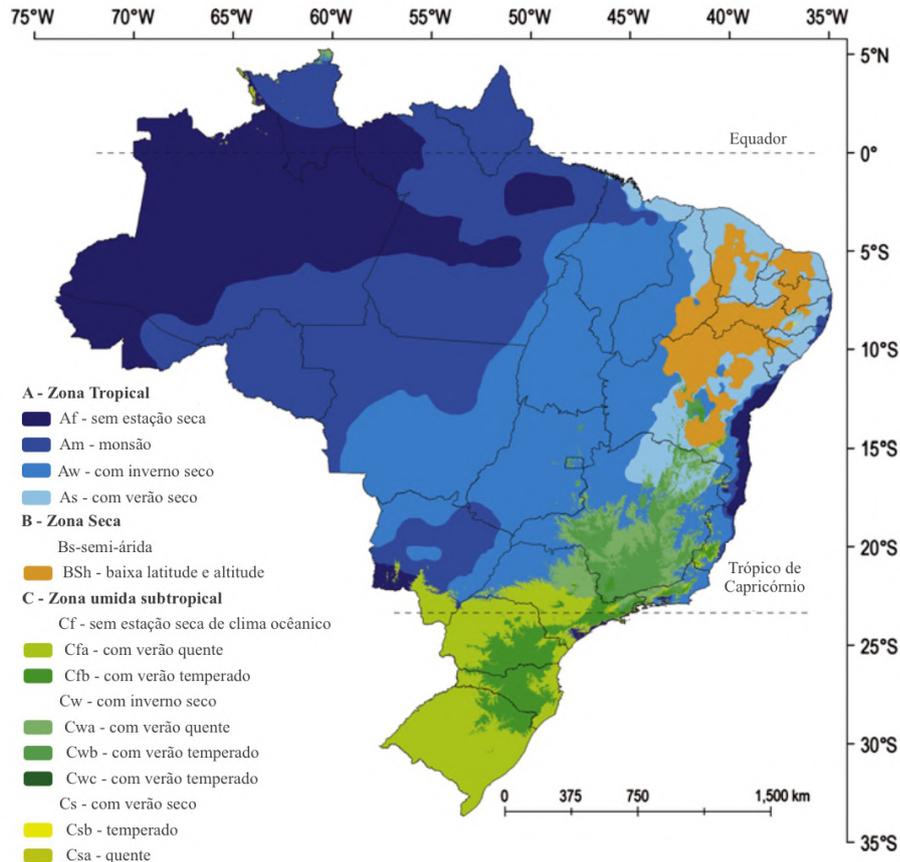
Expandindo o primeiro processo da Figura 8, tem-se que inicialmente são definidos dois locais, um de teste (com estação meteorológica do agricultor com dados para o teste dos modelos) e um de treinamento (com estações meteorológicas com dados históricos para o treinamento dos modelos). Abordagens semelhantes são encontradas nas pesquisas de Kisi (2015), Sanikhani et al. (2019) e Ferreira e Cunha (2020).

A pesquisa de Ferreira e Cunha (2020) também elabora que os locais de treinamento devem ser selecionados na mesma região climática dos locais de teste, para que as variações dos dados climáticos entre os dois locais sejam minimizadas. Esta classificação climática é obtida a partir da classificação *Köppen* (KÖPPEN; GEIGER, 1930).

O estudo realizado por Alvares et al. (2013) elabora um mapa climático da classificação *Köppen* com uma resolução espacial de um hectare, o qual pode ser utilizado em diferentes

sistemas de Informações Geográficas, também conhecidos como *Geographic Information Systems (GIS)*. O mesmo mapa é utilizado no estudo de Ferreira et al. (2019b). A Figura 9 ilustra o mapa de classificação climática, elaborado na pesquisa de Alvares et al. (2013).

Figura 9 – Mapa de Classificações Climáticas



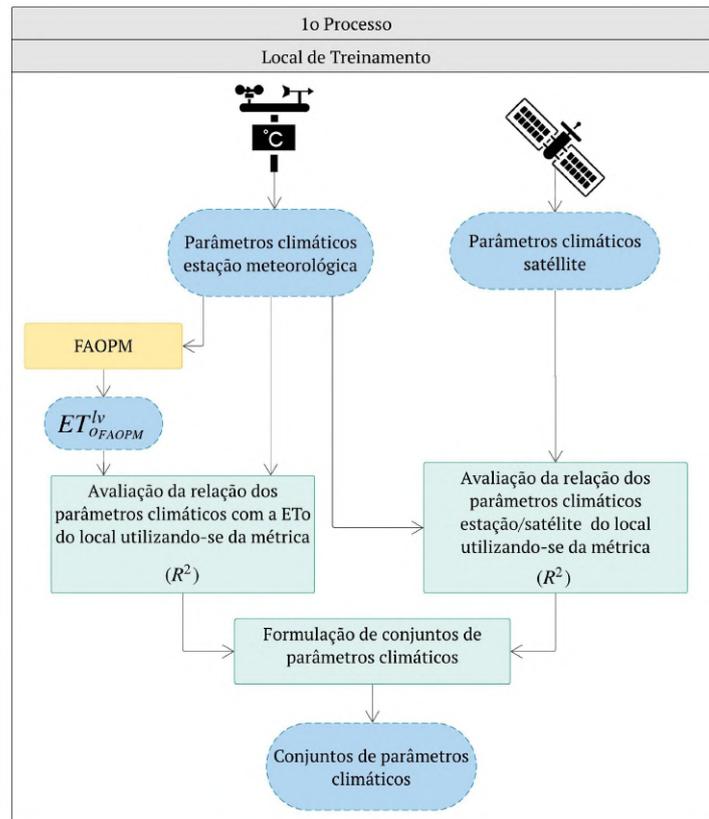
Fonte: Adaptado de Alvares et al., 2013

Para que as estações meteorológicas de treinamento sejam avaliadas, o sistema ARCGIS Pro é utilizado. Neste sistema o mapa produzido por Alvares et al. (2013) é adicionado, onde também são colocadas as coordenadas geográficas das estações meteorológicas das regiões estudadas, bem como, as coordenadas dos locais de teste. Neste sentido, ao selecionar o local no mapa, o sistema retornará a classificação climática correspondente, sendo assim as estações meteorológicas que se mantiverem mais próximas dos locais de teste são selecionadas como local de treinamento.

Após a definição da estação meteorológica de treinamento são recuperados os parâmetros climáticos da respectiva estação, para que o parâmetro de evapotranspiração de referência (ET_o) seja calculado para o local. Para isso o método estabelecido por Allen et al. (1998) elaborado na Equação 1 foi implementado no ambiente de desenvolvimento *Excel*.

Após o cálculo do parâmetro de evapotranspiração do local de treinamento, são recuperados os parâmetros climáticos medidos por satélites, onde o coeficiente de determinação (R^2) é utilizado para avaliar a relação entre os parâmetros climáticos obtidos do local de teste, pelas estações meteorológicas, e pela ET_o calculada.

Figura 10 – Diagrama da avaliação dos parâmetros climáticos



Fonte: Autor, 2020

Esta análise é necessária para que seja avaliada a contribuição dos parâmetros climáticos em relação à ET_o . A mesma análise pode ser encontrada na pesquisa realizada por Antonopoulos e Antonopoulos (2017).

O coeficiente R^2 também é utilizado para avaliar a relação entre os parâmetros climáticos medidos pelas estações meteorológicas e os satélites, assim como a pesquisa elaborada por Monteiro, Sentelhas e Pedra (2018).

A saída do primeiro processo consiste nos conjuntos de parâmetros climáticos que serão utilizados para treinar os modelos, bem como as séries temporais dos parâmetros de evapotranspiração de referência, que serão utilizadas como objetivo e validação dos modelos de

aprendizagem de máquina. A Figura 10 ilustra as etapas e os resultados gerados no primeiro processo.

Uma vez que os parâmetros foram avaliados e os conjuntos de parâmetros foram definidos, inicia-se o segundo processo, no qual são avaliadas diferentes arquiteturas dos modelos SVM e MLP treinados com os conjuntos de parâmetros obtidos no primeiro processo, No qual a função de ativação dos neurônios do modelo MLP consiste na tangente hiperbólica elaborada na equação 2, a seguir.

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (2)$$

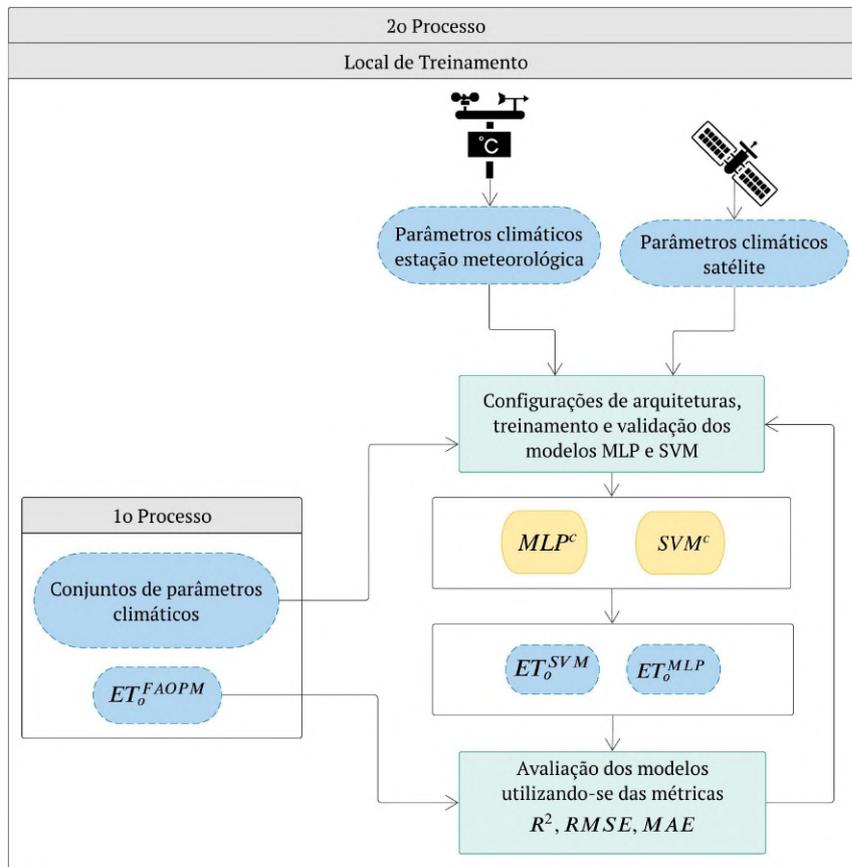
Diferentes pesquisadores elaboram diferentes configurações dos modelos SVM e MLP em seus estudos, sendo assim as arquiteturas dos modelos de aprendizagem de máquina utilizadas nesta pesquisa foram selecionados de acordo método de tentativa e erro no qual são recuperados cada padrão de arquitetura encontrada na literatura pesquisada, para então ser realizada uma composição e a partir destas composições são realizadas avaliações para se definir a arquitetura que possui a melhor performance para o problema em questão. O mesmo foi realizado por autores como Kişi e Öztürk (2007), Reis et al. (2019) e Ferreira e Cunha (2020).

Uma vez que a arquitetura dos modelos foi selecionada, inicia-se o treinamento dos modelos no local de treinamento, no qual os conjuntos de parâmetros climáticos estabelecidos no processo anterior são utilizados como variáveis de entrada dos modelos e como de saída é utilizada a variável ET_o^{FAOPM} do local.

Após o treinamento dos modelos, denotados por SVM^c e MLP^c , avalia-se suas respectivas acurácias, cujos resultados denominados de ET_{oSVM} e ET_{oMLP} , são comparados com a variável de referência ET_o^{FAOPM} utilizando-se das métricas de avaliação $RMSE$, MAE e R^2 .

A saída do segundo processo consiste nos modelos MLP e SVM treinados. O Diagrama da Figura 11 ilustra a sequência desta segunda etapa.

Figura 11 – Diagrama do estudo, treinamento e validação dos modelos MLP e SVM no local de treinamento



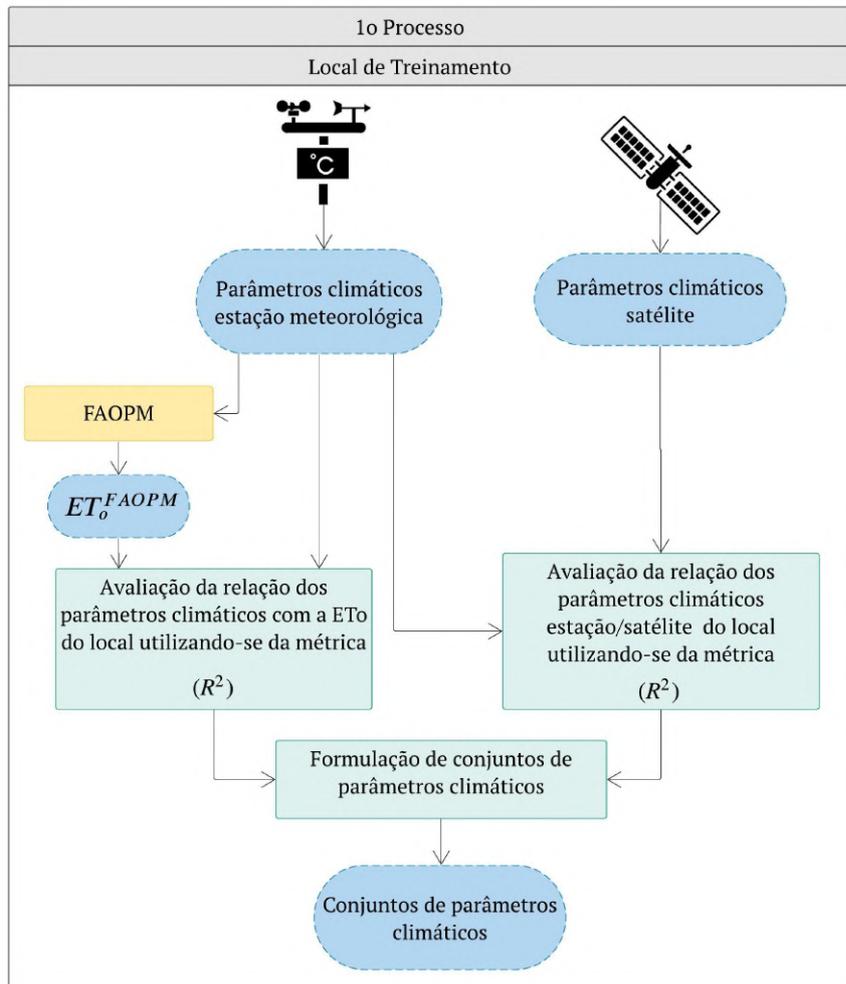
Fonte: Autor, 2020

Por fim o terceiro processo consiste na avaliação da acurácia dos modelos, testando-os no local de teste do experimento, utilizando-se novamente dos grupos de parâmetros obtidos no primeiro processo.

Para isso, são recuperados os parâmetros climáticos da estação meteorológica do local, na qual a ET_o^{FAOPM} é calculada novamente, no ambiente de desenvolvimento *Excel*.

Após o cálculo da ET_o do local, são recuperados os parâmetros climáticos obtidos por satélites, na mesma localização da estação meteorológica, que serão utilizados como entrada dos modelos.

Figura 12 – Diagrama de teste dos modelos MLP e SVM no local de teste



(Fonte: Autor, 2020)

Duas avaliações são feitas utilizando-se do coeficiente de determinação R^2 . A primeira consiste na avaliação dos parâmetros climáticos obtidos pela base de dados de satélites em relação aos parâmetros climáticos da estação local, assim como na primeira etapa, e a segunda avaliação é realizada entre a ET_o do local de teste em relação a ET_o do local de treinamento.

Após estas avaliações, os modelos de aprendizagem de máquina são aplicados com os dados do local, onde os resultados obtidos, denominados de ET_o^{SVM} e ET_o^{MLP} , são comparados com a ET_o^{FAOPM} utilizando-se das métricas de avaliação $RMSE$, MAE e R^2 , conforme ilustra o Diagrama da Figura 12.

4.1 Métricas de Avaliação

As métricas utilizadas para a avaliação dos resultados dos modelos de aprendizagem de máquina foram selecionadas de acordo com a literatura, na qual três se destacam.

As métricas consistem no coeficiente de determinação (R^2), elaborado na equação 3, a raiz quadrada do erro médio ($RMSE$), elaborado na equação 4 e, o erro médio absoluto (MAE), elaborado na equação 5. O uso destas métricas pode ser observado nas pesquisas de Antonopoulos e Antonopoulos (2017), Traore, Luo e Fipps (2016), Pandey, Nyori e Pandey (2017), Zhang, Gong e Wang (2018), Ferreira et al. (2019a) e Ferreira e Cunha (2020).

$$R^2 = \frac{\sum_{i=1}^n (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\left(\sum_{i=1}^n (P_i - \bar{P})^2\right) \left(\sum_{i=1}^n (O_i - \bar{O})^2\right)}} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - O_i)^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |P_i - O_i| \quad (5)$$

Onde n representa o número de observações, P_i o valor estimado do parâmetro, \bar{P} a sua média, O_i é o valor observado do parâmetro e \bar{O} a sua média.

4.2 Parâmetros Climáticas e Bases de Dados

Para se obter os dados climáticos de estações meteorológicas é utilizada a rede de estações do Instituto Nacional de Meteorologia (INMET) e pode ser acessada *online* através do *link* (<https://bdmep.inmet.gov.br>). Assim como nos trabalhos de Monteiro, Sentelhas e Pedra (2018) e Ferreira et al. (2019b), são extraídos os parâmetros de temperatura mínima (T_{min}), Temperatura máxima (T_{max}), Horas de Insolação (R_h), Umidade relativa do ar (RH) e velocidade do vento a uma altura de 10 metros (u_{10}), em uma escala diária.

Estes parâmetros são necessários para que a ET_o do local de treinamento seja obtida utilizando-se a Equação 1 e as equações 6, 7, 8 e 9 para se obter os parâmetros de temperatura média (T_m), radiação solar (R_a), radiação solar extraterrestre (R_s) e velocidade do vento a 2 metros (u_2), conforme Allen et al. (1998), durante o período de 01/01/2009 até 30/06/2020.

$$T_m = \frac{T_{max} + T_{min}}{2} \quad (6)$$

Onde, T_m representa a temperatura média em (°C) e T_{max} e T_{min} a temperatura máxima e mínima em (°C), respectivamente.

$$R_a = \frac{24(60)}{\pi} 0.082 d_r [\omega_s \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \sin(\omega_s)] \quad (7)$$

Onde, R_a representa a radiação solar extraterrestre em ($MJ m^{-2} h^{-1}$). O parâmetro d_r representa a distância inversa relativa entre a terra e o sol, δ a declinação solar, φ a latitude do local de estudo, em radianos, ω_s o ângulo da hora do por do sol.

$$R_s = \left(0.25 + 0.50 \frac{n}{N}\right) R_a \quad (8)$$

Onde, R_s representa a radiação solar em ($MJ m^{-2} d^{-1}$), n representa as horas de insolação, N representa o máximo de horas de insolação em um dia e R_a a radiação solar extraterrestre.

$$u_2 = u_z \frac{4.87}{\ln(67.8 z - 5.42)} \quad (9)$$

Onde, u_2 representa a velocidade do vento a 2 m em (m/s), u_z a velocidade do vento na altura z , assim como z representa a altura de medição da velocidade do vento.

Para se obter os dados climáticos de satélites, o repositório de dados da *National Aeronautics and Space Administration* (NASA/POWER) (STACKHOUSE et al., 2015) é utilizado, que pode ser acessado *online* pelo link (<https://power.larc.nasa.gov>).

A base de dados NASA/POWER consiste em parâmetros climáticos estimados por diferentes projetos, como *NASAs Fast Longwave and Shortwave Radiative Fluxes (FLASHFlux)*, *NASA's Modern Era Retro-analysis for Research and Applications assimilation model (MERRA-2)* e o projeto GEOS, em uma escala global de 0.5° de latitude e 0.5° longitude (STACKHOUSE et al., 2015).

As variáveis extraídas da base de dados NASA/POWER consiste na Temperatura mínima (T_{min}), Temperatura máxima (T_{max}), Temperatura média (T_m), Radiação solar extraterrestre (R_a), Radiação solar (R_s), Umidade relativa do ar (RH) e velocidade do vento a uma altura de 2 metros (u_2), em uma escala diária, durante o período de 01/01/2009 a 30/06/2020.

A terceira fonte de dados fornece os dados para a validação dos resultados obtidos nos locais treinamento nos locais de teste. Estes dados são fornecidos por agricultores que possuem estações meteorológicas em suas localidades que seguem as normas de instalação assim como as especificações do INMET.

As fontes de dados do INMET e NASA/POWER são divididas em dois conjuntos. O primeiro conjunto utilizado para treinamento e validação dos (algoritmo)s no local de treinamento e o segundo para os testes no local de teste de cada experimento.

O conjunto de treinamento e validação é dividido em dois subconjuntos, no qual o primeiro constitui-se de 75%, voltados para o treinamento, e o segundo é formado com 25% para a validação. A divisão de 75% – 25% segue o mesmo padrão encontrado nos trabalhos de Pandey, Nyori e Pandey (2017), Ferreira et al. (2019a). O conjunto de teste em local de treinamento constitui-se de período de um ano (01/06/2019 até 30/06/2020).

O teste dos modelos nos locais de teste de cada experimento utiliza-se de um conjunto de dados de um ano, que varia de acordo com a disponibilidade de dados de cada agricultor.

4.3 Experimentos

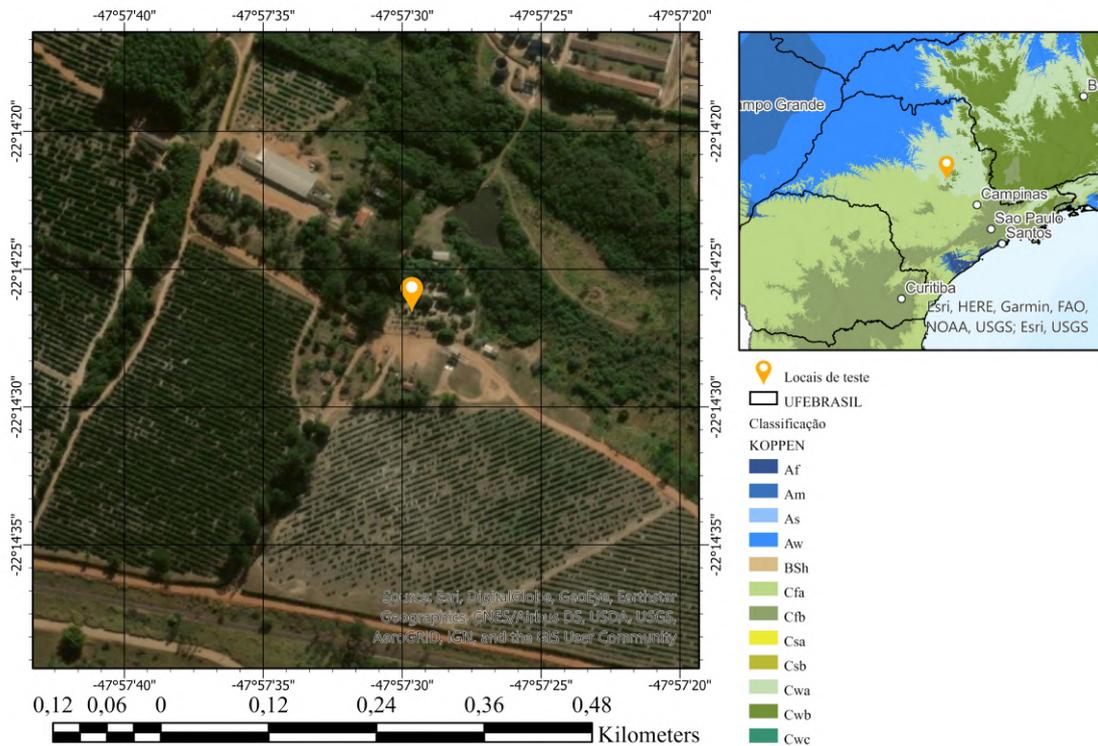
Para a realização dos experimentos foi desenvolvido um *script* em linguagem de programação *Python* no ambiente de desenvolvimento *online Google Colab*, utilizando-se dos pacotes *Tensor Flow* e *Scikit Learning*, onde todos os dados são importados e processados.

O primeiro experimento localiza-se na cidade de Brotas, no Estado de São Paulo, nas coordenadas de Latitude 22.24S e Longitude 47.95W a 756 metros acima do nível do mar. Neste experimento são capturados os parâmetros de temperatura média (T_m), radiação solar (R_s), umidade relativa média (RH_m) e velocidade a 2 metros (u_2). A classificação climática deste local é caracterizada como Temperado de inverno seco e verão quente (Cwa), conforme elaborado por (ALVARES et al., 2013).

Este experimento localiza-se no cinturão citrícola brasileiro que contempla cidades dos estados de Minas Gerais e São Paulo, tais cidades são divididas em cinco regiões denominadas de Noroeste, Norte, Centro, Sul e Sudoeste. Esta região é possuí uma área plantada com citros de 465.635 hectares, distribuídos em 9.845 propriedades (CITRICULTURA, 2016).

Em São Paulo a área dedicada ao plantio de laranja é representada por 62% da área de produção de laranjas no Brasil, dentre os quais 98% das laranjas produzidas é destinada a exportação (BARROS; BARROS; CYPRIANO, 2016). A Figura 13 ilustra o local do experimento.

Figura 13 – Localização do experimento de Brotas



Fonte: Autor, 2020

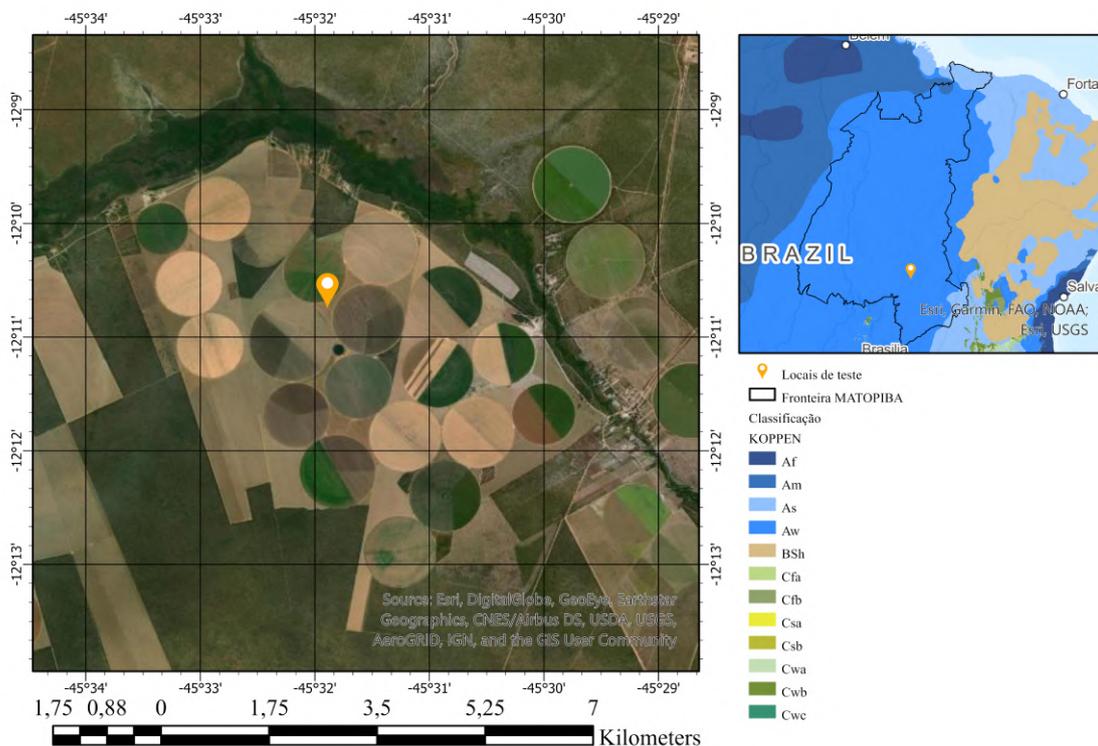
Neste experimento foram concedidos um conjunto de dados meteorológicos, que consiste no período de 20/09/2018 a 13/12/2019 para a que fosse realizado os testes dos modelos no local de teste.

Como parte do projeto *Smart Water Management Platform (SWAMP)* (KAMIENSKI et al., 2019), o experimento MATOPIBA localiza-se e na cidade de Barreiras, no Estado da Bahia, nas coordenadas de Latitude $12.17S$ e Longitude $45.53W$ a 733 metros acima do nível do mar. Neste experimento são capturados os parâmetros de temperatura mínima e máxima (T_{min} , T_{max}), radiação solar (R_s), umidade mínima e máxima (RH_{min} , RH_{max}) e velocidade do vento a 2 metros (u_2). Este experimento possui a classificação climática Tropical de savana (Aw), conforme elaborado por Alvares et al. (2013).

A região de MATOPIBA é a grande fronteira agrícola do Brasil, onde são contempladas regiões dos estados de Maranhão, Tocantins, Piauí e Bahia. O clima favorável desta região permite o cultivo das principais culturas de grãos e fibras, tais quais a soja, arroz e algodão (EMBRAPA, 2017).

Está é uma área de suma complexidade, que representa um desafio para a (agricultura) sustentável onde seus 337 municípios chegam a cerca de 73 milhões de hectares, além dos 324 mil estabelecimentos agrícolas e 35 terras indígenas. Ainda assim grande parte do seu impulso produtivo advém do acesso às tecnologias que hoje são utilizadas, como cultivares adaptados a condições edafoclimáticas e boas práticas para a utilização de fertilizantes, corretivos e defensivos, sistemas conservacionistas de manejo e a integração lavoura-pecuária-floresta (EMBRAPA, 2017). A figura 14 ilustra o local do experimento.

Figura 14 – Localização do experimento de MATOPIBA



Fonte: Autor, 2020

Neste segundo experimento foi utilizado um conjunto de dados meteorológicos relativos ao período de 01/01/2019 a 20/04/2020 para fosse realizado os testes dos modelos no local de teste.

5 Resultados

Neste capítulo são apresentados os resultados obtidos em cada etapa da pesquisa, no qual os primeiros resultados são relativos a definição dos locais de teste dos respectivos experimentos, posteriormente são apresentados os resultados das arquiteturas dos modelos de aprendizagem de máquina, passado pelos resultados dos treinamento dos modelos e, por fim os resultados da aplicação dos modelos nos locais de teste.

5.1 Definição dos Locais de Treinamento

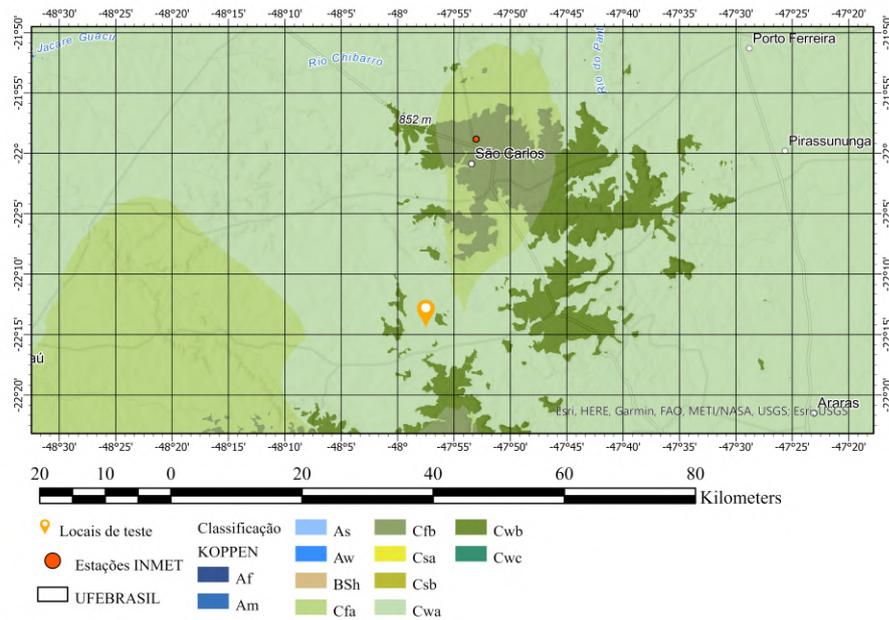
Utilizando-se do sistema ARCGIS Pro foi possível avaliar as estações meteorológicas do INMET mais próximas dos locais de teste, que serviram de local de treinamento e validação dos modelos de aprendizagem de máquina. A Figura 15 ilustra o local selecionado para o treinamento e validação dos modelos no experimento de Brotas, assim como a Figura 16 ilustra o local selecionado para o experimento de MATOPIBA.

5.1.1 Experimento de Brotas

A estação selecionada para o experimento de Brotas encontra-se à distância de 32.41 km do local de teste, nas coordenadas de Latitude 21.98*S*, Longitude 47.88*W* a uma altitude de 859.29m acima do nível do mar.

A classificação climática deste local é caracterizado por um clima temperado sem estação seca de verão morno (Cfb). Esta estação foi selecionada por representar a menor distância do local de teste e se localizar em uma região climática com características semelhantes.

Figura 15 – Avaliação do local de teste e treinamento para o experimento de Brotas

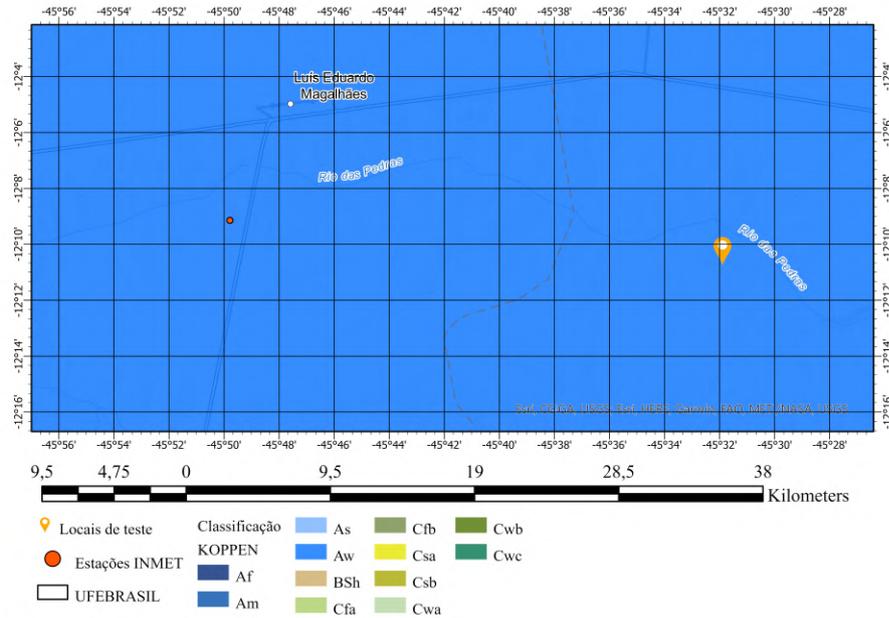


Fonte: Autor, 2020

5.1.2 Experimento de MATOPIBA

A estação selecionada para o experimento de MATOPIBA encontra-se à distância de 33.89 km do local de teste, nas coordenadas de Latitude 12.15S e Longitude 45.83W à altitude de 760.68 m acima do nível do mar. A classificação climática deste local é caracterizado por um clima tropical de savana (Aw), assim como a região climática do local de teste.

Figura 16 – Avaliação do local de teste e treinamento para o experimento de MATOPIBA



Fonte: Autor, 2020

5.2 Conjunto de Parâmetros

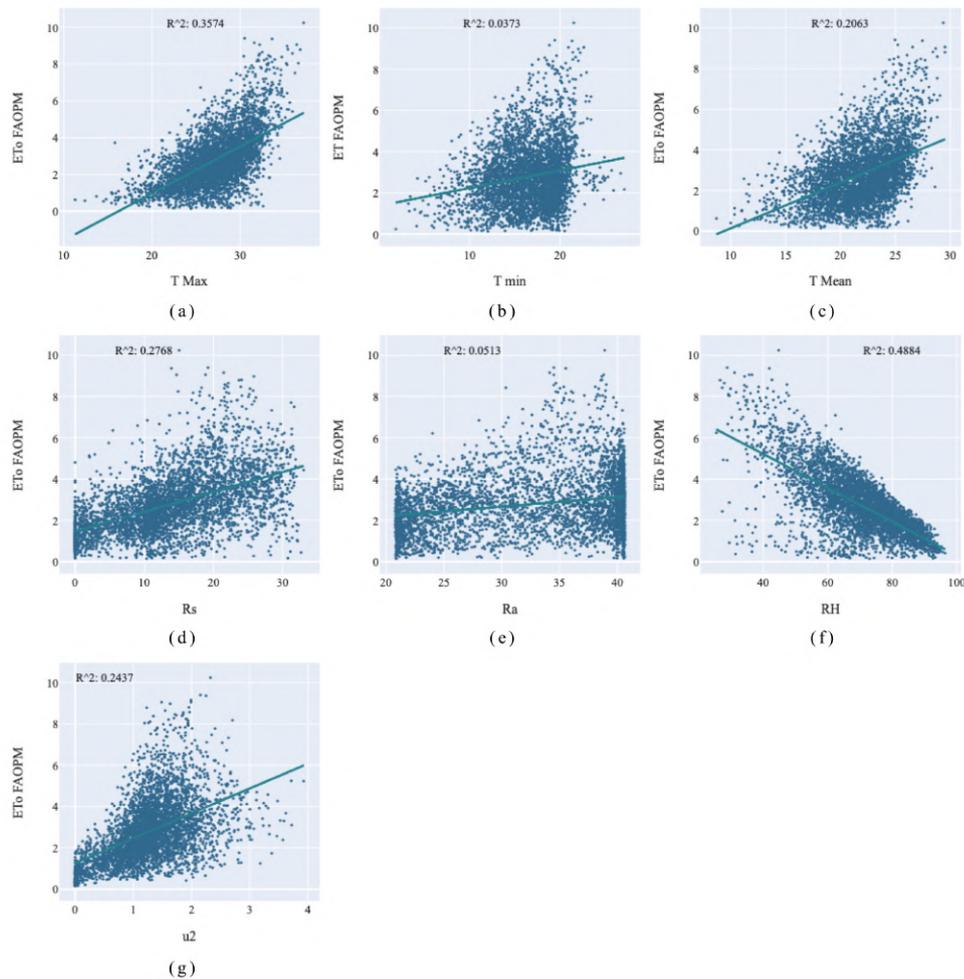
Conforme elaborado no capítulo 4, e exibido na Figura 10, é necessário avaliar os parâmetros climáticos obtidos das bases de dados utilizando-se do coeficiente de determinação R^2 , para então testar diferentes conjuntos de parâmetros para calcular a ET_o do local, que posteriormente são utilizados para treinar os modelos de aprendizagem de máquina no local de treinamento e também para testar os modelos no local de teste.

5.2.1 Experimento de Brotas

Os resultados da execução do primeiro processo, elaborado no capítulo 4 e ilustrado na Figura 10, realtivo ao experimento de Brotas é elaborado a seguir.

A Figura 17 exhibe o coeficiente de determinação entre os parâmetro climáticos e a evapotranspiração. A relação entre o parâmetro T_{max} e ET_o - Figura 17 (a) - possui um R^2 de 0.3574, enquanto os parâmetros T_{min} - Figura 17 (b) - e T_m - Figura 17 (c) - possuem um coeficiente R^2 de 0.0373 e 0.2063, respectivamente em relação a ET_o .

Figura 17 – Gráficos de avaliação dos parâmetros climáticos INMET e a ET_o calculada no experimento Brotas



Fonte: Autor, 2020

O coeficiente R^2 dos parâmetros R_a - Figura 17 (d) - e R_s - Figura 17 (e) - relacionados com a ET_o , obtiveram um resultado de 0.2768 e 0.0513, enquanto o coeficiente de determinação do parâmetro de umidade relativa RH_m - Figura 17 (f) - ficou em 0.4884 e o R^2 do parâmetro u_2 - Figura 17 (g) - ficou em 0.2437.

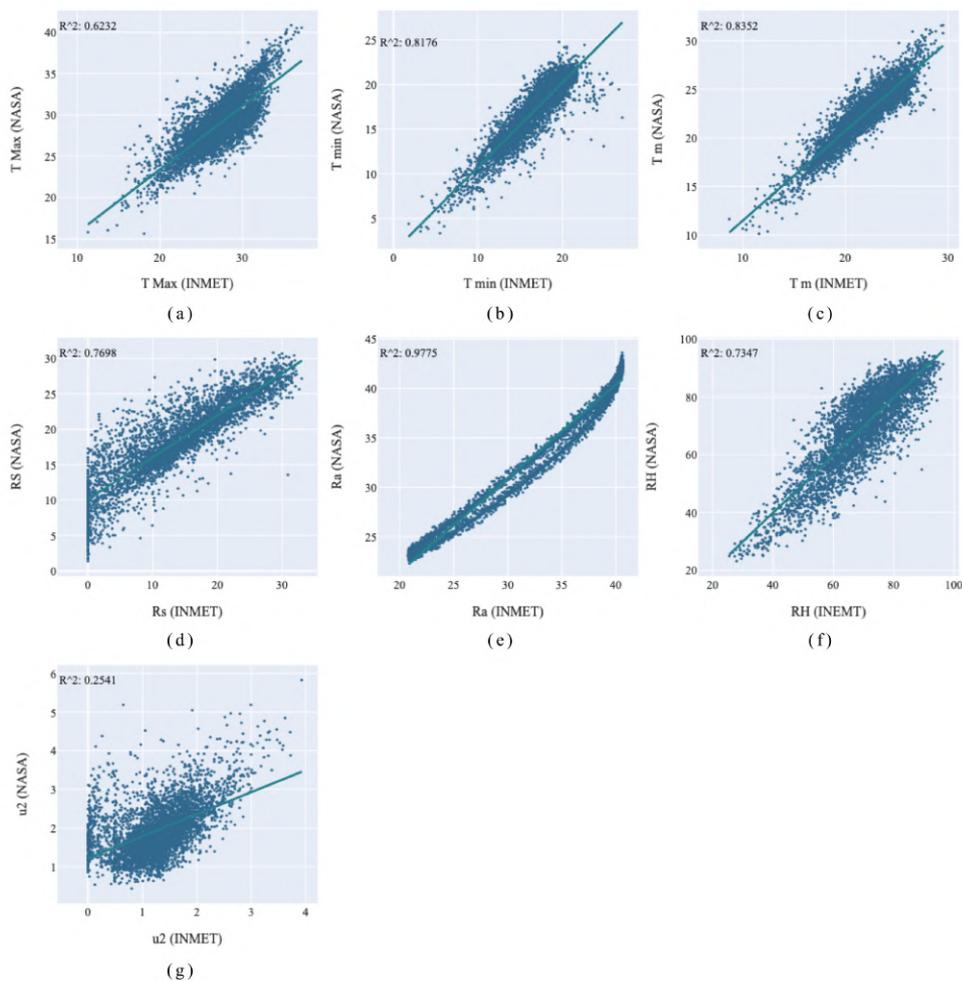
Tais avaliações demonstram que a evapotranspiração de referência (ET_o) deste local possui uma relação maior com os parâmetros de umidade relativa (RH), temperatura máxima (T_{max}), radiação solar (R_s), velocidade do vento (u_2) e temperatura média (T_m). Os parâmetros de temperatura mínima (T_{min}) e radiação solar extra terrestre (R_a) possuem relação menor com a evapotranspiração de referência. A Tabela 3 exibe os resultados obtidos.

Tabela 3 – Resultados do coeficiente de determinação entre a ET_o calculada e os parâmetros climáticos INMET para o experimento de Brotas

ET_o calculada	R^2
Parâmetros climáticos INMET	
T_{max}, ET_o	0.3574
T_{min}, ET_o	0.0373
T_m, ET_o	0.2063
R_s, ET_o	0.2768
R_a, ET_o	0.0513
RH, ET_o	0.4884
u_2, ET_o	0.2437

Fonte: Autor, 2020

Figura 18 – Gráficos de avaliação dos parâmetros climáticos NASA/POWER e INMET no experimento Brotas



Fonte: Autor, 2020

A Figura 18 exibe os resultados do coeficiente de determinação utilizado entre os parâmetros climáticos obtidos pela base de dados do INMET e a base de dados NASA/POWER. O coeficiente de determinação do parâmetro de temperatura máxima - Figura 18 (a) - resultou em um R^2 de 0.6232, enquanto R^2 dos parâmetros de temperatura mínima - Figura 18 (b) - ficou em 0.8176, sendo que o R^2 da temperatura média - Figura 18 (c) - resultou em 0.8352.

O coeficiente de determinação dos parâmetros de radiação resultou em um R^2 de 0.7698 para a radiação solar - Figura 18 (d) - e um R^2 de 0.9775 para a radiação solar extraterrestre - Figura 18 (e) - e, para os parâmetros de umidade média - Figura 18 (f) - o R^2 ficou em 0.7347 e do parâmetro de velocidade do vento - Figura 18 (g) - em 0.2541 .

A partir destas avaliações é possível observar que os parâmetros climáticos obtidos pela base de dados NASA/POWER possuem uma relação adequado, quando comparados os resultados obtidos pela avaliação dos parâmetros de radiação solar extraterrestre (R_a), temperatura média (T_m), temperatura mínima (T_{min}), radiação solar (R_s), umidade relativa (RH) e temperatura máxima (T_{max}), enquanto a relação entre os parâmetros de velocidade do vento (u_2) ainda se mostra um desafio para os modelos que determinam tais parâmetros a partir de imagens de satélite. A Tabela 4 exibe os resultados obtidos.

Tabela 4 – Resultados do coeficiente de determinação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER para o experimento de Brotas

Parâmetros climáticos INMET Parâmetros climáticos NASA/POWER	R^2
T_{max}, T_{max}	0.6232
T_{min}, T_{min}	0.8176
T_m, T_m	0.8352
R_s, R_s	0.7698
R_a, R_a	0.9775
RH, RH	0.7347
u_2, u_2	0.2541

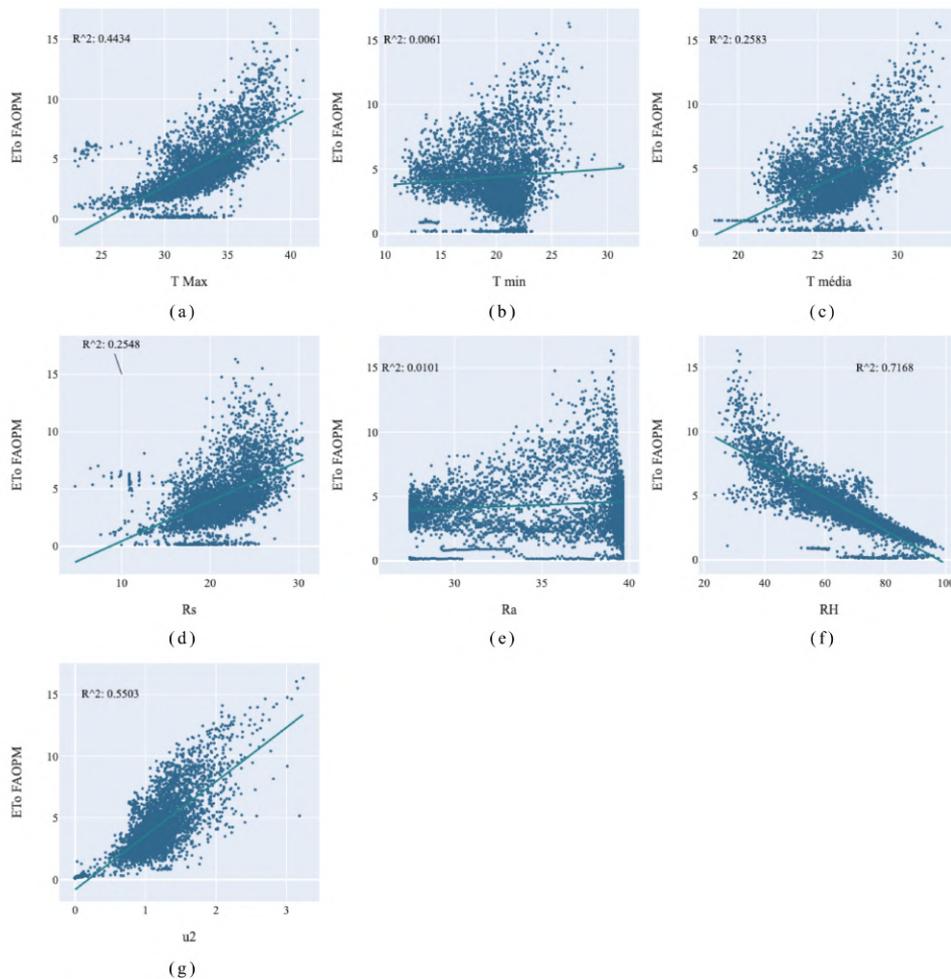
Fonte: Autor, 2020

5.2.2 Experimento de MATOPIBA

Os resultados da execução do primeiro processo, elaborado no capítulo 4 e ilustrado na Figura 10, realtivo ao experimento de MATOPIBA é elaborado a seguir.

A Figura 19 exibe os resultados obtidos do coeficiente de determinação utilizado entre os parâmetros climáticos obtidos da base de dados do INMET e a ET_o calculada para o experimento de MATOPIBA.

Figura 19 – Gráficos de avaliação dos parâmetros climáticos INMET e a ET_o calculada no experimento MATOPIBA



Fonte: Autor, 2020

O parâmetro de T_{max} - Figura 19 (a) - obteve um R^2 de 0.4434, para esta região, enquanto o R^2 do parâmetro T_{min} - Figura 19 (b) - resultou em 0.0061 e o R^2 do parâmetro T_m - Figura 19 (c) - ficou em 0.2583.

Os resultados do coeficiente de determinação dos parâmetros R_s - Figura 19 (d) - e R_a - Figura 19 (e) - com a ET_o ficaram em 0.2548 e 0.0101 respectivamente, enquanto o parâmetro RH_m - Figura 19 (f) - obteve um R^2 de 0.7168 e o parâmetro u_2 - Figura 19 (g) - ficou em 0.5503.

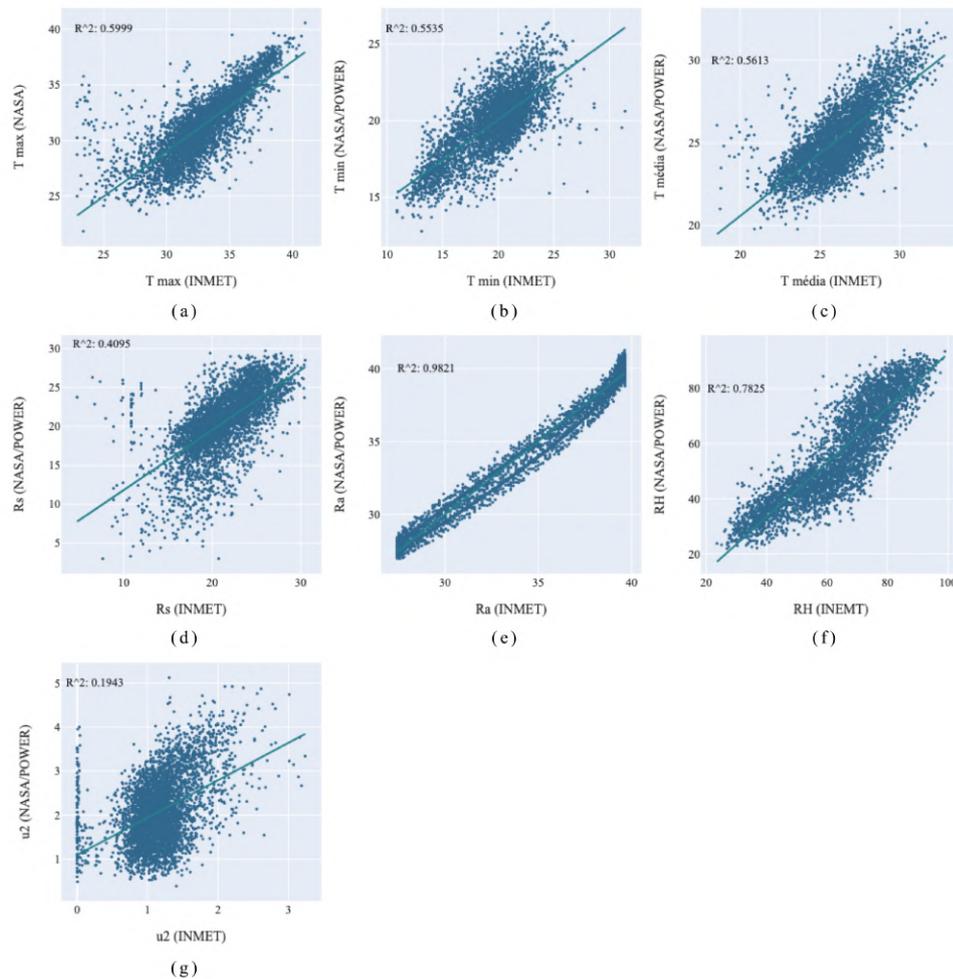
Estes resultados demonstram que a evapotranspiração de referência (ET_o) do local possui uma relação maior com os parâmetros de umidade relativa (RH), velocidade do vento (u_2), temperatura máxima (T_{max}), temperatura média (T_m) e radiação solar (R_s). Entretanto os parâmetros de radiação solar extraterrestre (R_a) e temperatura mínima (T_{min}) possuem uma relação menor com a ET_o do local. A Tabela 5 exibe os resultados obtidos.

Tabela 5 – Resultados do coeficiente de determinação entre a ET_o calculada e os parâmetros climáticos INMET para o experimento de MATOPIBA

ET_o calculada Parâmetros climáticos INMET	R^2
ET_o, T_{max}	0.4434
ET_o, T_{min}	0.0061
ET_o, T_m	0.2583
ET_o, R_s	0.2548
ET_o, R_a	0.0101
ET_o, RH	0.7168
ET_o, u_2	0.5503

Fonte: Autor, 2020

Figura 20 – Gráficos de avaliação parâmetros climáticos NASA/POWER e INMET no experimento MATOPIBA



Fonte: Autor, 2020

A Figura 20 exibe que o coeficiente de determinação dos parâmetros de temperatura resultou em um R^2 de 0.599, 0.5535 e 0.5613, T_{max} - Figura 20 (a) - T_{min} - Figura 20 (b) - e T_m - Figura 20 (c) -, respectivamente.

Os resultados para os parâmetros de radiação mostram um R^2 de 0.4095 para o parâmetro de radiação solar - Figura 20 (d) - e 0.9821 para o parâmetro de radiação solar extraterrestre - Figura 20 (e) -. O resultado do parâmetro de umidade - Figura 20 (f)- ficou em 0.7825 e para o parâmetro de velocidade do vento - Figura 20 (g)- resultou em um R^2 de 0.1943.

A partir destas avaliações é possível inferir que os parâmetros de radiação solar extraterrestre (R_a), umidade relativa (RH), temperatura máxima (T_{max}), temperatura média (T_m), temperatura mínima (T_{min}) e radiação solar (R_s) são melhores modelados a partir dos imagens

de satélites para esta região ao contrário do parâmetro de velocidade do vento (u_2), que seguiu o mesmo padrão da avaliação do experimento de Brotas. A Tabela 6 exibe os resultados obtidos.

Tabela 6 – Resultados do coeficiente de determinação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER para o experimento de MATOPIBA

Parâmetros climáticos INMET Parâmetros climáticos NASA/POWER	R^2
T_{max}, T_{max}	0.5999
T_{min}, T_{min}	0.5535
T_m, T_m	0.5613
R_s, R_s	0.4095
R_a, R_a	0.9821
RH, RH	0.7825
u_2, u_2	0.1943

Fonte: Autor, 2020

5.2.3 Definição dos conjuntos de parâmetros climáticos

Por fim, o ultimo resultado da execução do primeiro processo, tem-se as definições dos conjuntos de parâmetros climáticos que foram utilizados nos experimentos de Brotas e MATOPIBA. Os conjuntos são apresentados na Tabela 7 a seguir.

Tabela 7 – Tabela de conjuntos de parâmetros utilizados

Conjunto	Parâmetros Climáticos
C1	$T_{max}, T_{min}, T_m, R_a, R_s, RH, u_2$
C2	$T_{max}, T_{min}, T_m, R_a, R_s, RH$
C3	$T_{max}, T_{min}, T_m, R_a, R_s$

Fonte: Autor, 2020

5.3 Arquitetura dos Algoritmos

Conforme o segundo processo, elaborado no capítulo 4 e ilustrado na Figura 11, é necessário avaliar as melhores configurações de arquiteturas para modelos SVM e MLP, por tentativa e erro a (conforme elaborado no capítulo 4). Sendo assim, as sessões a seguir apresentam os resultados encontrados.

5.3.1 Arquitetura SVM

A arquitetura do modelo Máquina de Vetores de Suporte (SVM) selecionada consiste em um algoritmo de kernel de base radial (RBF), com o parâmetro $C = 5$, o parâmetro $\epsilon = 0.1$. Arquiteturas semelhantes podem ser encontradas nas pesquisas de Pandey, Nyori e Pandey (2017), Tangune e Escobedo (2018), Ferreira et al. (2019a) e Gocić et al. (2015), que utilizam o algoritmo SVM em regiões climáticas semelhantes e com os mesmos parâmetros climáticos utilizados nesta pesquisa.

5.3.2 Arquitetura MLP

A arquitetura do algoritmo MLP consiste em duas camadas escondidas de 10 neurônios cada uma, tendo como função de ativação a função tangente hiperbólica, denotada pela equação 2. Para a calibração do MLP, o algoritmo de treinamento estimativa adaptativa de momento (Adam) foi selecionado. Arquiteturas semelhantes são encontradas nas pesquisas realizadas por Tezel e Buyukyildiz (2016), Nourani, Elkiran e Abdullahi (2019), Ferreira et al. (2019a) e Ferreira e Cunha (2020), nas quais tais arquiteturas são utilizadas com os mesmos parâmetros climáticos além de serem utilizadas nas mesmas regiões climáticas desta pesquisa.

5.4 Treinamento e validação dos Modelos no Local de Treinamento

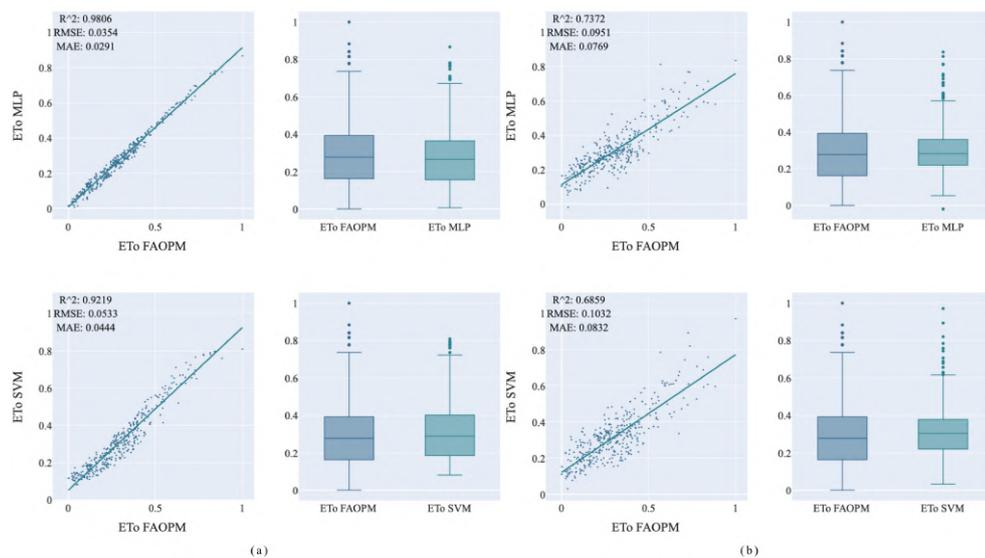
Os resultados obtidos na segunda etapa Figura 11 a partir dos modelos MLP e SVM, treinados com os conjuntos de parâmetros climáticos $C1$, são exibidos na Figura 21 para o experimento de Brotas e na Figura 24 para o experimento de MATOPIBA. Os resultados obtidos com o conjunto de parâmetros $C2$ são exibidos na Figura 22 para o experimento de Brotas e na Figura 25 para o experimento de MATOPIBA. Por fim os resultados obtidos com o conjunto $C3$ são exibidos na Figura 23 para o experimento de Brotas e na Figura 26 para o experimento de MATOPIBA.

5.4.1 Experimento de Brotas

Os resultados da execução do segundo processo no experimento de Brotas é elaborado a seguir.

A Figura 21 (a) exibe os resultados obtidos pelos modelos MLP e SVM treinados com o conjunto de parâmetros $C1$ obtidos a partir de da estação meteorológica INMET selecionada na primeira etapa da pesquisa. Para este caso, o coeficiente de determinação R^2 do modelo MLP ficou em 0.9806, enquanto as métricas $RMSE$ e MAE ficaram em 0.0354 e 0.0291, respectivamente. As métricas do modelo SVM ficaram com R^2 em 0.9219, $RMSE$ em 0.0533 e MAE em 0.0444.

Figura 21 – Avaliações dos modelos que utilizaram-se do conjunto $C1$ no experimento de Brotas



Fonte: Autor, 2020

Os modelos treinados com o mesmo conjunto de dados, porém obtidos a partir de satélites, exibido na Figura 21 (b), obtiveram as métricas R^2 , $RMSE$ e MAE em 0.7372, 0.0951 e 0.0769, para o modelo MLP e 0.6859, 0.1032 e 0.0832 para o modelo SVM.

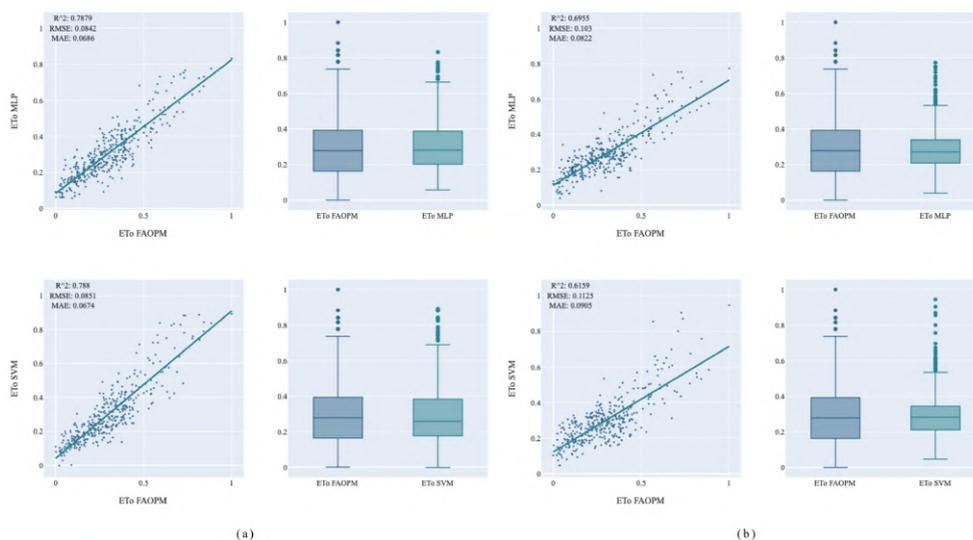
Tais resultados elaboram que os modelos de fato podem ser utilizados para modelar a evapotranspiração de referência, obtendo-se resultados semelhantes ao método original (FAOPM). Contudo pode ser observado aqui uma insuficiência nos resultados obtidos com os modelos que utilizaram-se dos parâmetros climáticos obtidos por satélites. A tabela 8 exibe os resultados obtidos pelos modelos, no local de teste do experimento de Brotas.

Tabela 8 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C1$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C1	R^2	0.9806	0.9219	0.7372	0.6859
	RMSE	0.0354	0.0533	0.0951	0.1032
	MAE	0.0291	0.0444	0.0769	0.0832

Fonte: Autor, 2020

Figura 22 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de Brotas



Fonte: Autor, 2020

Os resultados dos modelos treinados com o conjunto de parâmetros climáticos $C2$, obtidos a partir de uma estação meteorológica, são exibidos na Figura 22 (a). O modelo MLP ficou com um R^2 em 0.7879, $RMSE$ em 0.0842 e MAE em 0.0686. O modelo SVM ficou com R^2 em 0.788, $RMSE$ em 0.0851 e MAE em 0.0674.

A Figura 22 (b) exhibe os resultados obtidos pelos modelos MLP e SVM treinados com o mesmo conjunto de parâmetros climáticos, porém, obtidos por satélites. Neste caso o modelo MLP ficou com um R^2 em 0.6955, $RMSE$ em 0.103 e MAE em 0.0822, enquanto o modelo SVM ficou com as métricas R^2 em 0.6159, $RMSE$ em 0.1125 e MAE em 0.0905.

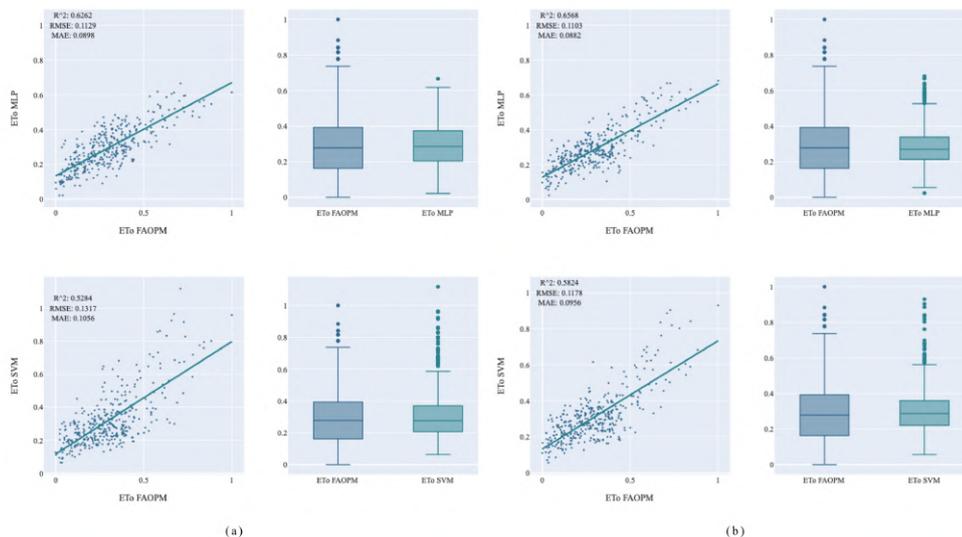
O uso do conjunto de parâmetros $C2$ já mostra uma queda na acurácia dos modelos, onde apesar dos modelos serem capazes de modelar a evapotranspiração de referência do local, os medidas tendem a ser subestimadas, tanto para o uso de dados de estações meteorológicas, quanto para o uso de dados de satélite. A tabela 9 exhibe os resultados obtidos pelos modelos, no local de teste do experimento de Brotas, utilizando-se do conjunto de parâmetros $C2$.

Tabela 9 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C2$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C2	R^2	0.7879	0.788	0.6955	0.6159
	RMSE	0.0842	0.0851	0.103	0.1125
	MAE	0.0686	0.0674	0.0822	0.0905

Fonte: Autor, 2020

Figura 23 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de Brotas



Fonte: Autor, 2020

Na figura 23 (a) são exibidos os resultados obtidos pelos modelos MLP e SVM treinados com o conjunto de parâmetros $C3$. O modelo MLP, treinado com os parâmetros obtidos por uma estação meteorológica, obteve as métricas R^2 em 0.6262, $RSME$ em 0.1129 e MAE em

0.0898, enquanto o modelo, SVM que utilizou da mesma fonte de dados, resultou nas métricas R^2 em 0.5284, $RSME$ em 0.1317 e MAE em 0.1056.

A Figura 23 (b) exibe os resultados do modelo MLP treinado com o conjunto de parâmetros $C3$, obtidos por satélites. Neste caso, o modelo MLP obteve as métricas R^2 em 0.6568, $RSME$ em 0.1103 e MAE em 0.0882, assim como o modelo SVM ficou com as métricas R^2 em 0.5824, $RSME$ em 0.1178 e MAE em 0.0956.

Seguindo o padrão do uso de conjunto de parâmetros $C2$, os modelos treinados com o conjunto de parâmetros $C3$ também tem a tendência a subestimar as medidas de evapotranspiração. Contudo neste caso há um resultado que deve ser observado. Ao contrário dos resultados anteriores, os modelos de aprendizado de máquina treinados com dados de satélites possuem métricas melhores do que os mesmos modelos treinados com dados de estações meteorológicas. A tabela 10 exibe os resultados obtidos pelos modelos, no local de teste do experimento de Brotas, utilizando-se do conjunto de parâmetros $C3$.

Tabela 10 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C3$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C3	R^2	0.6262	0.5284	0.6568	0.5824
	RMSE	0.1129	0.1317	0.1103	0.1178
	MAE	0.0898	0.1056	0.0882	0.0956

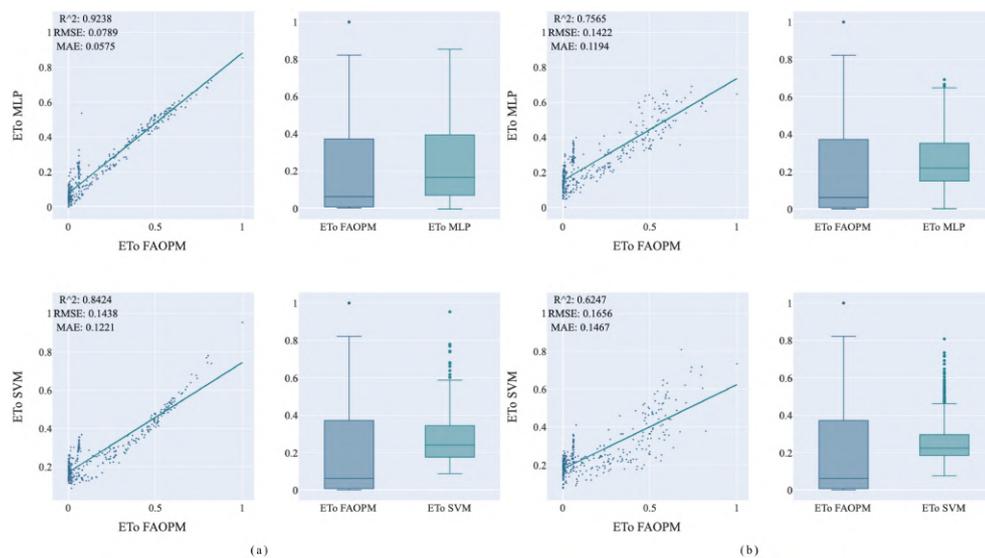
Fonte: Autor, 2020

5.4.2 Experimento de MATOPIBA

Os resultados obtidos pela execução do segundo processo no experimento de MATOPIBA é elaborado a seguir.

A Figura 24 (a) exibe os resultados obtidos pelos modelos treinados com os conjunto de parâmetros climáticos $C1$, estes obtidos a partir de uma estação meteorológica. O modelo MLP ficou com as métricas R^2 em 0.9238, $RMSE$ em 0.0789 e MAE 0.0575. O modelo SVM, treinado com o mesmo conjunto de parâmetros obteve as métricas R^2 em 0.8424, $RMSE$ em 0.1438 e MAE em 0.1221.

Figura 24 – Avaliações dos modelos que se utilizaram do conjunto $C1$ no experimento de MATOPIBA



Fonte: Autor, 2020

Os modelos treinados com o conjunto de parâmetros $C1$ obtidos a partir de satélites são exibidos na Figura 24 (b). Neste caso o modelo MLP obteve as métricas R^2 , $RSME$ e MAE em 0.7565, 0.1422 e 0.1194, respectivamente, enquanto o modelo SVM, obteve as métricas R^2 , $RSME$ e MAE em 0.6247, 0.1656 e 0.1467.

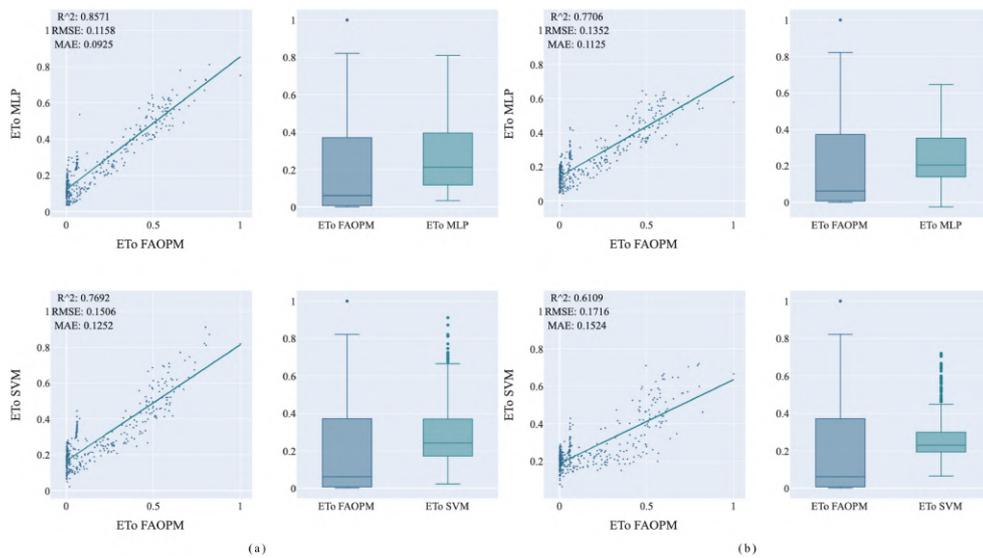
Neste experimento os modelos seguiram o mesmo padrão de conseguir modelar a evapotranspiração de referência com acurácia, quando utilizado todos os parâmetros de estações meteorológicas. Os modelos treinados com dados de satélites mostraram uma acurácia melhor do que a obtida com o mesmo conjunto de parâmetros no experimento de Brotas, contudo a insuficiência destes dados ainda persiste. A tabela 11 exibe os resultados obtidos pelos modelos, no local de teste do experimento de MATOPIBA, utilizando-se do conjunto de parâmetros $C1$.

Tabela 11 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C1$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C1	R^2	0.9238	0.8424	0.7565	0.6247
	RMSE	0.0789	0.1438	0.1422	0.1656
	MAE	0.0575	0.1221	0.1194	0.1467

Fonte: Autor, 2020

Figura 25 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de MATOPIBA



Fonte: Autor, 2020

A Figura 25 (a) exibe os resultados obtidos pelos modelos MLP e SVM, treinados com o conjunto de dados $C2$, obtidos por uma estação meteorológica. Neste caso o modelo MLP obteve as métricas, R^2 de 0.8571, $RMSE$ de 0.1158 e MAE de 0.0925, enquanto modelo SVM obteve as métricas R^2 de 0.7692, $RMSE$ de 0.1106 e MAE de 0.1252.

Os resultados dos modelos treinados com o mesmo conjunto de dados, porém obtidos a partir de satélites, são exibidos na Figura 25 (b). Neste caso o modelo MLP obteve as métricas R^2 , $RMSE$ e MAE de 0.7706, 0.1352 e 0.1125, respectivamente, enquanto o modelo SVM treinado com o mesmo conjunto de parâmetros, obtidos a partir de satélites, obteve as métricas R^2 de 0.6109, $RMSE$ de 0.1716 e MAE de 0.1524.

A partir destas avaliações é possível observar, que assim como no experimento de Brotas, há uma queda na acurácia dos modelos treinados com os parâmetros climáticos obtidos por satélites o que reforça a incapacidade de somente os parâmetros $C2$ serem utilizados para modelar a evapotranspiração de referência deste local. Contudo houve uma ligeira melhora no modelo MLP treinado com os parâmetros de satélite, indicando que os parâmetros do conjunto $C2$ podem ser melhores quando obtidos por esta tecnologia.

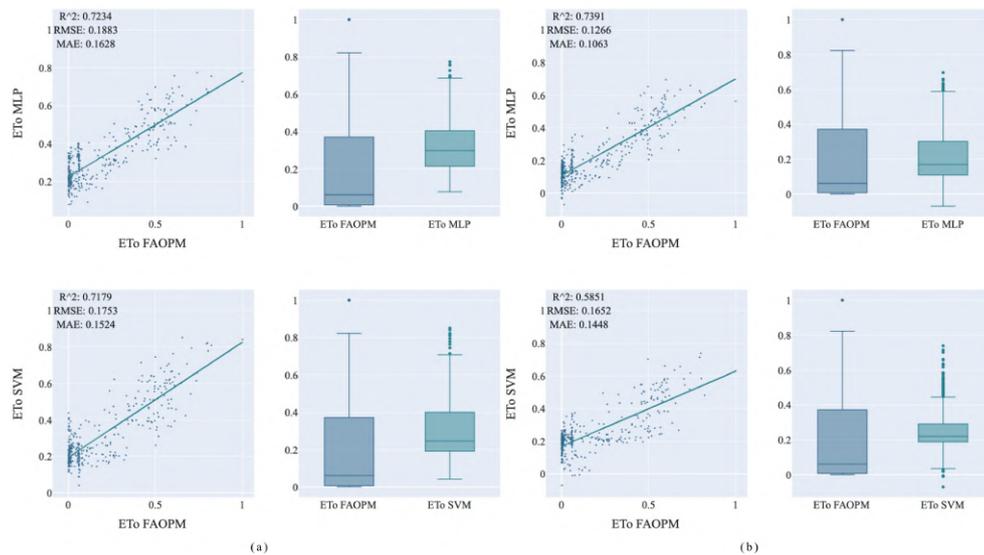
A tabela 12 exibe os resultados obtidos pelos modelos, no local de teste do experimento de MATOPIBA, utilizando-se do conjunto de parâmetros $C2$.

Tabela 12 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C2$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C2	R^2	0.8571	0.7692	0.7706	0.6109
	RMSE	0.1158	0.1106	0.1352	0.1716
	MAE	0.0925	0.1252	0.1125	0.1524

Fonte: Autor, 2020

Figura 26 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de MATOPIBA



Fonte: Autor, 2020

A Figura 26 (a) exibe os resultados obtidos pelos modelos treinados com o conjunto de parâmetros $C3$, recuperados a partir de uma estação meteorológica. Neste caso, o modelo MLP obteve as métricas R^2 , $RMSE$ e MAE de 0.7234, 0.1883 e 0.1628, enquanto o modelo SVM obteve as métricas R^2 , $RMSE$ e MAE de 0.7179, 0.1753 e 0.1524.

Os resultados obtidos pelos modelos, treinados com o mesmo conjunto de dados porém recuperados por satélites, são exibidos na Figura 26 (a). Neste caso, o modelo MLP obteve as métricas R^2 em 0.7391, $RMSE$ em 0.1266 e MAE em 0.1063, enquanto o modelo SVM obteve as métricas R^2 em 0.5851, $RMSE$ em 0.1652 e MAE em 0.1448.

Estes resultados reforçam os resultados obtidos no experimento de Brotas, no qual os modelos treinados com o conjunto de parâmetros $C3$ possuem maior acurácia do que os mode-

los treinados com este conjunto de parâmetros, porém com dados obtidos por estações meteorológicas. Ainda assim deve-se observar que mesmo neste caso a evapotranspiração de referência modelada é subestimada pelos modelos.

A tabela 13 exhibe os resultados obtidos pelos modelos, no local de teste do experimento de MATOPIBA, utilizando-se do conjunto de parâmetros *C3*.

Tabela 13 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros *C3*

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C3	R^2	0.7234	0.7179	0.7391	0.5851
	RMSE	0.1883	0.1753	0.1266	0.1656
	MAE	0.1628	0.1524	0.1063	0.1448

Fonte: Autor, 2020

5.5 Aplicação dos modelos no local de teste

Conforme elaborado na terceira etapa ilustrado na Figura 12, os modelos aprendizagem de máquina treinados no local de treinamento são aplicados no local de teste, para modelar a ET_o do local.

Os resultados obtidos para a terceira etapa, para o experimento de Brotas, são apresentados nas Figuras 28, 29 e 30. Os resultados desta etapa, para o experimento de MATOPIBA, são apresentados nas Figuras Figuras 32, 33 e 34.

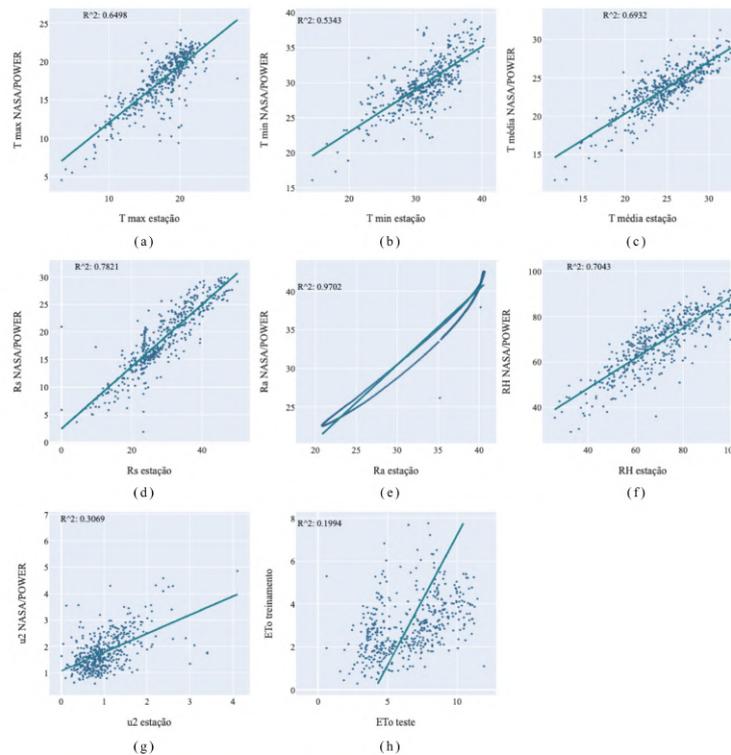
5.5.1 Experimento de Brotas

Os resultados obtidos a partir da execução do terceiro processo no experimento de Brotas é elaborado a seguir.

Os resultados das avaliações de relação de parâmetros climáticos no local de teste são apresentados na Figura 27 (a). A partir destas avaliações é possível notar uma relação entre os parâmetros climáticos obtidos pela base de dados NASA/POWER e a estação local, onde o parâmetro T_{max} - Figura 27 (a) - apresentou um coeficiente R^2 de 0.6498, o parâmetro T_{min} - Figura 27 (b) - um coeficiente R^2 de 0.5343, enquanto o parâmetro T_m - Figura 27 (c) -

apresentou um R^2 de 0.6932, o parâmetro R_s - Figura 27 (d) - um coeficiente de determinação de 0.7821, o parâmetro R_a - Figura 27 (e) - obteve um R^2 de 0.9702, o parâmetro RH - Figura 27 (f) - um R^2 de 0.7043 e u_2 - Figura 27 (g) - coeficiente R^2 de 0.3069.

Figura 27 – Gráficos de avaliação de parâmetros climáticos NASA/POWER e estação local e ET_o do local de treinamento e teste do experimento Brotas



Fonte: Autor, 2020

Contudo, a avaliação de relação entre o parâmetro de ET_o do local de teste e o local de treinamento - Figura 27 (h) - mostrou um baixo coeficiente de determinação que resultou em 0.1994. A Tabela 14 resume os resultados obtidos.

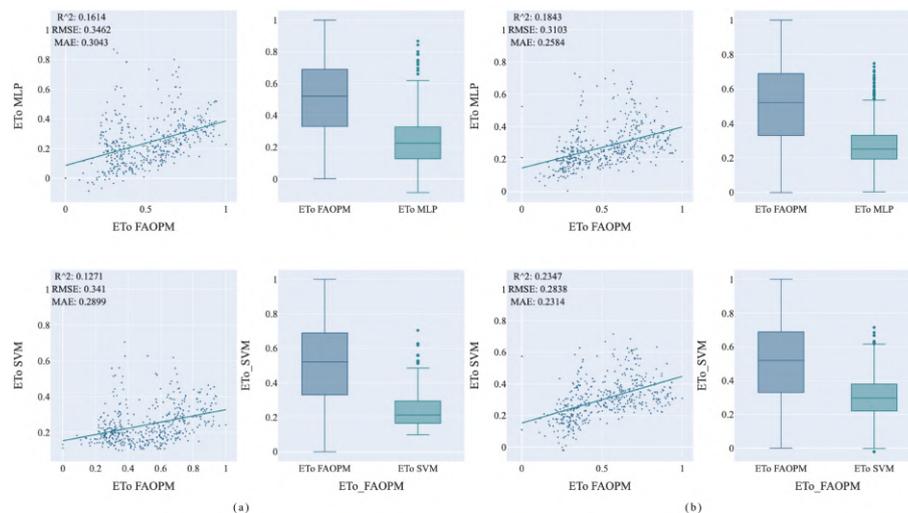
Tabela 14 – Avaliações de relação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER e a evapotranspiração de referência do local de teste e treinamento no experimento de Brotas

Parâmetros climáticos INMET Parâmetros climáticos NASA/POWER	R^2
T_{max}, T_{max}	0.6498
T_{min}, T_{min}	0.5343
T_m, T_m	0.6932
R_s, R_s	0.7821
R_a, R_a	0.9702
RH, RH	0.7043
u_2, u_2	0.3069
ET_o, ET_o	0.1994

Fonte: Autor, 2020

A aplicação do modelo MLP e SVM utilizando-se do conjunto de dados $C1$ no experimento de Brotas é ilustrado na Figura 28 (a). Os modelos treinados com os dados do INMET na estação do local de treinamento e aplicado no local de teste com os dados da estação meteorológica resultou nas métricas R^2 de 0.1614, $RMSE$ de 0.3462 e MAE de 0.3043 para o modelo MLP aplicado e métricas R^2 de 0.1271, $RMSE$ de 0.341 e MAE de 0.2899 para o modelo SVM.

Figura 28 – Avaliações dos modelos que utilizaram-se do conjunto $C1$ no experimento de Brotas



Fonte: Autor, 2020

Os modelos de aprendizagem de máquina treinados com o conjunto de parâmetros $C1$, com dados obtidos a partir da base de dados NASA/POWER são apresentados na Figura 28 (b).

Neste caso o modelo MLP obteve métricas R^2 de 0.1843, $RMSE$ de 0.3103 e MAE de 0.2584. O modelo SVM, por sua vez, obteve as métricas R^2 de 0.2347, $RMSE$ de 0.2838 e MAE de 0.2314.

Ao contrário das simulações realizadas no local de treinamento, os resultados dos modelos aplicados no local de teste mostram diferenças menores entre os modelos que utilizam-se dos dados da estação local e os dados da base de dados NASA/POWER do local. A Tabela 15 resume as métricas obtidas nestes testes.

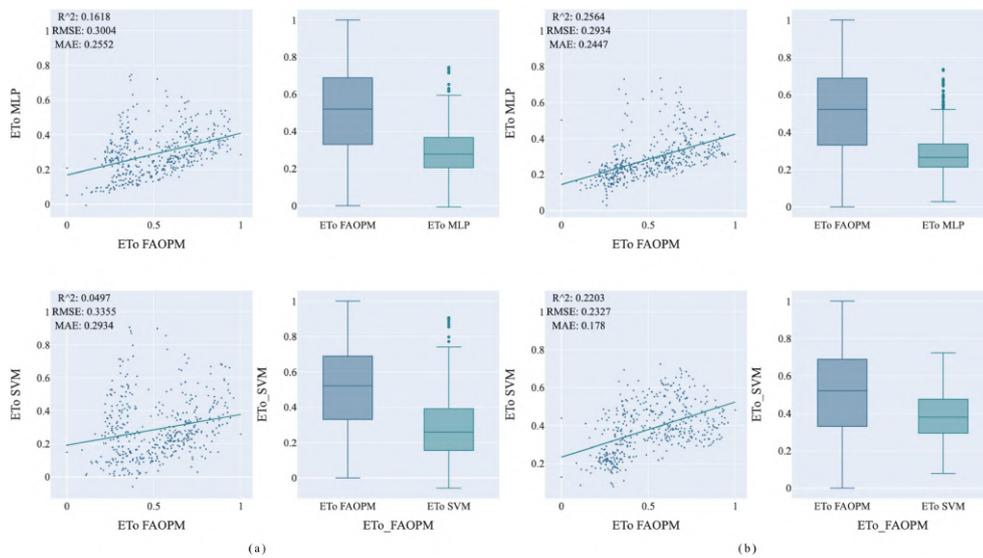
Tabela 15 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C1$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C1	R^2	0.1614	0.1271	0.1843	0.2347
	RMSE	0.3462	0.2899	0.3103	0.2314
	MAE	0.3043	0.2899	0.2584	0.2314

Fonte: Autor, 2020

Os resultados do experimento de Brotas em que houve a aplicação dos modelos MLP e SVM no local de teste utilizando-se do conjunto de parâmetros $C2$ e foram treinados com a base de dados INMET são apresentados na 29 (a) a seguir. Nesta simulação o modelo MLP obteve as métricas R^2 de 0.1618, $RMSE$ de 0.3004 e MAE de 0.2552. A aplicação do modelo SVM resultou nas métricas R^2 de 0.0497, $RMSE$ de 0.3355 e MAE de 0.2943.

Figura 29 – Avaliações dos modelos que se utilizaram do conjunto $C2$ no experimento de Brotas



Fonte: Autor, 2020

Quando treinados com o conjunto de dados $C2$, porém com os dados obtidos da base de dados NASA/POWER, as métricas do modelo MLP resultam em um R^2 de 0.2564, $RMSE$ de 0.2934 e MAE de 0.2447. O modelo SVM para o mesmo conjunto de parâmetros resultou nas métricas R^2 de 0.2203, $RMSE$ de 0.2327 e MAE de 0.178.

Neste caso os resultados mostram uma performance melhor do lado dos modelos que se utilizando dos dados obtidos pela base de dados NASA/POWER, ainda sim os resultados mostram-se inferiores aqueles obtidos no melhor caso do mesmo modelo aplicado no local de teste. A Tabela 16 resume os resultados das métricas obtidas nesta simulação.

Tabela 16 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C2$

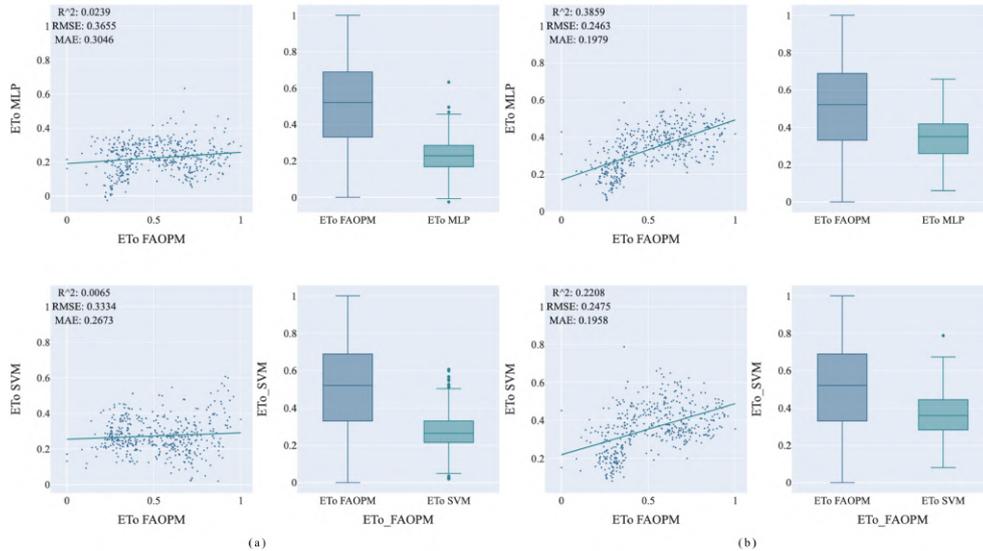
Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C2	R^2	0.1618	0.0497	0.2564	0.2203
	RMSE	0.3004	0.3355	0.2934	0.2327
	MAE	0.2552	0.2934	0.2447	0.178

Fonte: Autor, 2020

O resultado da aplicação dos modelo de aprendizagem de máquina utilizando-se do conjunto $C3$ com dados obtidos a partir das estações meteorológicas é ilustrado na Figura 30

(a) a seguir. O modelo MLP obteve as métricas R^2 de 0.0239, $RMSE$ de 0.3655 e MAE de 0.3046, enquanto o modelo SVM obteve as métricas R^2 de 0.0065, $RMSE$ de 0.3334 e MAE de 0.2673, mostrando uma relação quase inexistente ente os resultados dos modelos e a ET_o do local.

Figura 30 – Avaliações dos modelos que utilizaram-se do conjunto $C3$ no experimento de Brotas



Fonte: Autor, 2020

O modelo MLP aplicado com o conjunto de parâmetros $C3$ e dados obtidos a partir da base de dados NASA/POWER mostrou métricas melhores, quando comparado com a aplicação do modelo com os dados da estação meteorológica local. Tais métricas resultaram em um R^2 de 0.3859, $RMSE$ de 0.2463 e MAE de 0.1979. O modelo SVM também obteve métricas melhores, na mesma comparação, onde tais métricas resultaram em um R^2 de 0.2208, $RMSE$ de 0.2475 e MAE de 0.1958. A Tabela 17 resume os resultados das métricas obtidas.

Tabela 17 – Resultados obtidos pelos modelos no local de teste do experimento de Brotas utilizando-se do conjunto de parâmetros $C3$

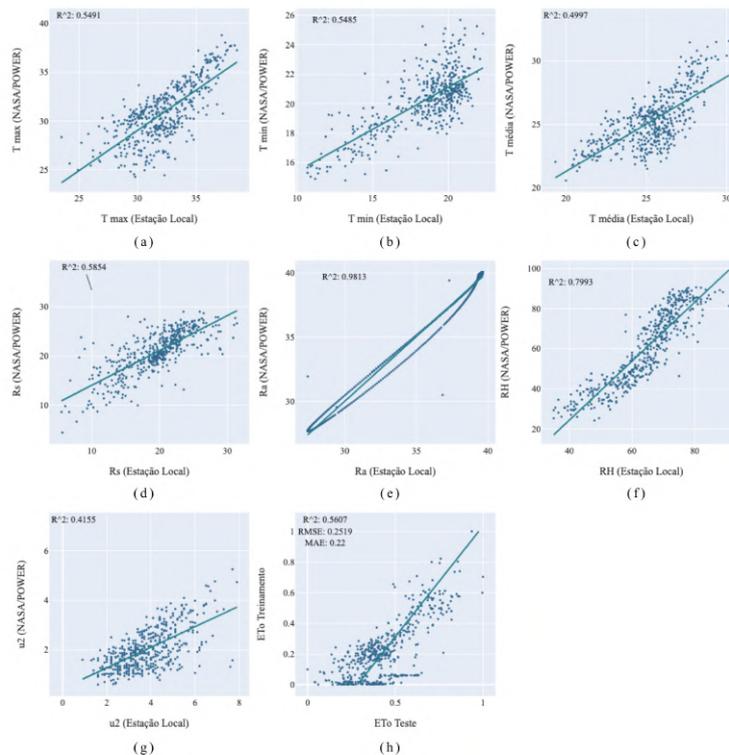
Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C3	R^2	0.0239	0.0065	0.3859	0.2208
	RMSE	0.3655	0.3334	0.2463	0.2475
	MAE	0.3046	0.2673	0.1979	0.1958

Fonte: Autor, 2020

5.5.2 Experimento de MATOPIBA

Os resultados das avaliações de parâmetros realizadas no local de teste do experimento de MATOPIBA são apresentados na Figura 31 a seguir. Neste experimento os parâmetros T_{max} - Figura 31 (a) - mostrou um coeficiente de determinação R^2 de 0.5491, o parâmetro T_{min} - Figura 31 (b) -, resultou em um R^2 de 0.5485, enquanto o parâmetro T_m , Figura 31 (c), obteve um R^2 de 0.4997. Os parâmetros R_s , - Figura 31 (d) -, e R_a , - Figura 31 (f) -, obtiveram um parâmetro R^2 de 0.5854 e 0.9813, respectivamente. O parâmetro RH , por sua vez, resultou em um R^2 de 0.7993 e o parâmetro R^2 de 0.4155.

Figura 31 – Gráficos de avaliação de parâmetros climáticos NASA/POWER e estação local e ET_o do local de treinamento e teste do experimento MATOPIBA



Fonte: Autor, 2020

A avaliação dos parâmetros de evapotranspiração de referência entre o local de teste e o local de treinamento resultou em um R^2 de 0.5607, mostrando uma relação melhor do que aquela realizada entre os parâmetros de ET_o do experimento de Brotas. A Tabela 18 resume os resultados obtidos pela avaliação de relação de parâmetros.

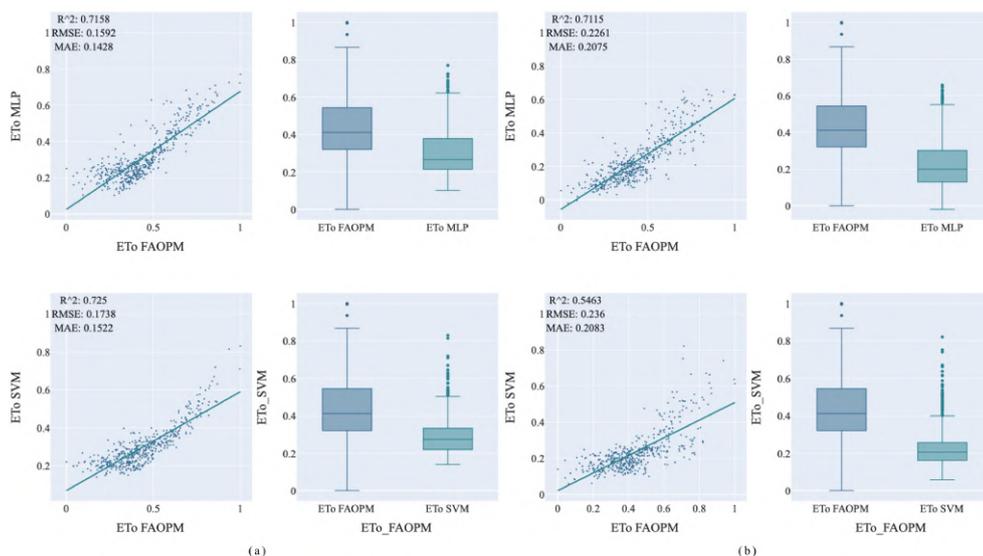
Tabela 18 – Resultados da avaliação de relação entre os parâmetros climáticos INMET e os parâmetros climáticos NASA/POWER e a ET_o do local de treinamento e teste do experimento de MATOPIBA

Parâmetros climáticos INMET Parâmetros climáticos NASA/POWER	R^2
T_{max}, T_{max}	0.5491
T_{min}, T_{min}	0.5485
T_m, T_m	0.4997
R_s, R_s	0.5854
R_a, R_a	0.9813
RH, RH	0.7993
u_2, u_2	0.4155
ET_o, ET_o	0.5607

Fonte: Autor, 2020

A simulação do experimento de MATOPIBA em que os modelos de aprendizagem de máquina foram aplicados com o conjunto de parâmetros $C1$ e os dados obtidos pela estação meteorológica local resultaram nas métricas R^2 de 0.7158, $RMSE$ de 0.1592 e MAE de 0.1428 para o modelo MLP e métricas R^2 de 0.725, $RMSE$ de 0.1738 e MAE de 0.1522 para o modelo SVM.

Figura 32 – Avaliações dos modelos que se utilizaram do conjunto $C1$ no experimento de MATOPIBA



Fonte: Autor, 2020

As simulações realizadas com o mesmo conjunto de parâmetros, porém com o uso de dados da base de dados NASA/POWER, resultaram nas métricas R^2 de 0.7115, $RMSE$ de

0.2261 e *MAE* de 0.2075 para o modelo MLP. A aplicação do modelo SVM nesta simulação obteve as métricas R^2 de 0.5463, *RMSE* de 0.236 e *MAE* de 0.2083.

Semelhante aos resultados das simulações no local treinamento, os resultados dos modelos em que os dados de estações meteorológicas são utilizando, geram métricas mais precisas do que aqueles que se utilizam de dados climáticos de satélites. A Tabela 19 resume os resultados obtidos neste cenário.

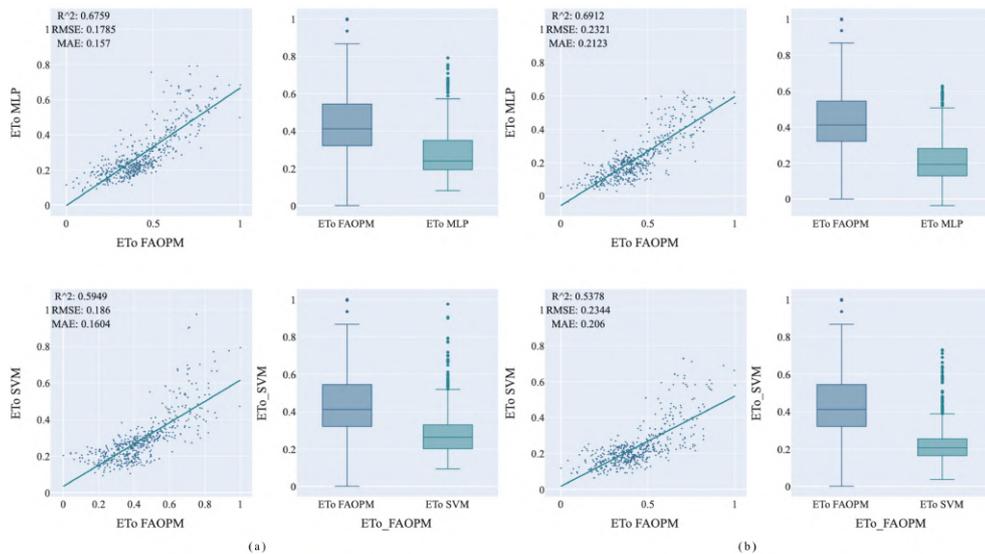
Tabela 19 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros C1

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C1	R^2	0.7158	0.725	0.7115	0.5463
	RMSE	0.1592	0.1738	0.2261	0.236
	MAE	0.1428	0.1522	0.2075	0.2083

Fonte: Autor, 2020

As simulações realizadas com o modelo MLP utilizando-se do conjunto C2 e os a base de dados da estação local, obteve as métricas R^2 de 0.6759, *RMSE* de 0.1785 e *MAE* de 0.157, enquanto o modelo SVM obteve as métricas R^2 de 0.5949, *RMSE* de 0.186 e *MAE* de 0.1604, para o mesmo cenário.

Figura 33 – Avaliações dos modelos que se utilizaram do conjunto C2 no experimento de MATOPIBA



Fonte: Autor, 2020

O modelo MLP, aplicado com dados obtidos a partir da base de dados NASA/POWER, obteve as métricas R^2 de 0.6912, $RMSE$ de 0.2321 e MAE de 0.2123, enquanto o modelo SVM obteve as métricas R^2 de 0.5378, $RMSE$ de 0.2344 e MAE de 0.206. A Tabela 20 resume os resultados obtidos neste cenário.

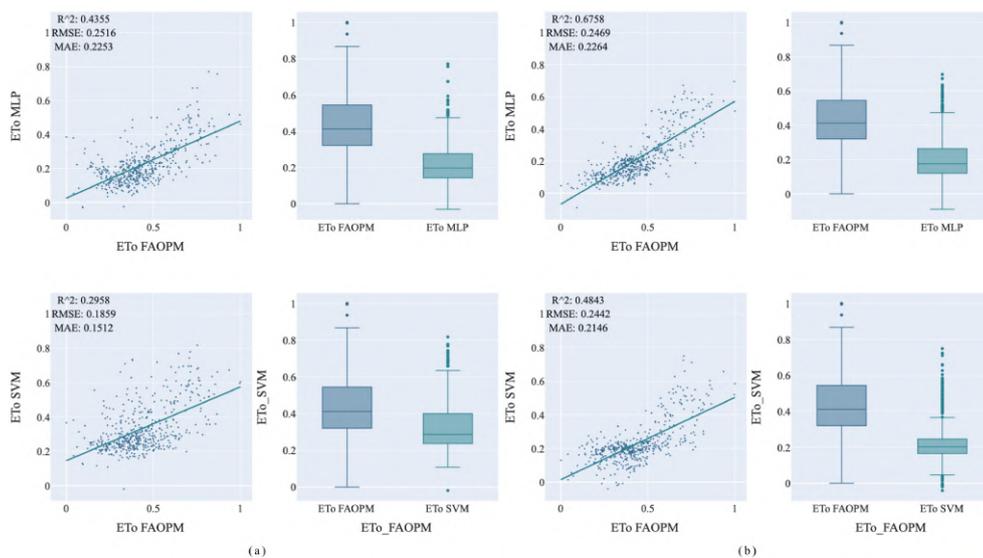
Tabela 20 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros $C2$

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C2	R^2	0.6759	0.5949	0.6912	0.5378
	RMSE	0.1785	0.186	0.2321	0.2344
	MAE	0.157	0.1604	0.2123	0.206

Fonte: Autor, 2020

Por fim os resultados obtidos pela aplicação do modelo MLP no local de teste, utilizando-se do conjunto de parâmetros $C3$ e os dados obtidos pela estação local, resultaram nas métricas R^2 de 0.4355, $RMSE$ de 0.2516 e MAE de 0.2253. O modelo SVM obteve as métricas R^2 de 0.2958, $RMSE$ de 0.1859 e MAE de 0.1512.

Figura 34 – Avaliações dos modelos que se utilizaram do conjunto $C3$ no experimento de MATOPIBA



Fonte: Autor, 2020

Os modelos de aprendizagem de máquina que se utilizaram do conjunto $C3$, porém com os dados obtidos a partir da base de dados NASA/POWER, resultaram nas métricas R^2 de

0.6758, *RMSE* de 0.4269 e *MAE* de 0.2264 para o modelo MLP, enquanto o modelo SVM obteve as métricas R^2 de 0.4843, *RMSE* de 0.2442 e *MAE* de 0.2146. A Tabela 21 resume os resultados obtidos neste cenário.

Tabela 21 – Resultados obtidos pelos modelos no local de teste do experimento de MATOPIBA utilizando-se do conjunto de parâmetros C3

Conjunto de Parâmetros	Métricas de Avaliação	Estação Meteorológica		Satélite	
		MLP	SVM	MLP	SVM
C3	R^2	0.4355	0.2958	0.6758	0.4843
	RMSE	0.2516	0.1859	0.2469	0.2442
	MAE	0.0.2253	0.1512	0.2264	0.2146

Fonte: Autor, 2020

A partir dos resultados mostrados na Tabela 21 é possível observar que o padrão dos modelos que utilizam-se de dados da base de dados NASA/POWER superam os modelos que se utilizam de dados de estações meteorológicas, contudo a aplicação dos modelos no local de teste resulta em métricas de raiz quadrada do erro médio e erro médio absoluto maiores.

6 Discussão

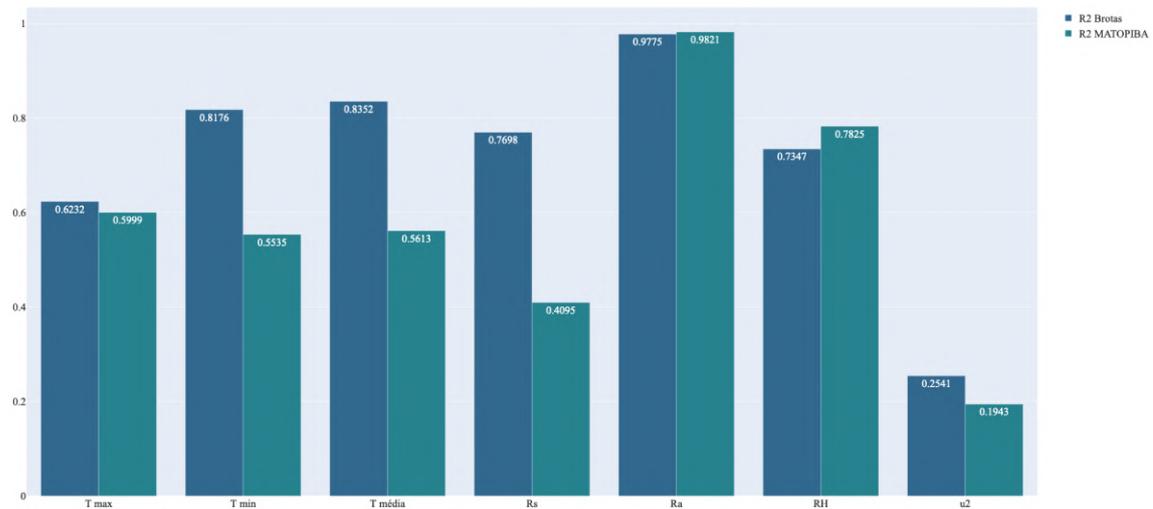
Os resultados obtidos pela avaliação dos parâmetros climáticos, elaborados nas subseções 5.2.1 e 5.2.2, seguem a mesma linha dos resultados evidenciados por Monteiro, Sentelhas e Pedra (2018), os quais elaboram que os parâmetros climáticos T_{max} , T_m , R_s e R_a obtidos pela base de dados NASA/POWER são satisfatórios, uma vez que estes podem ser facilmente derivados dos sensores instalados em satélites, contudo, os parâmetros de T_{min} , e u_2 ainda precisam de melhorias uma vez que a incidência de nuvens tende a obstruir as imagens o que levam a ser parâmetros mais imprecisos.

Apesar dos resultados elaborados por Monteiro, Sentelhas e Pedra (2018), o parâmetro RH obtido pela base de dados NASA/POWER mostrou um coeficiente de determinação superior a 0.7 em todos os locais, quando comparado com o mesmo parâmetro obtido por estações meteorológicas, o que indica uma melhoria nos modelos que estimam este parâmetro.

Nesta pesquisa os parâmetros obtidos pela base de dados NASA/POWER mostram R^2 acima de 0.7 para os parâmetros T_{min} , T_m , R_a , R_{xs} e RH e um R^2 menor que 0.5 para os parâmetros T_{max} , e u_2 , conforme elaborado na subseção 5.2.1 referente ao experimento de Brotas e ilustrado na Figura 35. Este resultado pode ser explicado devido a baixa precisão de parâmetros climáticos obtidos pela base de dados utilizada em regiões de altitude, conforme elaborado por Monteiro, Sentelhas e Pedra (2018), indicando que os modelos de reanálise destes parâmetros ainda necessitam de melhorias.

No experimento de MATOPIBA, os parâmetros climáticos obtidos pela base de dados NASA/POWER mostraram um coeficiente R^2 menor que 0.7 para os parâmetros T_{max} , T_{min} , T_m , R_s e u_2 , com exceção dos parâmetros R_a e RH , conforme elaborado na subseção 5.2.2 referente ao experimento de MATOPIBA e ilustrado na Figura 35 (b). Este resultado pode ser explicado pela interferência de nuvens nas medições de satélites, como elabora autores como Zhang et al. (2019), Khand et al. (2019) e Shafi et al. (2019b), a qual acaba sendo comum em regiões de clima tropical, como no caso do experimento de MATOPIBA.

Figura 35 – Coeficiente de determinação do experimento de Brotas e MATOPIBA



Fonte: Autor, 2020

Tais resultados evidenciam o potencial de uso de parâmetros climáticos obtidos a partir de satélites. Não obstante outro resultado promissor do uso deste parâmetros pode ser visto na pesquisa de Monteleone et al. (2020), no qual diferentes cenários de simulações foram avaliados para se obter a recomendação de irrigação em uma cultura de fibras.

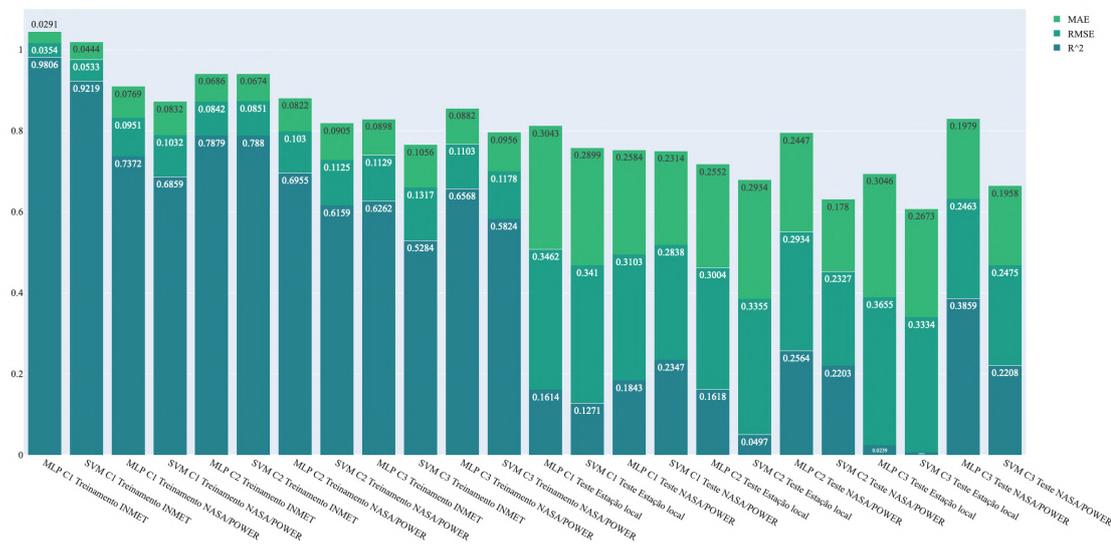
Entre os modelos de aprendizagem de máquina estudados, os melhores resultados foram obtidos pelo modelo MLP, apresentando os maiores valores da métrica R^2 e os menores valores das métricas $RMSE$ e MAE , para todos os cenários realizados no local de treinamento. O mesmo pode ser observado nas pesquisas de Adeloey, Rustum e Kariyama (2011), Goyal et al. (2014), Pandey, Nyori e Pandey (2017), Zhang, Gong e Wang (2018), Heddarn et al. (2018) e Ferreira et al. (2019a). Ainda assim, os resultados podem ser melhorados com a avaliação de modelos para a redução de ruídos e dados fora do padrão que podem ser gerados por erros nas medições de estações meteorológicas.

Dentre os diversos conjuntos de dados testados, os melhores resultados foram obtidos pelo modelos treinados com todos os parâmetros climáticos, representados pelo conjunto $C1$ e parâmetros obtidos por estações meteorológicas. Resultados semelhantes podem ser observados nas pesquisas de Antonopoulos e Antonopoulos (2017), Pandey, Nyori e Pandey (2017), Tangune e Escobedo (2018), dentre os quais obtiveram melhores resultados de seus modelos, utilizando-se da maior quantidade de parâmetros climáticos.

Os modelos treinados com o conjunto de parâmetros $C2$ mostrou melhores métricas com a partir do modelo MLP em todos dos cenários de simulação nos locais de treinamento, tanto para o experimento de Brotas, quanto para o experimento de MATOPIBA.

Os modelos treinados com o conjunto de dados $C3$ apresentaram resultados satisfatórios para a modelagem da ET_o , nos quais a medida de evapotranspiração de referência é modelada a partir de parâmetros climáticos limitados a temperatura e radiação, para todas as simulações nos locais de teste.

Figura 36 – Métricas de todos os experimentos de Brotas



Fonte: Autor, 2020

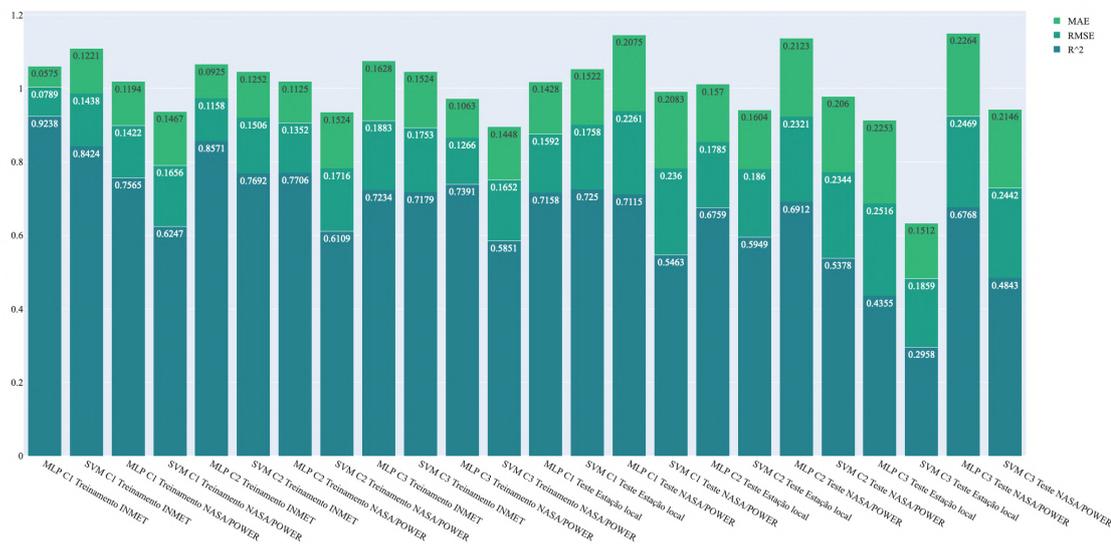
Um resultado interessante que pode ser observado, são os modelos treinados com o conjunto de parâmetros $C3$ com dados obtidos a partir da base de dados NASA/POWER, no qual forneceram métricas R^2 maiores e métricas $RMSE$ e MAE menores, quando comparados com as métricas dos mesmos modelos treinados com os parâmetros das estações meteorológicas. Tal padrão pode ser observado tanto no experimento de Brotas, quanto no experimento de MATOPIBA, para as simulações realizadas no local de treinamento, como pode ser observado na Figura 36 e na Figura 37.

Este resultado estende a pesquisa da modelagem da evapotranspiração de referência a partir de modelos de aprendizagem de máquina, uma vez que tais resultados vão de encontro com as pesquisas de Antonopoulos e Antonopoulos (2017), Traore, Luo e Fipps (2016), Pandey, Nyori e Pandey (2017) Ferreira et al. (2019a) e Ferreira e Cunha (2020), as quais obtiveram resultados semelhantes, utilizando-se de modelos de MLP ou SVM treinados com os parâmetros

climáticos relacionados a temperatura e radiação solar, porém em outras regiões climáticas, e com dados obtidos a partir de estações meteorológicas, mostrando que o modelo MLP que utiliza-se de parâmetros relacionados com a temperatura e a radiação obtidos por satélites pode fornecer resultados melhores independente da região climática.

A Aplicação dos modelos de aprendizagem de máquina no local de teste do experimento de Brotas apresentou as piores métricas, indicando que o modelo treinado no local de treinamento selecionado não pode ser replicado para o local de teste, como ilustrado na Figura 36.

Figura 37 – Métricas de todos os experimentos de MATOPIBA



Fonte: Autor, 2020

Uma explicação para tal resultado advém da baixa relação entre a ET_o dos locais de treinamento e teste. Esta baixa relação pode ser avaliada a partir da classificação climática dos locais estudados, onde ambos pertencem a mesma classe climática, porém diferem na sub-classificação, o que indica que um modelo treinado em um local de classificação temperada de estação seca e verão morno pode não ser adequado a um local de classificação de clima temperado de inverno seco e e verão quente.

No experimento de MATOPIBA houve uma melhor relação entre a ET_o do local de treinamento e o local de teste como pode ser observado na Figura 37. Esta melhor relação pode ser associada ao fato de ambos os locais pertencerem a uma mesma classificação climática.

Não obstante a aplicação dos modelos treinados no local de treinamento e aplicados no local de teste obtiveram métricas melhores, quando comparados com o experimento de Brotas,

indicando que neste local os modelos treinados no local de treinamento pode ser estendido para os locais de teste.

7 Conclusão

A presente pesquisa propôs o uso de modelos de aprendizagem de máquina Máquina Vetor de Suporte e Perceptron de Múltiplas Camadas para modelar a evapotranspiração de referência a partir de parâmetros climáticos medidos por satélites. A pesquisa elabora uma nova abordagem para o uso de parâmetros climáticos recuperados da base de dados NASA/POWER, assim como, contribuí para a literatura de uso de modelos de aprendizagem de máquina para a modelagem da evapotranspiração de referência.

A metodologia empregada evidência que ainda há parâmetros climáticos obtidos a partir de satélites que necessitam suas técnicas de modelagem sejam melhoradas em certas regiões climáticas, mas, seu potencial de uso é evidente. Contudo, uma solução para a melhoria dos resultados seria utilizar os parâmetros mais precisos de cada fonte de dados e combina-los nos modelos estudados.

Os resultados obtidos também evidenciam o potencial de uso dos modelos de aprendizagem de máquina estudados para se obter a medida de evapotranspiração de referência, utilizando-se de parâmetros climáticos medidos a partir de satélites nos locais de treinamento.

Os melhores resultados foram obtidos pelo modelo MLP em todos os testes realizados, sendo que no experimento de Brotas e MATOPIBA este modelo mostrou uma acurácia excelente, quando treinado com todos os parâmetros climáticos obtidos por estações meteorológicas. Já os testes realizados com os parâmetros climáticos obtidos pela base de dados NASA/POWER há uma tendência de se subestimar a evapotranspiração de referência em todos os experimentos.

A aplicação dos modelos estudados nos locais de teste mostram que os resultados do experimento de Brotas são insatisfatórios, uma vez que os modelos não foram capazes de modelar a evapotranspiração do local.

Já os resultados obtidos na aplicação dos modelos no experimento de MATOPIBA são promissores, uma vez que o padrão de modelos aplicados com o conjunto $C1$ com dados obtidos por estações meteorológicas serem os mais precisos e os modelos aplicados com o conjunto $C3$ com dados de satélite superarem os modelos que se utilizam do mesmo conjunto de dados porém obtidos por estações meteorológicas.

Como pesquisa futura novos modelos de aprendizagem de máquina podem ser avaliados, bem como a extensão desta pesquisa em outras regiões climáticas, onde tais modelos podem ser treinados com diferentes escalas temporais de parâmetros climáticos, além de outras bases de dados de satélites para a aquisição destes mesmos parâmetros poderem ser avaliadas.

Outras pesquisas também podem utilizar os mesmos modelos, porém com parâmetros obtidos por satélites, veículos aéreos não tripulados, estações meteorológicas e sondas de solo, com o intuito de avaliar diferentes escalas espaciais, o que poderia melhorar os resultados do experimento de Brotas.

Outra possibilidade de trabalhos futuros, a partir dos resultados obtidos no experimento de MATOPIBA, diz respeito a aplicação dos modelos em um talhão experimental com o intuito de se avaliar o impacto das medidas de evapotranspiração de referência obtidas a partir dos modelos de aprendizagem de máquina para se obter a recomendação de irrigação de uma cultura.

REFERÊNCIAS

- ABTEW, Wossenu. Evapotranspiration measurements and modeling for three wetland systems in South Florida 1. **JAWRA Journal of the American Water Resources Association**, Wiley Online Library, v. 32, n. 3, p. 465–473, 1996.
- ADELOYE, Adebayo J; RUSTUM, Rabee; KARIYAMA, Ibrahim D. Kohonen self-organizing map estimator for the reference crop evapotranspiration. **Water Resources Research**, Wiley Online Library, v. 47, n. 8, 2011.
- ADNAN, Muhammad et al. Estimating evapotranspiration using machine learning techniques. **International Journal of Advanced Computer Science and Applications**, v. 8, n. 9, p. 108–113, 2017.
- ADNAN, Rana Muhammad et al. Reference Evapotranspiration Modeling Using New Heuristic Methods. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 22, n. 5, p. 547, 2020.
- AGARWAL, Abhishek et al. Efficient hierarchical-PCA dimension reduction for hyperspectral imagery. In: IEEE. 2007 IEEE International Symposium on Signal Processing and Information Technology. [S.l.: s.n.], 2007. P. 353–356.
- ALLEN, Richard G et al. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. **Fao, Rome**, v. 300, n. 9, p. d05109, 1998.
- ALPAYDIN, Ethem. **Introduction to machine learning**. [S.l.]: MIT press, 2020.
- ALVARES, Clayton Alcarde et al. Köppen’s climate classification map for Brazil. **Meteorologische Zeitschrift**, E. Schweizerbart’sche Verlagsbuchhandlung, v. 22, n. 6, p. 711–728, 2013.
- ANA. **Atlas irrigação: Uso da água na agricultura irrigada**. [S.l.]: ANA Brasília, 2017.
- ANA. **Conjuntura dos recursos hídricos no Brasil 2019: informe anual / Agência Nacional de Águas**. Brasília, 2019. P. 100.
- ANTONOPOULOS, Vassilis Z; ANTONOPOULOS, Athanasios V. Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. **Computers and Electronics in Agriculture**, Elsevier, v. 132, p. 86–96, 2017.
- ANTONY, Anish Paul et al. A Review of Practice and Implementation of the Internet of Things (IoT) for Smallholder Agriculture. **Sustainability**, Multidisciplinary Digital Publishing Institute, v. 12, n. 9, p. 3750, 2020.
- ATZORI, Luigi; IERA, Antonio; MORABITO, Giacomo. The internet of things: A survey. **Computer networks**, Elsevier, v. 54, n. 15, p. 2787–2805, 2010.

ATZORI, Luigi; IERA, Antonio; MORABITO, Giacomo. Understanding the Internet of Things: definition, potentials, and societal role of a fast evolving paradigm. **Ad Hoc Networks**, Elsevier, v. 56, p. 122–140, 2017.

BARROS, José Roberto Mendonça de; BARROS, Alexandre Lahoz Mendonça de; CYPRIANO, Marcelo Petersen. O mercado da citricultura no Brasil e as suas novas perspectivas, 2016.

BEZERRA, Sabrina. **Reservoir Computing com Hierarquia para Previsão de Vazões Médias Diárias**. Jul. 2016. Tese (Doutorado).

BLANEY, Harry French et al. Determining water requirements in irrigated areas from climatological and irrigation data, 1952.

BREIMAN, Leo. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001.

BROOMHEAD, David S; LOWE, David. **Radial basis functions, multi-variable functional interpolation and adaptive networks**. [S.l.], 1988.

CARDIM FERREIRA LIMA, Matheus et al. Automatic Detection and Monitoring of Insect Pests—A Review. **Agriculture**, Multidisciplinary Digital Publishing Institute, v. 10, n. 5, p. 161, 2020.

CHEN, T et al. guesstrin, C. Xgboost: A scalable tree boosting system. In: PROCEEDINGS of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (New York, NY, USA, 2016), KDD '16, ACM. [S.l.: s.n.], 2016. P. 785–794.

CITRICULTURA, Fundo de Defesa da. Inventário de árvores e estimativa da safra de laranja no cinturão citrícola de São Paulo e Triângulo/Sudoeste mineiro—retrato dos pomares em março/2016. **Araraquara: Fundecitrus**, 2016.

DAKER, Alberto. **A água na agricultura: irrigação e drenagem**. [S.l.]: F. Bastos, SD, 1973.

DJAMAN, Koffi et al. Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley. **Journal of Hydrology: regional studies**, Elsevier, v. 3, p. 139–159, 2015.

DOORENBOS, J; PRUITT, WO. Crop water requirements. FAO irrigation and drainage paper 24. **Land and Water Development Division, FAO, Rome**, v. 144, 1977.

EHLERS, Eduardo. **O que é agricultura sustentável**. [S.l.]: Brasiliense, 2009.

EMBRAPA. **Sobre o Matopiba**. [S.l.: s.n.], 2017. Disponível em: <<https://www.embrapa.br/tema-matopiba/sobre-o-tema>>. Acesso em: 24 out. 2020.

FAO. **The future of food and agriculture - Alternative pathways to 2050. Summary version**. Rome, 2018. P. 60. ISBN 9789251309896. Disponível em: <<http://www.fao.org/3/CA1553EN/ca1553en.pdf>>.

FENG, Yu et al. Modeling reference evapotranspiration using extreme learning machine and generalized regression neural network only with temperature data. **Computers and Electronics in Agriculture**, Elsevier, v. 136, p. 71–78, 2017.

FERREIRA, Candida. Gene expression programming: a new adaptive algorithm for solving problems. **arXiv preprint cs/0102027**, 2001.

FERREIRA, Lucas Borges; CUNHA, Fernando França da. New approach to estimate daily reference evapotranspiration based on hourly temperature and relative humidity using machine learning and deep learning. **Agricultural Water Management**, Elsevier, v. 234, p. 106113, 2020.

FERREIRA, Lucas Borges et al. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM—A new approach. **Journal of Hydrology**, Elsevier, v. 572, p. 556–570, 2019.

FERREIRA, Lucas Borges et al. Multivariate adaptive regression splines (MARS) applied to daily reference evapotranspiration modeling with limited weather data. **Acta Scientiarum. Agronomy**, SciELO Brasil, v. 41, 2019.

FRIEDMAN, Jerome H. Greedy function approximation: a gradient boosting machine. **Annals of statistics**, JSTOR, p. 1189–1232, 2001.

FRIEDMAN, Jerome H. Multivariate adaptive regression splines. **The annals of statistics**, JSTOR, p. 1–67, 1991.

GARCIA BENEVIDES, J; LOPEZ, DIAZ J. Formula para el calculo de La evapotranpiracion potencial adaptada al tropico (15oN a 15oS). **Agronomia Tropical, Maracay**, v. 20, n. 5, p. 335–340, 1970.

GIUSTOLISI, Orazio; SAVIC, Dragan A. A symbolic data-driven technique based on evolutionary polynomial regression. **Journal of Hydroinformatics**, IWA Publishing, v. 8, n. 3, p. 207–222, 2006.

GOCIC, Milan et al. Comparative analysis of reference evapotranspiration equations modelling by extreme learning machine. **Computers and Electronics in Agriculture**, Elsevier, v. 127, p. 56–63, 2016.

GOCIĆ, Milan et al. Soft computing approaches for forecasting reference evapotranspiration. **Computers and Electronics in Agriculture**, Elsevier, v. 113, p. 164–173, 2015.

GOODFELLOW, Ian et al. **Deep learning**. [S.l.]: MIT press Cambridge, 2016. v. 1.

GOYAL, Manish Kumar et al. Modeling of daily pan evaporation in sub tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS. **Expert systems with applications**, Elsevier, v. 41, n. 11, p. 5267–5276, 2014.

GRANATA, Francesco. Evapotranspiration evaluation models based on machine learning algorithms—A comparative study. **Agricultural Water Management**, Elsevier, v. 217, p. 303–315, 2019.

HAMON, W Russell. **Estimating potential evapotranspiration**. 1960. Tese (Doutorado) – Massachusetts Institute of Technology.

HARGREAVES, George H; ALLEN, Richard G. History and evaluation of Hargreaves evapotranspiration equation. **Journal of irrigation and drainage engineering**, American Society of Civil Engineers, v. 129, n. 1, p. 53–63, 2003.

HARGREAVES, George H; SAMANI, Zohrab A. Reference crop evapotranspiration from temperature. **Applied engineering in agriculture**, American Society of Agricultural e Biological Engineers, v. 1, n. 2, p. 96–99, 1985.

HAYKIN, Simon. Neural Networks and Learning Machine. In: [s.l.: s.n.], jan. 2008.

HEARST, Marti A. et al. Support vector machines. **IEEE Intelligent Systems and their applications**, IEEE, v. 13, n. 4, p. 18–28, 1998.

HEDDAM, Salim et al. Evolving connectionist systems (ECoSs): a new approach for modeling daily reference evapotranspiration (ET₀). **Environmental monitoring and assessment**, Springer, v. 190, n. 9, p. 516, 2018.

HORNIK, Kurt; STINCHCOMBE, Maxwell; WHITE, Halbert. Multilayer feedforward networks are universal approximators. **Neural networks**, Elsevier, v. 2, n. 5, p. 359–366, 1989.

HU, Qing-Fang et al. Global calibration of Hargreaves equation and its applicability in China. **Advances in Water Science**, China International Book Trading Corporation, P. O. Box 399 Beijing 100044 . . . , v. 22, n. 2, p. 160–167, 2011.

HUANG, Guang-Bin; ZHU, Qin-Yu; SIEW, Chee-Kheong. Extreme learning machine: a new learning scheme of feedforward neural networks. In: IEEE. 2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541). [S.l.: s.n.], 2004. v. 2, p. 985–990.

INITIATIVE, IEEE Internet et al. Towards a definition of the Internet of Things (IoT). **Revision-1, on-line: http://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf**. Accessed, v. 27, n. 2017, p. 479–501, 2015.

IRMAK, S et al. Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. **Journal of irrigation and drainage engineering**, American Society of Civil Engineers, v. 129, n. 5, p. 336–347, 2003.

ISPA. **PRECISION AG DEFINITION**. [S.l.: s.n.], 2020. Disponível em: <<https://www.ispag.org/about/definition>>. Acesso em: 24 nov. 2020.

IVAKHNENKO, AG; LAPA, VG. **Cybernetic predicting devices**. Kiev. [S.l.]: Naukova dumka, 1965.

JAMALUDDIN, Hishamuddin et al. Optimum grouping in a modified genetic algorithm for discrete-time, non-linear system identification. **Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering**, SAGE Publications Sage UK: London, England, v. 221, n. 7, p. 975–989, 2007.

JANG, J-SR. ANFIS: adaptive-network-based fuzzy inference system. **IEEE transactions on systems, man, and cybernetics**, IEEE, v. 23, n. 3, p. 665–685, 1993.

JENSEN, Marvin E; HAISE, Howard Ross. Estimating evapotranspiration from solar radiation. **Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division**, v. 89, p. 15–41, 1963.

JONES, CA. Crop growth models. **Management of Farm Irrigated Systems**, American Society of Agricultural Engineering, 1990.

JÚNIOR, Walter Maldonado; VALERIANO, Taynara Tuany Borges; SOUZA ROLIM, Glauco de. EVAPO: A smartphone application to estimate potential evapotranspiration using cloud gridded meteorological data from NASA-POWER system. **Computers and electronics in agriculture**, Elsevier, v. 156, p. 187–192, 2019.

KAMIENSKI, Carlos et al. Smart water management platform: Iot-based precision irrigation for agriculture. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 2, p. 276, 2019.

KASABOV, Nikola. Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 31, n. 6, p. 902–918, 2001.

KASABOV, Nikola K; SONG, Qun. DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. **IEEE transactions on Fuzzy Systems**, IEEE, v. 10, n. 2, p. 144–154, 2002.

KHAND, Kul et al. A modeling framework for deriving daily time series of evapotranspiration maps using a surface energy balance model. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 11, n. 5, p. 508, 2019.

KINGMA, Diederik P; BA, Jimmy. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014.

KISI, O; DEMIR, V. Evapotranspiration estimation using six different multi-layer perceptron algorithms. **Irrigation and Drainage Systems Engineering**, v. 5, n. 2, p. 1–6, 2016.

KISI, Ozgur. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. **Journal of Hydrology**, Elsevier, v. 528, p. 312–320, 2015.

- KIŞI, Özgür; ÖZTÜRK, Özgür. Adaptive neurofuzzy computing technique for evapotranspiration estimation. **Journal of Irrigation and Drainage Engineering**, American Society of Civil Engineers, v. 133, n. 4, p. 368–379, 2007.
- KITCHENHAM, Barbara. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. 2004, p. 1–26, 2004.
- KOHONEN, Teuvo et al. Engineering applications of the self-organizing map. **Proceedings of the IEEE**, IEEE, v. 84, n. 10, p. 1358–1384, 1996.
- KONCAR, N. **Optimisation methodologies for direct inverse neurocontrol**. 1997. Tese (Doutorado) – University of London.
- KÖPPEN, Wladimir; GEIGER, Rudolf. **Handbuch der klimatologie**. [S.l.]: Gebrüder Borntraeger Berlin, 1930. v. 1.
- LIKAS, Aristidis; VLASSIS, Nikos; VERBEEK, Jakob J. The global k-means clustering algorithm. **Pattern recognition**, Elsevier, v. 36, n. 2, p. 451–461, 2003.
- LIU, Tianfeng et al. Extreme learning machine based on particle swarm optimization for estimation of reference evapotranspiration. In: IEEE. 2017 36th Chinese Control Conference (CCC). [S.l.: s.n.], 2017. P. 4567–4572.
- MAKKINK, G F. Testing the Penman formula by means of lysimeters. **Journal of the Institution of Water Engineers**, v. 11, p. 277–288, 1957.
- MCCULLAGH, Peter. **Generalized linear models**. Routledge, 2018.
- MCCULLOCH, Warren S; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.
- MEHDIZADEH, Saeid; BEHMANESH, Javad; KHALILI, Keivan. Using MARS, SVM, GEP and empirical equations for estimation of monthly mean reference evapotranspiration. **Computers and electronics in agriculture**, Elsevier, v. 139, p. 103–114, 2017.
- MITCHELL, Tom M et al. **Machine learning**. [S.l.]: McGraw-hill New York, 1997.
- MOLIN, José Paulo; AMARAL, Lucas Rios do; COLAÇO, André. **Agricultura de precisão**. [S.l.]: Oficina de Textos, 2015.
- MONTEIRO, Leonardo A; SENTELHAS, Paulo C; PEDRA, George U. Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. **International Journal of Climatology**, Wiley Online Library, v. 38, n. 3, p. 1571–1581, 2018.
- MONTEITH, John L. Evaporation and environment. In: CAMBRIDGE UNIVERSITY PRESS (CUP) CAMBRIDGE. SYMPOSIA of the society for experimental biology. [S.l.: s.n.], 1965. v. 19, p. 205–234.

MONTELEONE, Sergio et al. Exploring the Adoption of Precision Agriculture for Irrigation in the Context of Agriculture 4.0: The Key Role of Internet of Things. **Sensors**, Multidisciplinary Digital Publishing Institute, p. 1–32, 2020.

NOURANI, Vahid; ELKIRAN, Gozen; ABDULLAHI, Jazuli. Multi-station artificial intelligence based ensemble modeling of reference evapotranspiration using pan evaporation measurements. **Journal of Hydrology**, Elsevier, v. 577, p. 123958, 2019.

LOUDIN, Ludovic et al. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. **Journal of hydrology**, Elsevier, v. 303, n. 1-4, p. 290–306, 2005.

ÖZBAYOĞLU, Gülhan; ÖZBAYOĞLU, M Evren. A new approach for the prediction of ash fusion temperatures: a case study using Turkish lignites. **Fuel**, Elsevier, v. 85, n. 4, p. 545–552, 2006.

PANDEY, PK; NYORI, Topi; PANDEY, Vanita. Estimation of reference evapotranspiration using data driven techniques under limited data conditions. **Modeling Earth Systems and Environment**, Springer, v. 3, n. 4, p. 1449–1461, 2017.

PEEL, Murray C; FINLAYSON, Brian L; MCMAHON, Thomas A. Updated world map of the Köppen-Geiger climate classification, 2007.

PENMAN, Howard Latimer. Natural evaporation from open water, bare soil and grass. **Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences**, The Royal Society London, v. 193, n. 1032, p. 120–145, 1948.

PEREIRA, Luis S et al. Crop evapotranspiration estimation with FAO56: Past and future. **Agricultural Water Management**, Elsevier, v. 147, p. 4–20, 2015.

PERERA, Kushan C et al. Forecasting daily reference evapotranspiration for Australia using numerical weather prediction outputs. **Agricultural and forest meteorology**, Elsevier, v. 194, p. 50–63, 2014.

PRIESTLEY, Charles Henry Brian; TAYLOR, Robert Joseph. On the assessment of surface heat flux and evaporation using large-scale parameters. **Monthly weather review**, v. 100, n. 2, p. 81–92, 1972.

QUINLAN, John R et al. Learning with continuous classes. In: WORLD SCIENTIFIC. 5TH Australian joint conference on artificial intelligence. [S.l.: s.n.], 1992. v. 92, p. 343–348.

REIS, Matheus Mendes et al. Empirical and learning machine approaches to estimating reference evapotranspiration based on temperature data. **Computers and Electronics in Agriculture**, Elsevier, v. 165, p. 104937, 2019.

ROSA, João Luís Garcia. Fundamentos da inteligência artificial. **Rio de Janeiro: LTC**, v. 1, 2011.

ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, American Psychological Association, v. 65, n. 6, p. 386, 1958.

RUSSELL, Stuart; NORVIG, Peter. Artificial intelligence: a modern approach, 2002.

SAGGI, Mandeep Kaur; JAIN, Sushma. Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. **Computers and electronics in agriculture**, Elsevier, v. 156, p. 387–398, 2019.

SAHRI, Zahriah; YUSOF, Rubiyah; WATADA, Junzo. FINNIM: Iterative imputation of missing values in dissolved gas analysis dataset. **IEEE Transactions on Industrial Informatics**, IEEE, v. 10, n. 4, p. 2093–2102, 2014.

SAMSURI, Saiful Farhan Mohd; AHMAD, Robiah; ZAKARIA, Mohd Zakimi. Reference Evapotranspiration Estimation in Tropical Region using Penman-Monteith Equation and Evolutionary Computation. In: IEEE. 2018 2nd International Conference on Smart Sensors and Application (ICSSA). [S.l.: s.n.], 2018. P. 152–157.

SANIKHANI, Hadi et al. Temperature-based modeling of reference evapotranspiration using several artificial intelligence models: application of different modeling scenarios. **Theoretical and Applied Climatology**, Springer, v. 135, n. 1-2, p. 449–462, 2019.

SANTOS, Eulanda Miranda dos et al. Teoria e aplicação de Support Vector Machines à aprendizagem e reconhecimento de objetos baseado na aparência. Universidade Federal de Campina Grande, 2002.

SAVVA, Andreas P; FRENKEN, Karen. **Crop water requirements and irrigation scheduling**. [S.l.]: FAO Sub-Regional Office for East e Southern Africa Harare, 2002.

SCHERER-WARREN, Morris; RODRIGUES, Lineu Neiva. Estimativa de Evapotranspiração Real por Sensoriamento Remoto: procedimento e aplicação em pivô central. **EMBRAPA Cerrados, Planaltina, DF**, 2013.

SHAFI, Uferah et al. Precision Agriculture Techniques and Practices: From Considerations to Applications. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 17, p. 3796, 2019.

SHAFI, Uferah et al. Precision agriculture techniques and practices: From considerations to applications. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 17, p. 3796, 2019.

SHIRI, Jalal. Evaluation of FAO56-PM, empirical, semi-empirical and gene expression programming approaches for estimating daily reference evapotranspiration in hyper-arid regions of Iran. **Agricultural water management**, Elsevier, v. 188, p. 101–114, 2017.

SINGH, VP; XU, C-Y. Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation. **Hydrological Processes**, Wiley Online Library, v. 11, n. 3, p. 311–323, 1997.

SPECHT, Donald F et al. A general regression neural network. **IEEE transactions on neural networks**, Citeseer, v. 2, n. 6, p. 568–576, 1991.

SPECHT, Donald F. Probabilistic neural networks. **Neural networks**, Elsevier, v. 3, n. 1, p. 109–118, 1990.

SRIVASTAVA, Prashant K et al. Performance assessment of evapotranspiration estimated from different data sources over agricultural landscape in Northern India. **Theoretical and Applied Climatology**, Springer, p. 1–12, 2020.

STACKHOUSE, PW et al. Surface meteorology and Solar Energy (SSE) Release 6.0 Methodology Version 3.2. 0 June 2. **Atmospheric Science Data Center**. **042d043b0435043a04420440043e043d043d044b~0438 044004350441044304400441: https://eosweb.larc.nasa.gov/sse/documents/SSE6Methodology.pdf (0414043004420430 043e04310440043004490435043d0438044f: 15.07. 2016)**, 2015.

STEINWART, Ingo; CHRISTMANN, Andreas. **Support vector machines**. [S.l.]: Springer Science & Business Media, 2008.

SUBEDI, Abhinaya; CHÁVEZ, José L. Crop evapotranspiration (ET) estimation models: a review and discussion of the applicability and limitations of ET methods. **Journal of Agricultural Science**, Canadian Center of Science e Education, v. 7, n. 6, p. 50, 2015.

SUYKENS, Johan AK; VANDEWALLE, Joos. Least squares support vector machine classifiers. **Neural processing letters**, Springer, v. 9, n. 3, p. 293–300, 1999.

TANGUNE, Bartolomeu Félix; ESCOBEDO, João Francisco. Reference evapotranspiration in São Paulo State: Empirical methods and machine learning techniques. **Int. J. Water Res. Environ. Eng.**, v. 10, p. 33–44, 2018.

TEZEL, Gulay; BUYUKYILDIZ, Meral. Monthly evaporation forecasting using artificial neural networks and support vector machines. **Theoretical and applied climatology**, Springer, v. 124, n. 1-2, p. 69–80, 2016.

THAYANANTHAN, Arasanathan et al. Pose estimation and tracking using multivariate regression. **Pattern Recognition Letters**, Elsevier, v. 29, n. 9, p. 1302–1310, 2008.

TORRES, Alfonso F; WALKER, Wynn R; MCKEE, Mac. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. **Agricultural Water Management**, Elsevier, v. 98, n. 4, p. 553–562, 2011.

TRAORE, Seydou; LUO, Yufeng; FIPPS, Guy. Deployment of artificial neural network for short-term forecasting of evapotranspiration using public weather forecast restricted messages. **Agricultural Water Management**, Elsevier, v. 163, p. 363–379, 2016.

TURC, L. Evaluation des besoins en eau d'irrigation, évapotranspiration potentielle. **Ann. Agron.**, v. 12, p. 13–49, 1961.

TURING, INTELLIGENCE BY AM. Computing machinery and intelligence-AM Turing. **Mind**, v. 59, n. 236, p. 433, 1950.

VALIANTZAS, John D. Simplified forms for the standardized FAO-56 Penman–Monteith reference evapotranspiration using limited weather data. **Journal of Hydrology**, Elsevier, v. 505, p. 13–23, 2013.

VAPNIK, V. **The nature of statistical learning theory**. New York: Springer, 1995.

WRIGHT, James L; JENSEN, Marvin E. Peak water requirements of crops in southern Idaho. **Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division**, v. 98, IR2, p. 193–201, 1972.

ZADEH, Lotfi A. Fuzzy sets. In: FUZZY sets, fuzzy logic, and fuzzy systems: selected papers by Lotfi A Zadeh. [S.l.]: World Scientific, 1996. P. 394–432.

ZHANG, Yu et al. Maize Crop Coefficient Estimated from UAV-Measured Multispectral Vegetation Indices. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 23, p. 5250, 2019.

ZHANG, Zixiong; GONG, Yicheng; WANG, Zhongjing. Accessible remote sensing data based reference evapotranspiration estimation modelling. **Agricultural Water Management**, Elsevier, v. 210, p. 59–69, 2018.