

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO
EM CIÊNCIA DOS MATERIAIS

THIAGO ALVES SPONTONI

**AVALIAÇÃO E MODELAGEM DA EVAPOTRANSPIRAÇÃO DE
REFERÊNCIA USANDO DIFERENTES TÉCNICAS DE APRENDIZADO DE
MÁQUINA PARA UMA SAVANA TROPICAL BRASILEIRA**

CAMPO GRANDE - MS

2023

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO
EM CIÊNCIA DOS MATERIAIS

THIAGO ALVES SPONTONI

AVALIAÇÃO E MODELAGEM DA EVAPOTRANSPIRAÇÃO DE
REFERÊNCIA USANDO DIFERENTES TÉCNICAS DE APRENDIZADO DE
MÁQUINA PARA UMA SAVANA TROPICAL BRASILEIRA

Orientador: Prof. Dr. Thiago Rangel Rodrigues

Tese apresentada ao Programa de Pós - Graduação em Ciência dos Materiais do Instituto de Física da Universidade Federal de Mato Grosso do Sul INFI/UFMS como parte dos requisitos para obtenção do título de Doutor em Ciência dos Materiais.

Campo Grande - MS

2023

AVALIAÇÃO E MODELAGEM DA EVAPOTRANSPIRAÇÃO DE REFERÊNCIA USANDO DIFERENTES TÉCNICAS DE APRENDIZADO DE MÁQUINA PARA UMA SAVANA TROPICAL BRASILEIRA

THIAGO ALVES SPONTONI

Tese apresentada ao Programa de Pós-Graduação em Ciência dos Materiais do Instituto de Física da Universidade Federal de Mato Grosso do Sul - INFI/UFMS como parte dos requisitos de para obtenção do título de Doutor em Ciência dos Materiais.

Banca Examinadora:

Titulares:

Prof^o. Dr. Thiago Rangel Rodrigues - Universidade Federal de Mato Grosso do Sul - UFMS

Prof^o. Dr. Fábio Mallmann Zimmer - Universidade Federal de Mato Grosso do Sul - UFMS

Prof^o. Dr. Thiago Meirelles Ventura - Universidade Federal de Mato Grosso - UFMT

Prof^o. Dr. Marcelo Sacardi Biudes - Universidade Federal de Mato Grosso - UFMT

Prof^o. Dr. Rafael da Silva Palácios - Universidade Federal do Pará - UFPA

Campo Grande - MS, 06 de setembro de 2023.

Agradecimentos

Agradeço a Deus pela vida, pelo grande amor que tem por mim e por colocar as pessoas certas em meu caminho;

Agradeço à minha base, fonte de toda inspiração, incentivo e alegria, Charlene Marcondes Avelar, àquela que me acompanha e direciona-me em todos os momentos;

Agradeço à minha filha, a pequena Ana Yara, a qual entendeu como ninguém a necessidade do papai em se abster de momentos de diversão para se dedicar aos estudos;

Aos meus pais, Antonio Claudio Spontoni e Celina Alves Spontoni e minha irmã Emory Alves Spontoni e a minha cunhada Marcela Vieira, os quais ajudaram com muitos conselhos, paciência e carinho;

Ao prof^o Dr. Thiago Rangel Rodrigues, que além de ser um ótimo orientador, inspirador deste trabalho, demonstrou muita paciência ao me conduzir pelos corretos caminhos, para que o sucesso fosse alcançado;

Ao prof^o Dr. Thiago Meirelles Ventura, por sua valiosa contribuição prestada nesta pesquisa;

Ao prof^o Dr. Fábio Mallmann Zimmer, por me recepcionar com toda paciência e afeto dentro do Programa de Pós-Graduação em Ciência dos Materiais;

Aos profs^o Dr. Marcelo Sacardi Biudes e Dr. Rafael da Silva Palácios, por dispensarem um tempo valioso na avaliação e recomendações sobre o projeto;

Agradeço à Universidade Federal de Mato Grosso do Sul - UFMS por toda sua contribuição na pesquisa, seja pelo apoio técnico, quanto ao apoio financeiro para a publicação do artigo;

Agradeço ao Instituto de Física - INFI e ao Programa de Pós-Graduação em Ciência dos Materiais - PPGCM por me acolherem como estudante desde o período de graduação;

Agradeço à Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul - Fundect, pela concessão da bolsa de estudos durante o período de realização deste trabalho;

A todos que, direta ou indiretamente, apoiaram a realização deste trabalho.

“Não é preciso saber tudo, mas é fundamental estar cercado por pessoas que sabem!”

Carmem Corrêa Miranda

Resumo

Elementos meteorológicos possuem diferentes influências no ambiente, podendo modificar significativamente o processo natural de desenvolvimento da vegetação e contribuindo assim para as mudanças climáticas. Nesse estudo foi investigado a utilização de variáveis meteorológicas para a determinação da Evapotranspiração de Referência (ET_o) em uma região de transição entre o Cerrado e o Pantanal. O objetivo foi o de aplicar técnicas de aprendizado de máquina para determinar a ET_o com o mínimo de variáveis possível. Os resultados indicam que a aplicação da inteligência artificial pode promover melhorias substanciais na modelagem ambiental, quando técnicas de previsão alternativas são empregadas, resultando em redução dos custos do projeto e obtenção de resultados mais confiáveis. A presente pesquisa buscou identificar a combinação mais eficiente de técnicas de aprendizado de máquina, tais como Redes Neurais Artificiais, *Random Forest* e *Support Vector Machines*. Como resultado foi desenvolvido um novo modelo, o qual depende de um menor número de variáveis climáticas em comparação com o método Penman-Monteith (método padrão para estimar a Evapotranspiração de Referência) e que consegue descrever de forma eficiente a ET_o. As técnicas de aprendizado de máquina se mostraram altamente eficazes nesse tipo de modelagem, devido à sua capacidade de processar grandes volumes de dados e identificar as melhores interações entre os parâmetros envolvidos. Além disso, quando as redes neurais artificiais são empregadas, foi obtida uma acurácia superior a 94% na determinação da ET_o ao se utilizar um número reduzido de variáveis em comparação com o método padrão.

Palavras-chave: Savana Tropical; Mudanças Climáticas; Combinação; Inteligência Artificial; Redes Neurais Artificiais.

Abstract

Meteorological elements have different influences on the environment and can significantly modify the natural vegetation development process, thus contributing to climate change. This study investigated the use of meteorological variables to determine Reference Evapotranspiration (ET_o) in a transition region between the Cerrado and the Pantanal. The aim was to apply machine learning techniques to determine ET_o with as few variables as possible. The results indicate that the application of artificial intelligence can promote substantial improvements in environmental modeling when alternative forecasting techniques are employed, resulting in reduced project costs and more reliable results. This research sought to identify the most efficient combination of machine learning techniques, such as Artificial Neural Networks, Random Forest and Support Vector Machines. As a result, a new model was developed, which depends on a smaller number of climatic variables compared to the Penman-Monteith method (the standard method for estimating reference evapotranspiration) and which manages to efficiently describe ET_o. Machine learning techniques have proved highly effective in this type of modeling, due to their ability to process large volumes of data and identify the best interactions between the parameters involved. Furthermore, when artificial neural networks are used, an accuracy of over 94% was obtained in determining ET_o when using a reduced number of variables compared to the standard method.

Keywords: Tropical Savannah; Climate Change; Combination; Artificial Intelligence; Artificial Neural Networks.

Sumário

1	Introdução	1
2	Fundamentação Teórica	4
2.1	O Processo de Evapotranspiração	4
2.2	Evaporação	4
2.3	Transpiração	5
2.4	Evapotranspiração (ET)	6
2.5	Evapotranspiração Potencial (ETp)	7
2.6	Evapotranspiração Real (ETr)	7
2.7	Evapotranspiração da Cultura (ETc)	7
2.8	Evapotranspiração Real da Cultura (ETRc)	8
2.9	Taxa de Evapotranspiração	8
2.10	Taxa de Evaporação (ET)	8
2.11	Coeficiente de Cultura (Kc)	8
2.12	Evapotranspiração de Referência (ETo)	8
2.12.1	Método de Hargreaves-Samani	9
2.12.2	Método de Thornthwaite	9
2.12.3	Método de Penman-Monteith (FAO-56)	10
2.13	Modelagem Matemática	12
2.14	Inteligência Artificial	13
2.14.1	A Inteligência Artificial Aplicada a Estudos Ambientais	14
2.15	<i>Machine Learning</i>	15
2.15.1	O <i>Machine Learning</i> na Análise de Sistemas Ambientais	15
2.15.2	Os Métodos de Aprendizagem de Máquina	16
2.15.3	Redes Neurais Artificiais (RNA)	17

2.15.4	<i>Random Forest (RF)</i>	18
2.15.5	<i>Support Vector Machine</i>	20
3	Material e Métodos	22
3.1	Área de Estudo	22
3.2	Medições Micrometeorológicas	23
3.3	Técnicas Utilizadas	23
3.4	Validação dos Modelos	25
3.4.1	R - Squared (R^2)	25
3.4.2	Root Mean Square Error (RMSE)	26
3.4.3	Mean Absolute Error (MAE)	26
4	Resultados e Discussões	27
5	Conclusões	38

Lista de Figuras

2.1	Representação esquemática de uma folha em corte, evidenciando sua estrutura interna.	5
2.2	Modelo de Regressão e Classificação.	17
2.3	Diagrama de funcionamento de uma Rede Neural Artificial.	18
2.4	Diagrama de funcionamento do algoritmo <i>Random Forest</i>	19
2.5	Esquema da função de transformação <i>Support Vector Machine</i>	20
2.6	Esquema do método <i>Support Vector Machine</i>	20
3.1	Localização da área de estudo.	23
4.1	Evolução temporal da Temperatura do Ar ($^{\circ}\text{C}$) durante 10 anos. Em destaque a Média Móvel de 15 dias.	27
4.2	Evolução temporal da Radiação Solar Global ($\text{MJ}/\text{m}^2.\text{d}$) durante 10 anos. Em destaque a Média Móvel de 15 dias.	28
4.3	Evolução temporal da Umidade Relativa do Ar (%) durante 10 anos. Em destaque a Média Móvel de 15 dias.	29
4.4	Evolução temporal da Velocidade do Vento (m/s) durante 10 anos. Em destaque a Média Móvel de 15 dias.	30
4.5	Comportamento do vento sob a influência das características do terreno.	31
4.6	Evolução temporal da Evapotranspiração de Referência (ET _o) durante 10 anos. Em destaque a Média Móvel de 15 dias.	31

4.7 Correlação entre os parâmetros com base na técnica utilizada. **(a)**, parâmetros GSR + T + WS, $y = 0,9356x + 0,1726$, $R^2 = 0,9450$; **(b)**, parâmetros GSR + RH + WS, $y = 0,9059x + 0,2636$, $R^2 = 0,9405$; **(c)**, parâmetros GSR + T + RH, $y = 0,8799x + 0,4695$, $R^2 = 0,9320$; **(d)**, parâmetros GSR + T + WS, $y = 0,9263x - 0,0085$, $R^2 = 0,8794$; **(e)**, parâmetros GSR + RH + WS, $y = 0,9004x + 0,0548$, $R^2 = 0,8642$; **(f)**, parâmetros GSR + T + RH, $y = 0,8761x + 0,1715$, $R^2 = 0,8471$; **(g)**, parâmetros GSR + T + WS, $y = 0,8987x + 0,0457$, $R^2 = 0,8261$; **(h)**, parâmetros GSR + RH + WS, $y = 0,8872x + 0,0755$, $R^2 = 0,8190$; **(i)**, parâmetros GSR + T + RH, $y = 0,8780x + 0,1816$, $R^2 = 0,8570$. . 36

Lista de Tabelas

4.1 Desempenho das técnicas utilizadas.	32
---	----

Lista de Abreviaturas e Siglas

ET Evapotranspiração

ET_o Evapotranspiração de Referência

ET_p Evapotranspiração Potencial

ET_r Evapotranspiração Real

ET_c Evapotranspiração da Cultura

ET_{Rc} Evapotranspiração Real da Cultura

K_c Coeficiente de Cultura

FAO Organização das Nações Unidas para Agricultura e Alimentação

RH Umidade Relativa do Ar (%)

IA Inteligência Artificial

ML *Machine Learning*

PM-FAO Penman-Monteith (FAO)

RF *Random Forest*

RBF *Radial Basis Function*

RNA Redes Neurais Artificiais

GSR Radiação Solar Global ($MJ/m^2.d$)

SVM *Support Vector Machine*

AM Aprendizado de Máquina

T Temperatura do Ar ($^{\circ}\text{C}$)

WS Velocidade do Vento (m/s)

Capítulo 1

Introdução

Devido a inúmeras e descontroladas, na maioria das vezes, intervenções que o homem vem praticando contra a natureza em todo planeta, as mudanças climáticas estão sendo cada vez mais percebidas por toda sociedade. Alterações em biomas naturais, cuja concentração de recursos e materiais de alto valor agregado é abundante, são praticadas através da retirada desenfreada de matérias, sem quaisquer preocupações futuras. Em nosso país, identifica-se que o desmatamento em regiões de grande potencial ecológico tem sido realizado muitas vezes por latifundiários e empresas do ramo da mineração, sendo esta prática intensificada devido ao alto valor comercial dos insumos extraídos [1].

Conseqüentemente, alterações na fauna e flora estão acontecendo e faz-se necessário que exista um equilíbrio entre o meio ambiente e a sociedade, para que de tal forma ocorra um desenvolvimento sustentável, visando o maior proveito de recursos e minimizando ao máximo sua extração, fortalecendo assim um modelo de sustentabilidade. Um fator importante a se destacar é de que a manutenção destes biomas está ligada ao equilíbrio climático e pode estar diretamente conectada ao comportamento do crescimento da vegetação nativa [2].

O Brasil possui uma posição privilegiada mundialmente, em relação à disponibilidade de recursos hídricos, visto que a vazão média anual dos rios em nosso território é em torno de 180 mil m^3/s , o que corresponde a aproximadamente 12% da disponibilidade de recursos hídricos do planeta [3]. Caso sejam computadas as vazões provenientes de territórios estrangeiros e que adentram o país (Região Amazônica: 86.321 mil m^3/s ; Uruguai: 878 m^3/s e Paraguai: 595 m^3/s), este índice de vazão média total chega na ordem de 267 mil m^3/s (18% da disponibilidade mundial) [3]. Sabe-se que a disponibilidade de água é dependente do clima, sendo influenciada diretamente pelo ciclo anual das chuvas, pela variação da temperatura e por diversos outros

fatores climáticos.

A região pantaneira sofre influência direta de três importantes biomas brasileiros: Amazônia, Cerrado e Mata Atlântica, além do bioma Chaco (nome dado ao Pantanal localizado no norte do Paraguai e leste da Bolívia). Apesar de sua beleza natural exuberante o local de transição do Cerrado vem sendo muito impactado pela ação humana, principalmente pela atividade agropecuária, especialmente nas áreas de planalto adjacente. Preocupações futuras relacionadas ao clima e suas possíveis mudanças, são temas de constantes debates na sociedade atual, pois existe uma grande preocupação com a frequência e a intensidade de eventos climáticos extremos de curta duração.

Assim, por poderem causar um grande impacto na população, eventos como ondas de calor extremas, precipitações intensas, enchentes, secas, entre outros, são de muito interesse entre pesquisadores, pois informações deste tipo podem evitar um grande impacto na sociedade, o que ocasionaria altos custos monetários e, em muitos casos, perdas econômicas e até humanas [4]. Diversos são os estudos e evidências científicas indicando que estas mudanças são um grande risco para os recursos hídricos do planeta [5].

Aliado a isso, temos ainda as variações climáticas que juntas afetam a economia nacional, bem como outros fatores, tais como àqueles relacionados a aspectos políticos e sociais, gerando um elevado custo para toda sociedade em diferentes ordens. Essa vulnerabilidade quando projetada para o futuro, leva a incertezas sobre a demanda de água, mas também pode ser utilizada para a definição políticas ambientais de uso e gerenciamento desse precioso recurso natural [6, 7].

Desempenhando um papel fundamental nos estudos relacionados ao ciclo da água e na gestão dos recursos hídricos, a Evapotranspiração (ET) é uma variável meteorológica crucial que representa a perda potencial de água da superfície do solo devido à evaporação e à transpiração das plantas sob condições ideais. Entender tal parâmetro é importante por várias razões significativas [8].

Esse indicador serve como uma medida padrão que ajuda a compreender e quantificar a demanda hídrica das culturas agrícolas. Ajuda a determinar as necessidades de irrigação em diferentes regiões e pode minimizar o uso da água na agricultura, o que é especialmente relevante em um contexto de crescente escassez de recursos hídricos [8].

Além disso, a Evapotranspiração (ET) desempenha um papel vital na avaliação do balanço hídrico de uma área específica. É um componente essencial em modelos hidrológicos e climáticos

que ajudam a prever a disponibilidade de água em bacias hidrográficas, o que é crítico para o gerenciamento sustentável dos recursos hídricos [9].

Estudos relacionados à conservação e ao uso eficiente da água, indicam que compreender a Evapotranspiração de Referência (ET_o) (quantidade de água que é perdida da superfície de uma cultura vegetal hipotética, conhecida como “grama padrão”, em condições de referência. A ET_o é usada como uma medida padrão para estimar a Evapotranspiração Real (ET_r) de culturas específicas em diferentes locais e condições) é essencial para o planejamento urbano e para garantir o uso responsável da água [10]. Seu entendimento fornece as bases necessárias para uma gestão eficaz dos recursos hídricos, contribuindo para a segurança alimentar, conservação dos recursos naturais e desenvolvimento sustentável [9].

Desta forma, considerando o desenvolvimento e aperfeiçoamento dos métodos de predição [11, 12, 13, 14], propõe-se neste trabalho aplicar diferentes técnicas de *Machine Learning*, tais como as Redes Neurais Artificiais (RNA), *Random Forest* (RF) e *Support Vector Machine* (SVM), para a obtenção de um modelo matemático analítico capaz de caracterizar a Evapotranspiração de Referência (ET_o) de uma região de transição do Cerrado, o qual possua uma menor dependência de variáveis climáticas quando comparado com o método padrão de Penman-Monteith (FAO) (PM-FAO).

Também objetivou-se avaliar as técnicas Redes Neurais Artificiais, *Random Forest* e *Support Vector Machine* quanto as suas *performances* para a descrição do sistema estudado, aplicando os conceitos de *Machine Learning* para modelar a Evapotranspiração de Referência (ET_o), comparando o método proposto com o método de Penman-Monteith (FAO) e, por fim, apresentar qual técnica é capaz de modelar o parâmetro em questão de maneira mais eficiente.

Capítulo 2

Fundamentação Teórica

2.1 O Processo de Evapotranspiração

ET é a perda de água do solo por evaporação e a perda de água da planta por transpiração, onde esses dois processos, que são simultâneos, precisam ser igualmente mensurados [8]. Normalmente expressa em milímetros por unidade de tempo (mm/d), essa taxa representa a quantidade de água perdida de um solo cultivado, caracterizando a lâmina de água evaporada ao longo de um determinado intervalo de tempo. A unidade de tempo pode ser hora, dia, mês, década ou até mesmo um ciclo inteiro da cultura [15].

2.2 Evaporação

Evaporação é o processo pelo qual a água líquida é convertida em vapor de água (vaporização) e removido da superfície de evaporação (remoção de vapor), sendo que essa água pode evaporar de uma variedade de superfícies, tais como lagos, rios, pavimentos, solos e vegetação húmida [16].

É necessária energia para mudar o estado das moléculas de água do líquido para o vapor, onde a radiação solar e, em menor medida, a temperatura ambiente do ar fornecem essa energia. A força motriz para remover o vapor de água da superfície de evaporação é a diferença entre a pressão de vapor de água à superfície de evaporação e a pressão de vapor atmosférica. À medida que a evaporação prossegue, o ar circundante se torna gradualmente saturado e o processo abrandará e poderá parar se o ar húmido não for transferido para a atmosfera [16, 17]. A substituição do ar saturado pelo ar mais seco depende muito da velocidade do vento. Assim, o

Radiação Solar Global ($MJ/m^2.d$) (**GSR**), a Temperatura do Ar ($^{\circ}C$) (**T**), a Umidade Relativa do Ar (%) (**RH**) e a Velocidade do Vento (m/s) (**WS**) são parâmetros climatológicos para avaliar o processo de Evapotranspiração (**ET**).

Outros fatores que podem afetar o processo de evaporação são a superfície do solo, o grau de sombreamento da cultura do dossel e a quantidade de água disponível na superfície de evaporação, sendo que chuvas frequentes, irrigação e a água transportada para cima em um solo de um lençol freático raso também podem molhar a superfície do solo [17]. Em locais onde o solo é capaz de fornecer água rápido o suficiente para satisfazer a demanda de evaporação, essa será determinada apenas pelas condições meteorológicas.

Contudo, se o intervalo entre as chuvas e a irrigação se tornar grande e a capacidade do solo de conduzir a umidade perto da superfície é pequena, o teor de água no solo cai e a superfície do solo seca. Sob estas circunstâncias a disponibilidade limitada de água exerce uma influência de controle sobre o solo. Na ausência de qualquer abastecimento de água à superfície do solo, a evaporação diminui rapidamente e pode cessar quase completamente dentro de alguns dias [18].

2.3 Transpiração

A transpiração consiste na vaporização da água líquida contida nos tecidos vegetais, seguida da remoção do vapor para a atmosfera. Culturas predominantemente perdem sua água através de estômatos [19]. Esses são pequenas aberturas na folha da planta através das quais os gases e o vapor de água passam (figura 2.1).

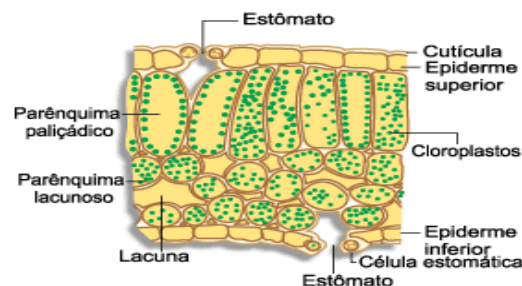


Figura 2.1: Representação esquemática de uma folha em corte, evidenciando sua estrutura interna.

Fonte: <http://fisiologiavegetalporhilde.blogspot.com/>

A água, junto com alguns nutrientes, é absorvida pelas raízes e transportada através da planta. A vaporização ocorre dentro da folha, ou seja, nos espaços intercelulares e a troca de

vapor com a atmosfera é controlada pela abertura estomática. Quase toda a água tomada é perdida pela transpiração e somente uma fração minúscula é usada dentro da planta [20].

A transpiração, como a evaporação direta, depende do fornecimento de energia, pressão de vapor gradiente e vento. Assim, o **GSR**, a **T**, a **RH** e a **WS** devem ser considerados ao avaliar a transpiração [21]. O teor de água do solo e a sua capacidade de conduzir a água para as raízes também determinam a taxa de transpiração, assim como o alagamento do solo salinidade da água. A taxa de transpiração também é influenciada pelas características da cultura, aspectos ambientais e práticas de cultivo, onde diferentes tipos de plantas apresentam distintas taxas de transpiração. Não só o tipo de cultura, mas também o desenvolvimento da cultura, ambiente e a gestão devem ser considerados na avaliação da transpiração [19, 20, 21].

2.4 Evapotranspiração (ET)

A evapotranspiração é um processo hidrológico fundamental que abrange a combinação da evaporação da água da superfície da Terra e da transpiração das plantas. Este fenômeno essencial envolve a transformação de água líquida em vapor de água na superfície do solo e a liberação subsequente de vapor d'água pelas plantas durante os processos metabólicos, como a fotossíntese. Apresenta-se como um elo crítico nos ciclos naturais da água e desempenha um papel vital na hidrologia global e na ecologia dos ecossistemas terrestres [22].

Sua quantificação é de fundamental importância em atividades ligadas a gestão de bacias hidrográficas, em modelagens meteorológicas e hidrológicas e, sobretudo, no manejo hídrico da agricultura irrigada. O conhecimento da Evapotranspiração (ET) é imprescindível para a agricultura, uma vez que permite a realização da irrigação adequada para cada cultura e época do ano, tornando assim o uso da água racional e eficiente [23].

O procedimento mais preciso para se medir a ET é feito com o uso de lisímetros, que consistem de tanques contendo solo representativo da área de interesse e são utilizados sobretudo para medir os componentes do balanço de água no sistema solo-planta-atmosfera, sob tratos culturais e condições climáticas específicas. Podem apresentar diversas formas e sistemas de medidas, porém os que usam mecanismos de pesagem são considerados os mais precisos e sensíveis para a determinação direta da ET, sendo conhecidos como modelos de balanço de massa de água [24].

Apesar de precisos, os lisímetros apresentam alguns problemas que dificultam sua operação

em escalas de tempo muito reduzidas, tornando-os instrumentos de manutenção trabalhosa, além do alto custo para aquisição [25]. Dessa forma, os valores de **ET** geralmente são estimados a partir de elementos climatológicos, existindo vários métodos que vão desde equações complexas de energia (Penman-Monteith (FAO) (**PM-FAO**)) à equações simplificadas, as quais utilizam apenas uma variável, como por exemplo, o método de Thornthwaite (sistema de classificação climática criado por Charles Warren Thornthwaite, no qual o fator mais importante é a evapotranspiração potencial e a sua comparação com a precipitação típica de uma determinada área) [26].

2.5 Evapotranspiração Potencial (**ETp**)

A Evapotranspiração Potencial (**ETp**), representa a quantidade máxima de água que teoricamente poderia ser perdida da superfície terrestre sob condições climáticas ideais, quando não há limitações de água no solo. Essa medida teórica é fundamental para entender o potencial de perda de água da superfície da Terra e é frequentemente utilizada em estudos de balanço hídrico e modelagem hidrológica [9].

2.6 Evapotranspiração Real (**ETr**)

A **ETr** se caracteriza pela quantidade efetiva de água perdida da superfície terrestre em uma localização específica durante um período determinado. Isso leva em consideração as limitações reais de água no solo, disponibilidade de água e outros fatores que afetam a taxa de evapotranspiração. A **ETr** é crucial para avaliar o uso sustentável da água em áreas agrícolas e na gestão de recursos hídricos [8].

2.7 Evapotranspiração da Cultura (**ETc**)

A Evapotranspiração da Cultura (**ETc**) é a quantidade de água perdida por uma cultura específica para a atmosfera durante seu ciclo de crescimento. Essa medida é essencial na agricultura para determinar as necessidades de irrigação das culturas, otimizando o uso da água e garantindo o crescimento saudável das plantas [10].

2.8 Evapotranspiração Real da Cultura (ETRc)

A quantidade de água efetivamente perdida por uma cultura específica em um local determinado é denominada Evapotranspiração Real da Cultura (ETRc), levando em consideração as condições reais, como umidade do solo, cobertura vegetal e clima. Essa medida é essencial para determinar a necessidade de irrigação em sistemas agrícolas e otimizar o uso da água [10].

2.9 Taxa de Evapotranspiração

Indica a quantidade de água que é perdida para a atmosfera em um período de tempo específico e em uma localização específica. Geralmente, é medida em milímetros por dia ou milímetros por ano, sendo uma métrica fundamental para entender as demandas hídricas em diferentes regiões [9].

2.10 Taxa de Evaporação (ET)

Quando aferida apenas a quantidade de água que é perdida apenas pela evaporação da superfície, excluindo a transpiração das plantas, temos a Taxa de Evaporação. Isso é frequentemente medido em corpos d'água, como lagos e reservatórios, e é crucial para avaliar a disponibilidade de água em reservatórios e corpos d'água [27].

2.11 Coeficiente de Cultura (Kc)

O Coeficiente de Cultura (Kc) é um fator multiplicador que ajusta a Evapotranspiração de Referência (ETo) para determinar a Evapotranspiração Real da Cultura (ETRc). Esse coeficiente varia de acordo com o tipo de cultura e seu estágio de desenvolvimento, permitindo um cálculo mais preciso da demanda hídrica das culturas [8].

2.12 Evapotranspiração de Referência (ETo)

Diversos trabalhos envolvendo a estimativa da Evapotranspiração de Referência (ETo), utilizando o método de Penman-Monteith (FAO) (PM-FAO) têm sido desenvolvidos [26]. Esse é considerado pela Organização das Nações Unidas para Agricultura e Alimentação (FAO) como

o método padrão para a estimativa de ET_o , sendo essa a ET de uma cultura bem adaptada ao local, crescendo sob as mesmas condições, usada como uma evapotranspiração padrão, através dos coeficientes de culturas, uma vez que combina componentes aerodinâmicos e de balanço de energia [8]. Sua validade é aceita em todos ambientes, visto que considera diversos parâmetros disponíveis (WS , GSR , RH e T).

Devido à escassa disponibilidade de estações meteorológicas de referência, definidas como estações agrometeorológicas localizadas em áreas com vegetação, capazes de medir todos os elementos climáticos envolvidos no cálculo da evapotranspiração Evapotranspiração (ET), estudos que se baseiam na avaliação de metodologias para a estimativa da Evapotranspiração (ET) utilizando o mínimo de dados comprovam que os resultados obtidos pelo método de Penman-Monteith (FAO) ($PM-FAO$) são aceitáveis para diferentes condições climáticas, tornando, assim, o método adequado para qualquer região [28, 29, 30, 31].

Entretanto, alguns métodos empíricos demandam uma quantidade menor de elementos climatológicos e podem ser considerados de boa aplicabilidade [32]. Apesar dos vários estudos avaliando diferentes métodos para estimativa de ET_o , as peculiaridades dos fatores climáticos de cada local exigem a avaliação destes métodos para maior exatidão na sua estimativa.

Alguns desses métodos são brevemente apresentados a seguir.

2.12.1 Método de Hargreaves-Samani

É uma abordagem empírica que estima a Evapotranspiração de Referência (ET_o) com base nas temperaturas máximas e mínimas diárias e na latitude do local. Esse método é simples e amplamente aplicável, mas não leva em consideração a umidade do ar, o que pode ser uma limitação em regiões com variações climáticas significativas [33].

2.12.2 Método de Thornthwaite

Esse método é baseado na temperatura média mensal e foi desenvolvido principalmente para estimar a Evapotranspiração Potencial (ET_p) em regiões temperadas. No entanto, ele tem limitações em áreas com climas tropicais ou desérticos e não leva em consideração a Umidade Relativa do Ar (%) [34].

2.12.3 Método de Penman-Monteith (FAO-56)

O método de Penman-Monteith é uma abordagem amplamente reconhecida e considerada uma referência para estimar a Evapotranspiração de Referência (ET_o). Ele leva em consideração várias variáveis meteorológicas, incluindo Temperatura do Ar ($^{\circ}C$), Umidade Relativa do Ar (%), Velocidade do Vento (m/s) e Radiação Solar Global ($MJ/m^2.d$). O método é mais complexo do que outros, mas oferece resultados mais precisos e é recomendado pela Organização das Nações Unidas para Agricultura e Alimentação (FAO) [8].

Vários fatores são levados em consideração para escolher o método de estimativa da Evapotranspiração de Referência (ET_o), como a disponibilidade dos dados meteorológicos e a escala de tempo pretendida [35]. Sendo assim, apesar do método ser considerado padrão pela Organização das Nações Unidas para Agricultura e Alimentação (FAO), o método de Penman-Monteith (FAO) (PM-FAO) pode muitas vezes ser limitado, já que emprega um número grande de elementos climáticos (WS, GSR, RH e T) que nem sempre estão disponíveis, tendo em vista o alto custo de aquisição dos instrumentos utilizados para as medições.

Para a estimativa de ET_o pelo método de Penman-Monteith (FAO), utiliza-se a seguinte expressão [8]:

$$ET_o = \frac{0,408\Delta(R_n - G) + \frac{\gamma 900U_2(e_s - e_a)}{T + 273}}{\Delta + \gamma(1 + 0,34U_2)} \quad (2.1)$$

em que:

- Δ é a declividade da curva de pressão de vapor em relação à temperatura ($KPa^{\circ}C^{-1}$);
- R_n é o saldo de radiação diário ($MJm^{-2}dia^{-1}$);
- G é o fluxo total diário de calor no solo ($MJm^{-2}dia^{-1}$);
- γ é o coeficiente psicrométrico ($KPa^{\circ}C^{-1}$);
- U_2 é a velocidade do vento a 2 metros de altura (ms^{-1});
- e_s é a pressão de saturação de vapor (kPa);
- e_a é a pressão atual de vapor (kPa);
- T é a temperatura média do ar ($^{\circ}C$).

Quando os valores de R_n , G , U_2 e T são medidos em estações meteorológicas, é preciso calcular os valores de Δ , γ , e_s e e_a .

Para isso, faz-se uso das seguintes relações:

$$\Delta = \frac{4098 \left[0,6108 \cdot \exp\left(\frac{17,27T}{T + 237,3}\right) \right]}{(T + 237,3)^2} \quad (2.2)$$

onde \exp é a base do logaritmo natural e , com $e = 2,71828$;

$$\gamma = 0,665 \cdot 10^{-3} \cdot Patm \quad (2.3)$$

em que $Patm$ é a pressão atmosférica local (kPa), calculada com base na altitude local (z);

$$Patm = 101,3 \left(\frac{293 - 0,0065z}{293} \right)^{5,26} \quad (2.4)$$

com z sendo a altitude local;

A diferença entre e_s e e_a é denominada déficit de saturação. Esses valores podem ser calculados através das seguintes expressões:

$$e_s = 0,6108 \cdot \exp\left(\frac{17,27T}{T + 237,3}\right) \quad (2.5)$$

e,

$$e_a = \frac{e_s \cdot UR}{100} \quad (2.6)$$

com UR sendo a a umidade relativa do ar (%), a qual é fornecida pela estação meteorológica.

Considerando as dificuldades encontradas na obtenção da Evapotranspiração de Referência ([ETo](#)), a ampla disponibilidade de dados de temperatura do ar nas estações meteorológicas e a

facilidade de obtenção de termômetros devido aos seus custos reduzidos, este trabalho adota o método padrão de Penman-Monteith (FAO) (**PM-FAO**) como parâmetro na comparação com métodos alternativos provenientes do *Machine Learning* (**ML**).

Aliadas aos métodos já apresentados, a Modelagem Matemática e algumas Técnicas de Inteligência Artificial (**IA**) são pontos extramamente úteis e fundamentais nessa pesquisa.

2.13 Modelagem Matemática

Define-se Modelagem Matemática como um processo de criação de modelos, através de estratégias de ação bem definidas, a qual busca representar a realidade com a subjetividade necessária ao evento estudado, ou seja, é uma tática para se obter explicações ou entendimentos de fatos reais [36].

Se destaca como uma aliada para auxiliar o detalhamento dos ecossistemas, sendo que a descrição dos ambientes, seja da perspectiva experimental, seja do ponto de vista teórico, constitui um dos mais intrigantes temas de pesquisa dos tempos modernos e um grande desafio à criatividade dos cientistas [37] e que dada a sua complexidade, usa-se a análise de sistemas para a sua compreensão, pois ela permite que fenômenos complexos sejam reduzidos em partes elementares, possibilitando a aplicação de métodos quantitativos [38].

Seguindo nesta linha de raciocínio, salienta-se o aspecto conectivo do conjunto, formando uma unidade, é um sistema estruturado de objetos e/ou atributos [39]. Esses objetos e atributos consistem de componentes ou variáveis (isto é, fenômenos que são passíveis de assumir magnitudes variáveis) que exibem relações discerníveis um com os outros e operam conjuntamente como um todo complexo, de acordo com determinado “padrão”. Percebe-se que para o estudo nesta área, importante se faz estar atento a multidisciplinaridade, a qual é necessária para uma compreensão mais abrangente dos sistemas naturais [40].

Observa-se que diversos aspectos da Matemática, da Física, da Química, das Ciências da Terra e, sobretudo, da Biologia, encontram-se presentes na investigação de sistemas naturais, em praticamente todas as abordagens propostas, mesmo aquelas mais simples, que procuram descrever o seu comportamento, observado no campo e no laboratório [37].

Aceitando como sendo mais adequada e completa ao contexto, pode-se afirmar que “um modelo é uma estruturação simplificada da realidade que supostamente apresenta, de forma generalizada, características ou relações importantes. São aproximações altamente subjetivas, por

não incluïrem todas as observações ou medidas associadas, mas são valiosos por obscurecerem detalhes acidentais e por permitirem o aparecimento dos aspectos fundamentais da realidade” [41]. Ressalta-se que um modelo matemático é quase sempre um sistema de equações ou inequações algébricas, diferenciais, integrais, etc., obtido através de relações estabelecidas entre as variáveis consideradas essenciais ao fenômeno sob análise [36].

Vale ainda destacar que devemos decidir sobre o grau de generalidade que iremos admitir e devemos ter em mente que nem toda hipótese que simplifica irá facilitar o tratamento posterior do modelo, tanto do ponto de vista teórico como do ponto de vista computacional [42]. Modelos geralmente se encontram fundamentados no conhecimento, ou nas pressuposições, das leis dos processos físicos e químicos e para uma análise determinística é preciso trabalhar com dados primários, ou seja, dados quantitativos que possuam, obrigatoriamente, unidades de medida que possam ser organizadas em classes [43].

Logo, faz-se necessário reestruturar os modelos existentes através de Computação Inteligente, explorando assim os aspectos não lineares, suas capacidades de adaptações e incertezas [44]. Pode-se dizer, em poucas palavras, que a modelagem matemática é um procedimento teórico/prático, motivado por seu criador a fim de entender a realidade, buscando meios para agir e transforma-lá. Para isso, técnicas de IA são grandes aliadas na procura de modelos mais eficazes, acarretando em resultados de maior confiabilidade.

2.14 Inteligência Artificial

A palavra inteligência vem do latim [45] que se divide em *inter* (entre) e *legere* (escolher), ou seja, inteligência é aquilo que o homem pode escolher entre uma coisa e outra, sendo que ela é o modo de resolver problemas ou de realizar tarefas. Dada esta definição, podemos considerar que Inteligência Artificial (IA) é todo tipo de inteligência capaz de beneficiar máquinas com alguma habilidade semelhante a dos humanos [46, 47].

Uma denominação mais simples apresenta a IA como uma disciplina que tem por objetivo o estudo e a construção de entidades artificiais com capacidades cognitivas semelhantes às dos seres humanos [48, 49]. É uma ciência que procura estudar e compreender o fenômeno da inteligência, e, por outro lado, um ramo da engenharia, na medida em que procura construir instrumentos para apoiar o ser humano [50].

Tais sistemas inteligentes têm como um dos seus propósitos, o de habilitar o computador

para executar funções que são desempenhadas pelo ser humano fazendo uso do conhecimento e do raciocínio. A capacidade inerente aos humanos de agir inteligentemente é habitualmente associada ao conhecimento que estes detêm, logo o conhecimento é um requisito essencial para a construção destes sistemas.

Ela é capaz de sistematizar e automatizar procedimentos, sendo importante para a aceleração e realização de atividades humanas. Localiza-se num campo universal, ou seja, pode ser aliada em qualquer área estudada, possuindo uma enorme variedade de subcampos, tais como demonstração de teoremas matemáticos e diagnóstico de doenças [51].

Diversos são os algoritmos capazes de analisar uma grande quantidade de dados de maneira mais rápida e precisa que o ser humano, sendo estes aptos a fazer previsões de aprendizado, com o objetivo de agrupar dados, fornecer informações, indicar padrões e tendências, ajudando cada vez mais as pessoas a tomarem decisões [52].

2.14.1 A Inteligência Artificial Aplicada a Estudos Ambientais

A Inteligência Artificial (IA) desempenha um papel crucial na previsão e monitoramento de mudanças climáticas. Algoritmos de aprendizado de máquina são capazes de analisar séries temporais complexas de dados climáticos, identificando tendências, padrões e correlações que podem passar despercebidos pelos métodos tradicionais. Isso resulta em previsões climáticas mais precisas, permitindo uma melhor preparação para eventos climáticos extremos [53].

Outro benefício notável da IA é sua contribuição para a conservação da biodiversidade. Através de diferentes algoritmos, é possível monitorar e rastrear populações de espécies ameaçadas, identificar habitats críticos e até mesmo prever o risco de extinção. Essa abordagem permite a implementação de estratégias de conservação mais eficazes [54].

Também é uma ferramenta fundamental na gestão sustentável de recursos naturais, como água, florestas e energia. Algoritmos de otimização podem ajudar a determinar a alocação mais eficiente desses recursos, garantindo seu uso sustentável e minimizando o impacto ambiental [55].

Além disso, a Inteligência Artificial (IA) tem um papel crítico na detecção de atividades ilegais, como desmatamento e caça furtiva. Sistemas de visão computacional e análise de dados de sensoriamento remoto podem identificar de forma autônoma padrões suspeitos e alertar as autoridades, possibilitando uma resposta mais rápida e eficaz [56].

2.15 *Machine Learning*

O Aprendizado de Máquina (AM) (do inglês *Machine Learning* (ML)) está diretamente ligado com a inteligência, pois “se um sistema é capaz de aprender a exercer determinada tarefa, então merece ser chamado de inteligente” [57]. Define-se como um aprendizado por experiência, aquele que conforme cada tarefa é realizada, o sistema aprende uma nova e melhor maneira de resolvê-la, estruturando-se do conhecimento já existente, para gerar um entendimento de aprendizado.

Um processo de aprendizagem inclui a aquisição de novas formas de conhecimento: o desenvolvimento motor e a habilidade cognitiva (através de instruções ou prática), a organização do novo conhecimento (representações efetivas) e as descobertas de novos fatos e teorias através da observação e experimentação. Desde o início da era dos computadores, tem sido realizadas pesquisas para implantar algumas destas capacidades em computadores. Resolver este problema tem sido o maior desafio para os pesquisadores de inteligência artificial. O estudo e a modelagem de processos de aprendizagem em computadores e suas múltiplas manifestações constituem o objetivo principal do estudo de aprendizado de máquinas [58].

Utilizando o conceito de *Machine Learning* (ML), ramo da inteligência artificial baseado na ideia de que sistemas podem modificar seu comportamento, aprendendo com dados, identificando padrões e tomando decisões com o mínimo de intervenção humana, ou seja, tornando-se autônomo a partir de sua própria experiência, acredita-se que regras lógicas possam ser estabelecidas para descrever um sistema complexo [59, 60]. Esses são definidos como locais em que não é possível inferir um comportamento [61], pois a interação dos parâmetros presentes produz saídas não previsíveis e não lineares, sendo esta uma possível explicação para o comportamento de fenômenos da natureza [61].

2.15.1 O *Machine Learning* na Análise de Sistemas Ambientais

Nos últimos anos, o *Machine Learning* (ML) emergiu como uma ferramenta poderosa na análise de sistemas ambientais. A capacidade de os algoritmos de ML aprenderem e identificarem padrões complexos em grandes conjuntos de dados tem se mostrado altamente vantajosa na compreensão, monitoramento e previsão de fenômenos ambientais críticos [62].

Um dos domínios em que o *Machine Learning* tem feito avanços significativos é o monitoramento das mudanças climáticas. Algoritmos de **ML** podem analisar vastas quantidades de dados climáticos, incluindo temperaturas, níveis de gases de efeito estufa e dados de satélites. Eles identificam correlações não triviais e fornecem previsões mais precisas das tendências climáticas e eventos extremos [62].

A conservação da biodiversidade também se beneficia do *Machine Learning* (**ML**). Através da análise de dados de sensoriamento remoto, pode identificar áreas de habitat crítico, monitorar populações de espécies ameaçadas e até mesmo prever riscos de extinção [63].

Outra aplicação importante está na gestão de recursos naturais. Algoritmos de *Machine Learning* (**ML**) podem otimizar o uso de recursos hídricos, prever padrões de desmatamento e auxiliar na alocação de recursos energéticos de forma mais eficiente e sustentável [64].

O *Machine Learning* (**ML**) também desempenha um papel crucial na detecção e prevenção de atividades ilegais relacionadas ao meio ambiente, como a caça furtiva e o desmatamento ilegal. Sistemas de visão computacional podem identificar automaticamente atividades suspeitas em imagens de satélite, permitindo uma resposta mais rápida por parte das autoridades [65].

2.15.2 Os Métodos de Aprendizagem de Máquina

Vários são os métodos de aprendizado de máquina [57] como, por exemplo, o método de aprendizado por conceito, que analisa hipóteses e demonstra qual é a correta. É um sistema complexo, mas com muita utilidade, pois é muito preciso quando comparado a ambientes de aprendizagem comuns. Tais métodos são categorizados pelo tipo de aprendizado utilizado, sendo eles o supervisionado, não supervisionado e por reforço.

O aprendizado supervisionado consiste na máquina em aprender uma função que mapeia uma entrada visando uma saída, com base em pares de entrada-saída de exemplo [51]. É capaz de inferir uma função a partir de dados de treinamento rotulados consistindo de um conjunto de exemplos de treinamento. Na análise preditiva existem dois tipos de modelos de aprendizado: os modelos de regressão e de classificação.

Os modelos de regressão preveem resultados com base em rótulos numéricos (são propriedades semelhantes que os objetos de uma classe possuem, onde cada atributo permite definir um intervalo de valores que as instâncias dessa propriedade podem apresentar), já os de classificação dividem seus dados em conjuntos finitos (azul claro e azul escuro), como pode ser observado na figura 2.2.

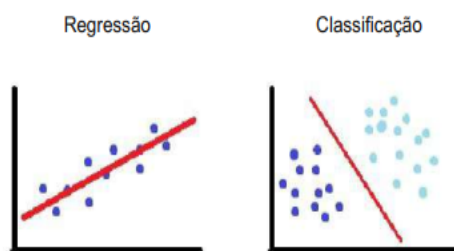


Figura 2.2: Modelo de Regressão e Classificação.

Já no aprendizado não supervisionado, temos um funcionamento independente de classificações, ou seja, capaz de prever resultados ao interpretar uma informação de entrada sem que nenhum tipo de classificação seja definida inicialmente [66], enquanto que no aprendizado por reforço os agentes inteligentes necessitam fazer exame de ações em um ambiente a fim maximizar a noção da recompensa cumulativa.

Difere da aprendizagem supervisionada em não necessitar de pares etiquetados da entrada-saída e em não necessitar ações para correções explícitas para o funcionamento. Seu foco é encontrar um equilíbrio entre a exploração (de território inexplorado) e a exploração (do conhecimento atual) [67, 68].

Devido a quantidade de dados disponíveis para a pesquisa, bem como os resultados calculados através do método padrão de Penman-Monteith (FAO) (PM-FAO), esse trabalho fará uso de métodos supervisionados, uma vez que procuramos modelar a Evapotranspiração de Referência (ET_o) através de uma quantidade menor de parâmetros, baseando-se nos valores calculados.

Dentre os métodos supervisionados, utilizou-se neste trabalho àqueles denominados Redes Neurais Artificiais (RNA), *Random Forest* (RF) e *Support Vector Machine* (SVM), os quais serão abordados a seguir.

2.15.3 Redes Neurais Artificiais (RNA)

Conhecidas como RNA, as Redes Neurais Artificiais, são uma grande rede de neurônios, organizadas geralmente em camadas, que tem semelhança com o funcionamento do cérebro humano, onde a primeira camada é aquela que recebe informações a serem classificadas, faz uso do aprendizado supervisionado para modificar a forma das conexões, de acordo com o informado e, por último, ativa os neurônios de saída (figura 2.3).

São algoritmos computacionais matemáticos baseados em estruturas neurais de organismos

inteligentes, capazes de reproduzir a atividade cerebral humana de maneira simples, os quais podem ser utilizadas para elaboração de um esquema artificial capaz de armazenar e ampliar conhecimento baseado em experiência (aprendizagem), disponibilizando saídas analíticas para diferentes aplicações [66, 69].

Podem ser utilizadas em situações complexas como, por exemplo, o Bioma Pantanal, local onde a previsibilidade das variáveis não é amplamente conhecida. Devido ao grande sucesso das RNA na modelagem de sistemas dinâmicos [70, 71] é sugestivo que a mesma possa ser uma ferramenta eficaz quando aplicada em análise ambiental, podendo ser capaz de apresentar soluções complexas, dependentes de vários critérios, o que seria em certos momentos inviável para as ferramentas usuais [72].

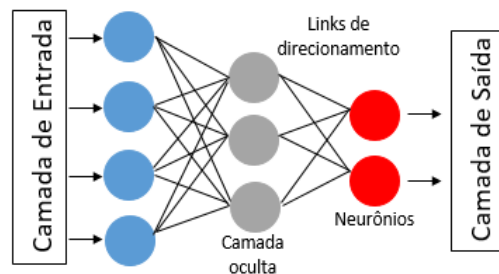


Figura 2.3: Diagrama de funcionamento de uma Rede Neural Artificial.

2.15.4 *Random Forest* (RF)

É um algoritmo de aprendizagem supervisionada, o qual cria uma “floresta” de modo aleatório. Essa “floresta” é a combinação (*ensemble*) de árvores de decisões, em sua maioria treinadas com o método de *bagging* (o método consiste em gerar subconjuntos de exemplos através de sorteio simples com reposição, sobre o conjunto de dados de treinamento original), consistindo de uma combinação de modelos de aprendizado que aumentam a qualidade do resultado geral (figura 2.4), portanto, o método gera várias árvores de decisões e as combina para a obtenção de uma predição melhor e mais estável [73].

Uma de suas vantagens consiste em poder ser utilizado tanto para classificação quanto para regressão, o que representa a maioria dos sistemas de aprendizagem de máquina atuais [73, 74]. A seguir, tem-se um diagrama de uma *Random Forest* (RF).

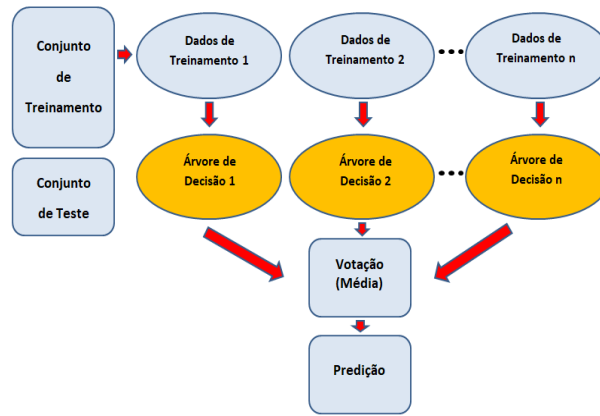


Figura 2.4: Diagrama de funcionamento do algoritmo *Random Forest*.

Em suas grandes majorias, um classificador de *Random Forest* (RF) possui todos os hiperparâmetros de uma árvore de decisão e também os hiperparâmetros de um classificador do tipo *bagging*, controlando assim as combinações de árvores [75]. Ao invés de se construir um classificador *bagging* e passá-lo para um classificador de árvore de decisão, pode-se usar a classe da RF, o que é mais conveniente e otimizado para as decisões.

O modelo adiciona aleatoriedade extra ao sistema, quando se está criando árvores. Ao invés de procurar pela melhor característica na partição de nodos, ele busca a melhor característica em um subconjunto aleatório das características. Desse modo é criada uma grande diversidade, a qual leva a geração de modelos mais consistentes e, portanto, quando se cria uma árvore no RF, apenas um subconjunto aleatório das características é considerado na partição de um nodo [74].

Outra importante qualidade da RF é a facilidade para definir a importância relativa de cada característica para a predição [76]. Através da inspeção da importância das características, pode-se decidir quais deixar de fora, já que estas não contribuem o suficiente para o processo. Isto é importante, porque uma regra geral em *Machine Learning* é que quanto mais características você tem, maiores são as chances do modelo sofrer um *Underfitting* (significa que o modelo não conseguiu aprender o suficiente a partir dos dados. Resulta em erros elevados tanto nos dados de treinamento quanto nos dados de teste) ou um *Overfitting* (ocorre quando um modelo aprende em excesso a partir dos dados de treinamento. Nesse caso, o modelo se adapta muito especificamente aos dados de treino, capturando até mesmo o ruído e as variações irrelevantes presentes neles. Como resultado, o modelo não consegue generalizar bem para novos dados que nunca foram vistos antes, levando a um desempenho significativamente inferior nos dados de teste em comparação aos dados de treinamento) [73].

2.15.5 *Support Vector Machine*

Support Vector Machine (SVM) consiste de uma classificação para dois conjuntos de dados, caso o delimitador entre as classes seja linear [77]. Entretanto, na prática, lidamos com limites não lineares. Assim, uma solução é mapear o conjunto de treinamento em seu espaço original (não linear) para um novo espaço de maior dimensão, denominado espaço de características (*feature space*), que é linear [78].

Logo, é preciso encontrar uma transformação não linear, $\varphi(x) = [\varphi_1(x), \dots, \varphi_m(x)]$, que seja capaz de mapear o espaço original das observações para um novo espaço de atributos m -dimensional, o qual pode ser muito maior do que o espaço original, onde nesse novo espaço, as observações passam a ser linearmente separáveis [78], conforme figura 2.5.



Figura 2.5: Esquema da função de transformação *Support Vector Machine*.

Obtida esta função de transformação, o problema de otimização recai pra uma *Support Vector Machine* (SVM) linear.

Os novos dados são mapeados no mesmo espaço e preditos como pertencentes a uma categoria, baseado em qual lado do espaço eles são colocados. Em outras palavras, o SVM encontra uma linha de separação, denominada hiperplano, entre dados de duas classes. Tal limitador busca maximizar a distância entre os pontos mais próximos em relação a cada uma das classes [79] (figura 2.6).

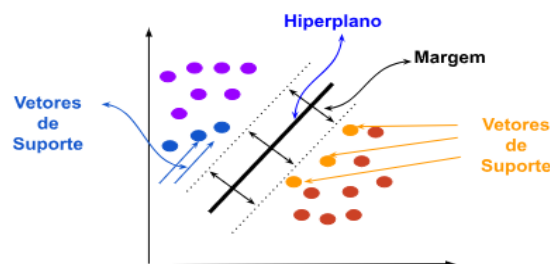


Figura 2.6: Esquema do método *Support Vector Machine*.

A distância entre o hiperplano e o primeiro ponto de cada classe é chamada de margem, de modo que a SVM coloque em primeiro lugar a classificação das classes, definindo assim cada ponto a uma determinada classe e, em seguida, maximiza esta margem, ou seja, ela primeiro

classifica as classes corretamente e depois, em função desta restrição, define a distância entre as margens [80].

Vale destacar que os dados utilizados neste trabalho sofrem variações ao longo do tempo, logo se faz necessário compreender a estrutura da coleção, o que é de grande importância pois podemos assim tomar decisões a partir dos esclarecimentos gerados com as análises.

Capítulo 3

Material e Métodos

Fundamentado na análise de diferentes variáveis climáticas, tais como Temperatura do Ar ($^{\circ}\text{C}$) (**T**), Umidade Relativa do Ar (%) (**RH**), Radiação Solar Global ($\text{MJ}/\text{m}^2.\text{d}$) (**GSR**) e Velocidade do Vento (m/s) (**WS**), coletadas *in loco*, as quais ajudam a descrever a Evapotranspiração de Referência (**ET_o**) de uma região de transição do Cerrado Mato-Grossense, apresenta-se uma proposta alternativa para entender como a Evapotranspiração de Referência (**ET_o**) é influenciada por tais elementos, compreendendo assim como são os efeitos dos agentes externos sobre a vegetação local.

3.1 Área de Estudo

A partir de dados meteorológicos extraídos na Fazenda Miranda (figura 3.1), localizada em $15^{\circ}43'53''\text{S}$ e $56^{\circ}04'18''\text{W}$, a 15 km SSE de Cuiabá, Mato Grosso, Brasil [81], coletas dos parâmetros **T**, **RH**, **GSR** e **WS** foram realizadas em intervalos de tempo iguais a 30 minutos, ao longo de 10 anos, a partir do ano de 2009. Esses dados apresentaram falhas amostrais inferiores a 5%.

O local é caracterizado por possuir uma vegetação do tipo floresta-pastagem, localmente conhecida como “campo sujo” [82, 83] (é um tipo fisionômico exclusivamente herbáceo arbustivo, com arbustos e subarbustos esparsos cujas plantas, muitas vezes, são constituídas por indivíduos menos desenvolvidos das espécies arbóreas do Cerrado).

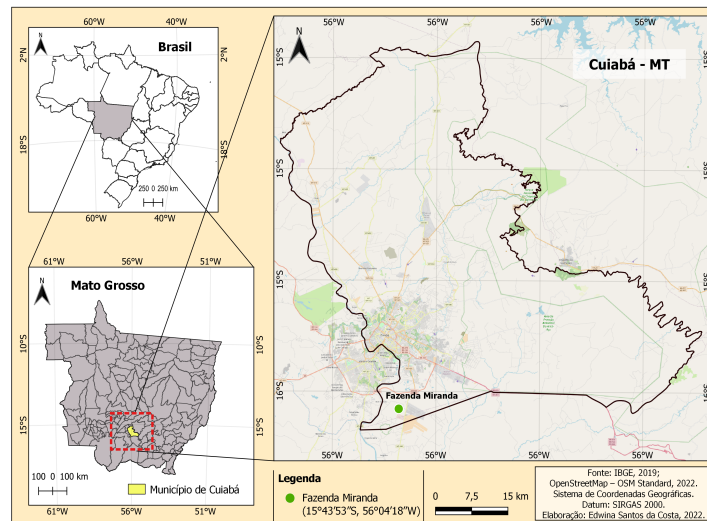


Figura 3.1: Localização da área de estudo.

3.2 Medições Micrometeorológicas

Tais medições foram realizadas pelos seguintes sensores: Termohigrômetro (HMP45AC, Vaisala Inc., Woburn, MA, EUA), utilizado para aferir a Temperatura do Ar ($^{\circ}\text{C}$) e a Umidade Relativa do Ar (%), instalado a 2 metros de altura em relação ao solo, Piranômetro (LI200X, LI-COR Biosciences, Inc., Lincoln, NE, USA) utilizado para medir a Radiação Solar Global ($\text{MJ}/\text{m}^2.\text{d}$), instalado a 5 metros de altura em relação ao solo e o Anemômetro (03101 R. M. Young Company) usado na medição da Velocidade do Vento (m/s), localizado a 10 metros de altura em relação ao solo e, fazendo uso da média aritmética simples dos parâmetros diários acima citados, juntamente com o Método de PM-FAO, foi possível determinar a Evapotranspiração de Referência (ET_0) dia a dia no período. Assim, os dados coletados foram tabulados e organizados de acordo com a sua evolução temporal, gerando correlações entre os 4 (quatro) parâmetros de entrada (T , GSR , RH e WS) e o parâmetro de saída ET_0 .

3.3 Técnicas Utilizadas

Algoritmos de *Machine Learning* foram elaborados, sendo feita a variação dos parâmetros de entrada e exigindo 1 (um) parâmetro de saída. Tais *scripts* foram treinados com 70% das variáveis coletadas em sua forma original (dados do ano de 2009 a 2016) e validados nos 30% restantes (anos de 2017 até 2019), gerando diferentes modelos analíticos capazes de prever a ET_0 no sistema analisado.

Faz-se importante analisar a relação entre os parâmetros gerados por um modelo e os valores reais observados, visto que um modelo para predição deve ser criado pensando em se alcançar dois principais objetivos: selecionar e avaliar os modelos. No primeiro ponto citado, a *performance* é observada para se escolher aquele com melhor desempenho, enquanto o segundo objetivo é utilizado para se buscar uma estimativa de erro de generalização para novas observações.

No algoritmo de Redes Neurais Artificiais (RNA), foram aplicados na camada de entrada os parâmetros coletados *in loco*, exigindo como saída os valores de Evapotranspiração de Referência (ET_o) já conhecidos através de Penman-Monteith (FAO) (PM-FAO). Neste mesmo caminho, utilizando uma menor quantidade de parâmetros, também foram realizadas as combinações entre os dados de entrada - um a um, dois a dois e três a três. Dessa forma foi possível analisar a melhor combinação, ou seja, àquela que trouxe o grau máximo de aproximação com a Evapotranspiração de Referência (ET_o) original, buscando assim o melhor modelo de saída, o qual foi dependente de uma menor quantidade de variáveis, acarretando desta forma em um erro menor após a sua validação e, conseqüentemente, uma maior confiança no método proposto.

O mesmo procedimento citado anteriormente foi realizado com as outras duas técnicas de *Machine Learning* (ML) estudadas, sendo elas a *Random Forest* (RF) e a *Support Vector Machine* (SVM). Prever um modelo confiável e que dependa de um número menor de parâmetros, pode ocasionar uma economia no sistema de coleta de dados, visto que, dependendo dos parâmetros a serem utilizados, o custo final na aquisição dos equipamentos medidores pode ser consideravelmente reduzido, além de trazer maior confiabilidade, uma vez que o erro agregado relativo as medições é menor.

Todos os testes foram realizados utilizando a linguagem de programação *Python* no ambiente *Google Colaboratory* (Google Colab). Também foram utilizadas bibliotecas de código aberto, como Keras e Scikit-learn, as quais possuem algoritmos de aprendizado de máquina implementados, como Redes Neurais Artificiais (RNA), *Support Vector Machine* (SVM) e *Random Forest* (RF).

Nos experimentos realizados, as Redes Neurais Artificiais (RNA) foram otimizadas através do teste de diversos, procurando assim encontrar a melhor configuração da rede. Foram testados diferentes números de neurônios, funções de ativações e algoritmos de treinamento. Tais operações foram executadas durante os testes, não se fazendo necessário salvar a combinação escolhida para determinado teste.

Já as *Random Forest* (RF) foram configuradas com 1000 (mil) árvores na floresta, sem limite de profundidade, apenas 2 (dois) como o número mínimo de amostras necessárias para dividir um nó interno e apenas 1 (um) como o número mínimo de amostras necessárias para estar em um nó folha. Essa configuração resulta em árvores de decisão mais flexíveis, apesar do aumento do tempo de processamento. No entanto, a quantidade de dados nesta pesquisa não impactou significativamente nos testes.

Para o *Support Vector Machine* (SVM), realizamos uma busca em grade com validação cruzada para definir os principais parâmetros: C, *epsilon* e *gamma*. Para cada parâmetro, verificamos três valores: 0,1, 10 e 1000 para C; 0,01, 1 e 10 para *epsilon*; 0,01, 1 e 10 para *gamma*. A melhor combinação foi utilizada para treinar o método, utilizando o kernel *Radial Basis Function* (RBF) (o kernel RBF permite que a SVM modele de forma eficiente relações não lineares complexas nos dados, tornando-o uma escolha comum em problemas de classificação ou regressão com fronteiras de decisão complexas).

3.4 Validação dos Modelos

Para avaliar cada protótipo gerado, calculou-se medidas de qualidade da previsão. Três medidas estatísticas foram utilizadas nessa etapa, as quais em conjunto puderam garantir a confiabilidade do modelo escolhido [84].

3.4.1 R - Squared (R^2)

A primeira medida calculada foi o Coeficiente de Determinação (R^2), que é o Coeficiente de Correlação Linear de Pearson ao quadrado entre os valores observados e previstos. Esse coeficiente está entre 0 (zero) e 1 (um) e quanto mais próximo de um, maior é a correlação entre os valores observados e previstos. Alta correlação significa, em geral, que quanto maior o valor observado na variável, maior tende a ser o previsto, de forma que está é uma medida de associação entre os valores observado e previsto, e não uma medida de erro do previsto em relação ao observado [85, 86].

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.1)$$

Esse é o fator resultante da divisão entre o quadrado das diferenças entre os valores observados e os valores estimados pela equação, e a soma dos quadrados das diferenças entre os valores observados e a média dos valores encontrados nos resultados, onde valores mais elevados indicam uma correlação mais robusta entre os dados.

3.4.2 Root Mean Square Error (RMSE)

O RMSE (Root Mean Square Error) foi a segunda métrica utilizada, o qual é a raiz quadrada da diferença ao quadrado média entre os valores observado (y_i) e previsto (y'_i):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (3.2)$$

Essa medida representa um erro médio de previsão, sendo que quanto maior for essa média, pior é a qualidade do modelo [85, 86].

3.4.3 Mean Absolute Error (MAE)

Por fim, como terceira e última métrica, foi calculado o erro MAE (Mean Absolute Error), que se baseia na diferença média em módulo (sem considerar o sinal) entre os valores observados (y_i) e previstos (y'_i):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (3.3)$$

Essa medida também representa um erro médio de previsão, sendo que quanto maior for essa média, pior é a qualidade do modelo [85, 86].

Capítulo 4

Resultados e Discussões

Compreender como cada parâmetro está associado com o valor da Evapotranspiração de Referência (ET_o) é de extrema importância, uma vez que a observação pode indicar qual(is) medida(s) possui(em) maior ou menos relação(ões) direta(s) com a variável de saída.

Assim, as imagens a seguir (figuras 4.1, 4.2, 4.3 e 4.4) trazem o comportamento temporal de cada variável de entrada coletada, ao longo de 10 (dez) anos, a partir do ano de 2009.

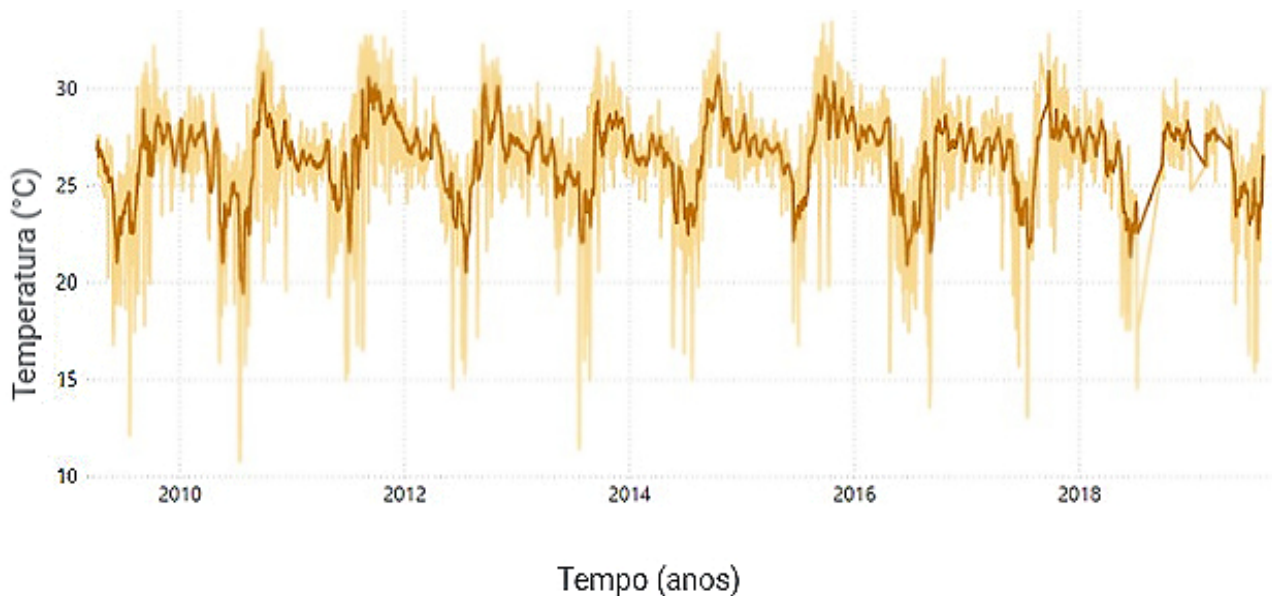


Figura 4.1: Evolução temporal da Temperatura do Ar (°C) durante 10 anos. Em destaque a Média Móvel de 15 dias.

Percebe-se que a variação temporal da Temperatura do Ar (°C) (T) ocorre numa amplitude limitada entre 10 °C e 35 °C (figura 4.1), mantendo uma média em torno de 27 °C ao longo do período estudado, indicando que o parâmetro possui considerável influência na ET_o [8].

No Brasil, estados do norte e nordeste geralmente apresentam uma amplitude térmica baixa,

porque os climas tropical e equatorial que atuam nas regiões são caracterizados por altas temperaturas ao longo do ano [87]. Já os estados do sudeste, sul e centro-oeste, em algumas estações do ano, apresentam uma amplitude térmica alta, sobretudo no inverno [87].

Isso acontece pelo tipo de relevo e vegetação do local, além do clima e das massas de ar que atuam nessas regiões e pela característica da água em resistir mais a mudanças de temperaturas, quando comparada com os sólidos, como por exemplo, a terra, sendo esta diferença de comportamento fundamental para se determinar a amplitude térmica de uma região [87, 88].

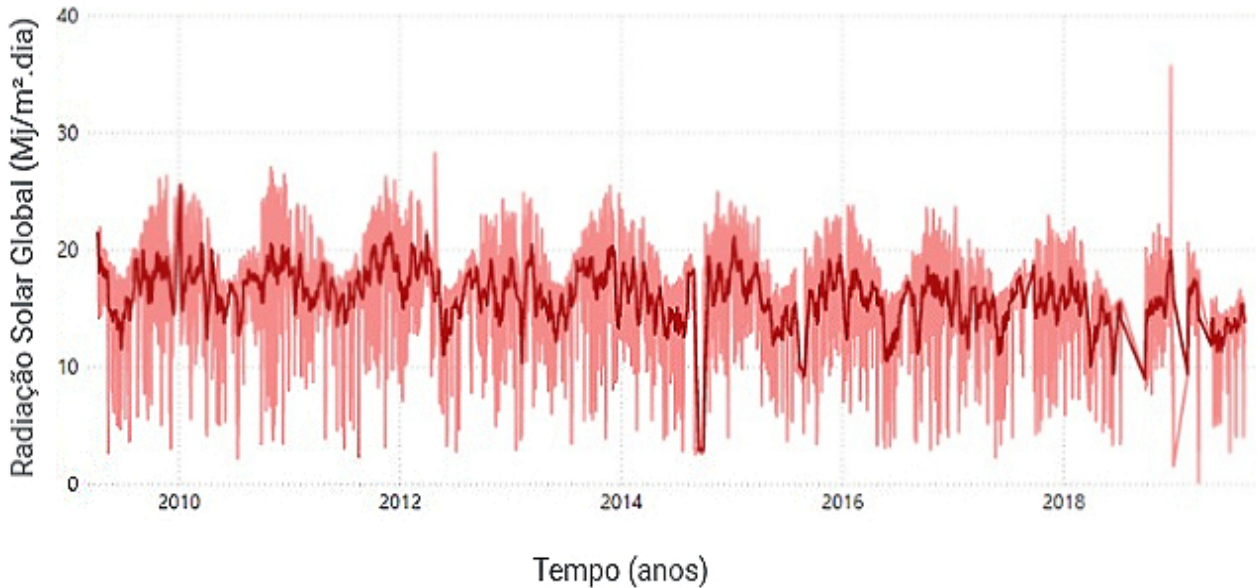


Figura 4.2: Evolução temporal da Radiação Solar Global ($MJ/m^2.d$) durante 10 anos. Em destaque a Média Móvel de 15 dias.

Influência maior em relação a variação anteriormente apresentada (**T**) é observada na figura 4.2, a qual mostra a evolução da Radiação Solar Global ($MJ/m^2.d$) (**GSR**), observando-se uma amplitude entre 3 e 30 ($MJ/m^2.d$), mantendo uma média em torno de 17 ($MJ/m^2.d$) no período em questão. Quando comparada com a **T**, entende-se que tal parâmetro é primordial para o cálculo da Evapotranspiração de Referência (**ET_o**) [8].

Variações na radiação solar (variações solares) é um termo que caracteriza as mudanças no tempo da radiação atual do Sol, sua distribuição espectral e os fenômenos que acompanham essas mudanças [88]. Faz-se uma diferenciação entre os componentes periódicos dessas mudanças, sendo a principal o ciclo solar de onze anos (a cada 11 anos, aproximadamente, os polos norte e sul da estrela se invertem) e as mudanças aperiódicas, ou seja, a energia solar que sai do Sol e atinge a superfície da Terra varia de tempos em tempos [88].

Essas mudanças na radiação solar são devidas à quantidade de energia emitida pelo Sol e

são detectáveis através das mudanças no brilho, variações no vento solar e no campo magnético [89]. Entre as causas que produzem esse fenômeno, a mais provável foi formulada por Horace Welcome Babcock (astrônomo estadunidense que inventou e construiu diversos instrumentos astronômicos e em 1953 foi o primeiro a propor a ideia da óptica adaptativa. Era especialista em espectroscopia e estudo do campo magnético de estrelas), o qual afirma que existe um elo entre os fenômenos eletrodinâmicos do Sol e os fenômenos eletromagnéticos [89, 90].

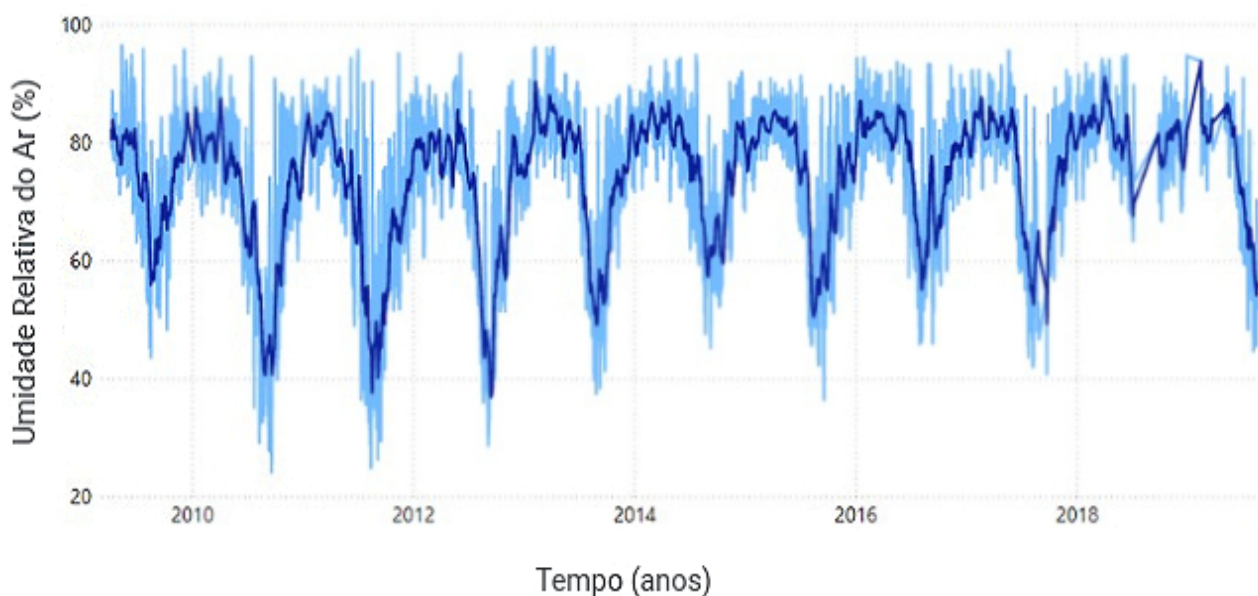


Figura 4.3: Evolução temporal da Umidade Relativa do Ar (%) durante 10 anos. Em destaque a Média Móvel de 15 dias.

Afetada diretamente pela **GSR**, a Umidade Relativa do Ar (%) (**RH**) apresenta também uma alta amplitude, entre 10% e 100% (figura 4.3), fixando uma média em torno de 75%. Este parâmetro, juntamente com a **GSR**, é considerado indispensável na obtenção do valor da **ETo** quando utilizado os meios tradicionais de cálculo [8].

A umidade do ar ou atmosférica é a quantidade de água existente no ar na forma de vapor. Dessa forma, é um dos mais relevantes elementos que atuam na atmosfera, pois a sua presença em maior ou menor grau influencia nas temperaturas, no regime de chuvas, na sensação térmica e até mesmo em nossa saúde [91].

Um dos fenômenos responsáveis pela variação da umidade em um local é a movimentação das massas de ar. Se essas estão carregadas de umidade, a tendência é aumentar os índices de chuva na região afetada, enquanto massas de ar seco tendem a conservar o ambiente igualmente sem umidade [91, 92]. Também se tem a vegetação, que, em alguns casos, também interfere na presença de umidade. Algumas regiões tendem a apresentar índices de umidade maiores

durante o ano, ao mesmo tempo em que também são responsáveis pela geração de umidade para outras regiões, uma vez que todo esse vapor d'água gerado é deslocado para outras áreas [91, 92].

Estes efeitos provenientes da umidade sobre o clima são observados tanto nas temperaturas quanto no regime de chuvas. A água, em razão de seu calor específico, tende a conservar por mais tempo as temperaturas, fazendo com que haja uma menor variação delas, ou seja, a amplitude térmica (diferença entre a maior e a menor temperatura local) é menor quanto maior for a umidade do ar [93]. Além disso, em regiões mais úmidas ou que estejam mais afetadas pela umidade, o regime de chuvas tende a ser maior, pois a saturação do ar que provoca a condensação é mais frequente [94].

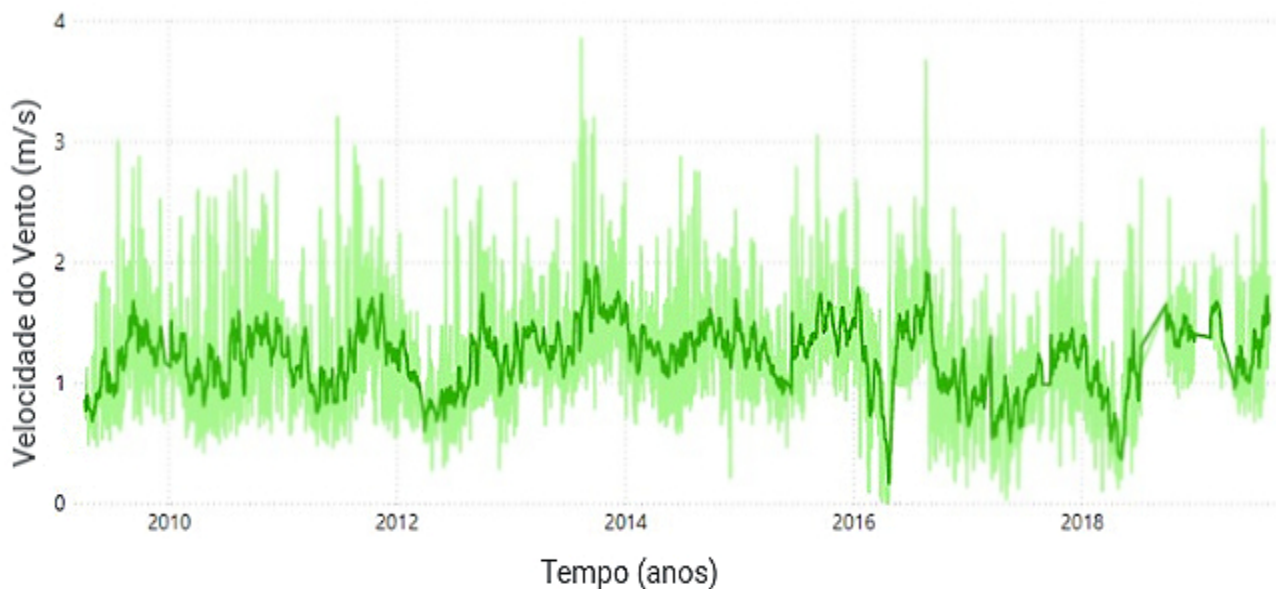


Figura 4.4: Evolução temporal da Velocidade do Vento (m/s) durante 10 anos. Em destaque a Média Móvel de 15 dias.

Oposto aos fatos anteriormente apresentados, observa-se através da figura 4.4 que a Velocidade do Vento (m/s) (WS) pode ser uma variável que não apresente influência significativa na ET_o , visto que sua amplitude oscila entre 0,5 m/s e 3,5 m/s, mantendo uma média em torno de 1,3 m/s.

Existem locais no globo terrestre nos quais os ventos jamais cessam de “soprar”, pois os mecanismos que os produzem (aquecimento no equador e resfriamento nos polos) estão sempre presentes na natureza [95]. As características topográficas (figura 4.5) de uma determinada região também influenciam no comportamento dos ventos, uma vez que, em uma determinada área, podem ocorrer diferenças de velocidade, ocasionando a redução ou aceleração na veloci-

dade do vento. Além das variações topográficas e de rugosidade do solo, a velocidade também varia seu comportamento de acordo com a altitude local [95].

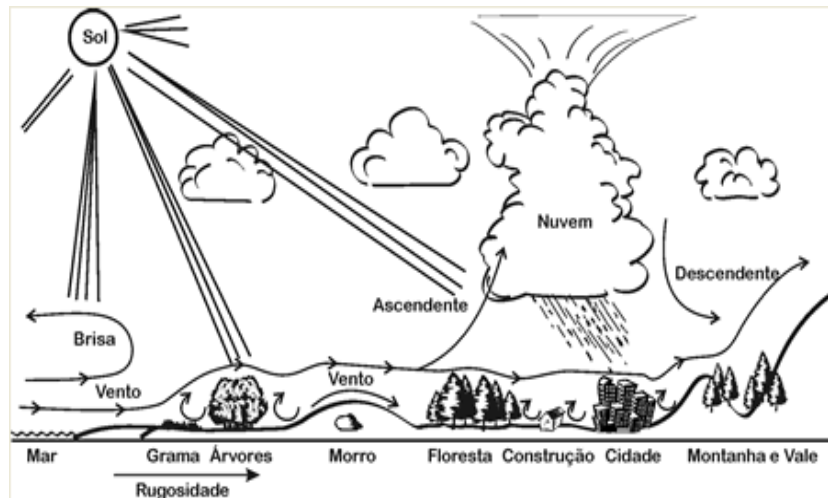


Figura 4.5: Comportamento do vento sob a influência das características do terreno.

Fonte: Atlas Eólico do Brasil, 1998.

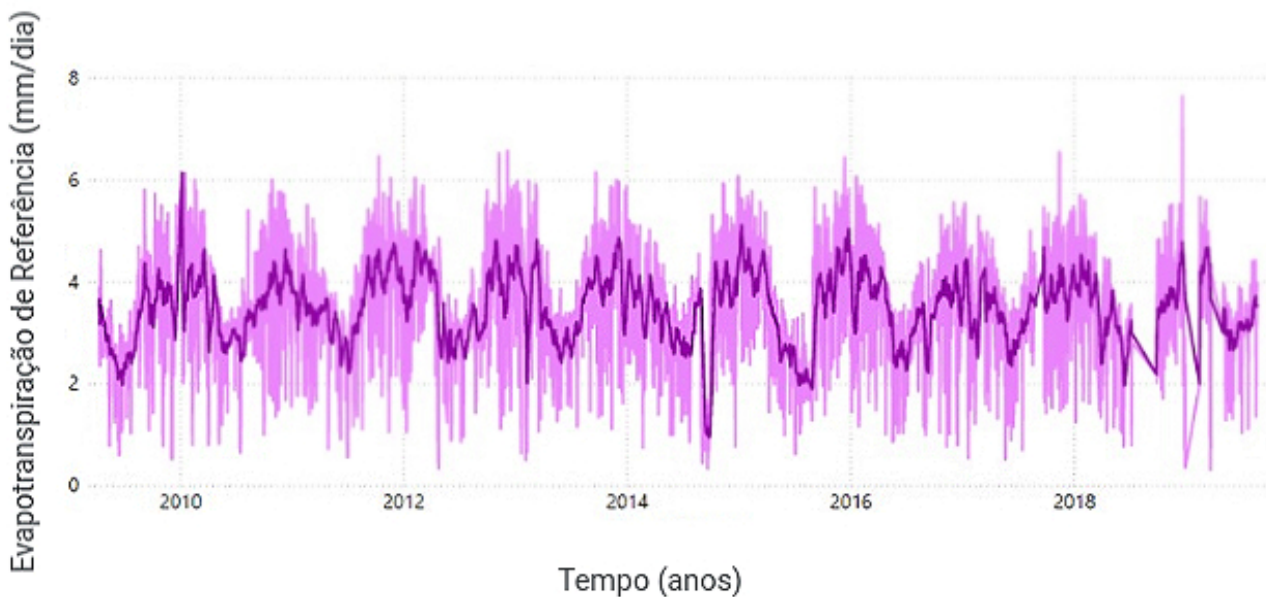


Figura 4.6: Evolução temporal da Evapotranspiração de Referência (ET_o) durante 10 anos. Em destaque a Média Móvel de 15 dias.

Temos pela figura 4.6, a evolução temporal da Evapotranspiração de Referência (ET_o), calculada através do método de Penman-Monteith (FAO) ($PM-FAO$), estabilizando em uma média de 3,3 mm/d.

A disponibilidade hídrica adequada é um dos fatores fundamentais para o desenvolvimento da vegetação, pois a falta de água pode resultar em deficiência hídrica e, por consequência, em

distúrbios fisiológicos negativos no sistema. Destaca-se, entre os principais fatores que afetam a disponibilidade hídrica de uma região, a precipitação e a evapotranspiração [96].

As mudanças climáticas indicam alterações no regime de precipitação em algumas regiões e assim interferem no aumento de temperatura no ambiente. Logo, por tais motivos, é muito importante monitorar as condições atmosféricas das regiões, pois elas contribuem diretamente para as tais variações [97].

Dessa maneira, a fim de prever a Evapotranspiração em uma certa localidade a partir da Temperatura do Ar ($^{\circ}\text{C}$) (**T**), Radiação Solar Global ($\text{MJ}/\text{m}^2.\text{d}$) (**GSR**), Umidade Relativa do Ar (%) (**RH**) e Velocidade do Vento (m/s) (**WS**), foram ajustados 45 modelos distintos, apresentados no Apêndice A. O objetivo dessa etapa foi o de encontrar o modelo de melhor qualidade para previsão fazendo uso de menos parâmetros em relação ao método de Penman-Monteith (FAO) (**PM-FAO**).

Quando os dados são observados ao longo do tempo é mais adequado dividir os dados em dois períodos, sendo os valores mais antigos os de treinamento e os mais recentes de validação. Uma vez separada a amostra em duas, utilizamos a amostra de treinamento para ajustar cada um dos modelos considerados. Então, valeu-se de cada modelo ajustado para prever a Evapotranspiração das observações na amostra de validação, obtendo assim os valores preditos por cada modelo [98].

Na tabela 4.1 são apresentados os 45 (quarenta e cinco) modelos ajustados, ordenados a partir dos melhores resultados do R^2 .

Tabela 4.1: Desempenho das técnicas utilizadas.

Legenda: Redes Neurais Artificiais (**RNA**), *Support Vector Machine* (**SVM**), *Random Forest* (**RF**), Radiação Solar Global ($\text{MJ}/\text{m}^2.\text{d}$) (**GSR**), Temperatura do Ar ($^{\circ}\text{C}$) (**T**), Umidade Relativa do Ar (%) (**RH**) e Velocidade do Vento (m/s) (**WS**).

Método	Parâmetros	R^2	RMSE	MAE
RNA	GSR + T + WS	0,9449	0,4330	0,1874
RNA	GSR + RH + WS	0,9405	0,4398	0,1934
RNA	GSR + T + RH	0,9320	0,4611	0,2126
RNA	GSR + T + RH + WS	0,9182	0,4824	0,2327
RNA	GSR + T	0,9170	0,4845	0,2347
RNA	GSR + WS	0,9112	0,4853	0,2355
RNA	GSR + RH	0,9072	0,4929	0,2430

Método	Parâmetros	R ²	RMSE	MAE
RNA	GSR	0,9025	0,5011	0,2511
SVM	GSR + T + RH + WS	0,8903	0,5310	0,2820
SVM	GSR + T + WS	0,8794	0,5443	0,2962
SVM	GSR + RH + WS	0,8642	0,5625	0,3164
SVM	GSR + T	0,8626	0,5450	0,3030
RF	GSR + T + RH	0,8571	0,5505	0,3030
RF	GSR + T + RH + WS	0,8535	0,5595	0,3131
SVM	GSR + T + RH	0,8471	0,5620	0,3158
SVM	GSR + RH	0,8420	0,5735	0,3289
SVM	GSR + WS	0,8411	0,5687	0,3234
RF	GSR + T	0,8314	0,5731	0,3284
RF	GSR + RH	0,8302	0,5756	0,3313
RF	GSR + T + WS	0,8261	0,5843	0,3414
SVM	GSR	0,8223	0,5723	0,3276
RF	GSR + RH + WS	0,8190	0,5885	0,3463
RF	GSR	0,7896	0,6019	0,3623
RF	GSR + WS	0,7853	0,6132	0,3760
SVM	T + RH + WS	0,7310	0,6205	0,3850
SVM	T + RH	0,7205	0,6305	0,3976
SVM	RH + WS	0,6857	0,6510	0,4238
RNA	T + RH + WS	0,6754	0,6530	0,4264
RF	T + RH	0,6734	0,6613	0,4373
RNA	T + RH	0,6624	0,6628	0,4393
RNA	RH	0,6577	0,6624	0,4388
SVM	RH	0,6574	0,6674	0,4454
RF	T + RH + WS	0,6528	0,6704	0,4494
RNA	RH + WS	0,6488	0,6713	0,4507
RF	RH + WS	0,5990	0,7015	0,4921
RF	RH	0,5881	0,7030	0,4943
SVM	T	0,5679	0,7161	0,5127

Método	Parâmetros	R ²	RMSE	MAE
SVM	T + WS	0,5600	0,7080	0,5012
RNA	T + WS	0,5103	0,7437	0,5531
RF	T + WS	0,4762	0,7524	0,5661
RF	T	0,4742	0,7568	0,5728
RNA	T	0,4560	0,7774	0,6044
SVM	WS	0,1036	0,8482	0,7194
RF	WS	0,0812	0,9064	0,8216
RNA	WS	0,0126	0,8850	0,7832

Tem-se na tabela 4.1 as técnicas e suas combinações correspondentes, com destaque para as 09 (nove) mais eficientes. Observa-se que a técnica de Redes Neurais Artificiais (RNA) gera resultados excelentes ao ser utilizada com diferentes combinações de três elementos que incluem o parâmetro GSR. Isso é evidenciado pelas métricas R^2 , RMSE e MAE, indicando que um número menor de variáveis é capaz de representar o sistema observado. A tabela também apresenta os resultados das outras duas técnicas utilizadas no estudo, as quais fazem uso das mesmas variáveis [99].

A utilização da técnica de RNA com a combinação dos parâmetros GSR + T + WS se destaca, apresentando métricas de $R^2 = 0,9449$, RMSE = 0,4330 e MAE = 0,1874. Esse modelo consegue representar com precisão a Evapotranspiração de Referência, mesmo sem a inclusão do parâmetro de Umidade Relativa do Ar (%). Isso ocorre devido às altas temperaturas frequentemente registradas na região, onde a RH tende a ter uma influência baixa nesse elemento climático, pois as técnicas de Inteligência Artificial (IA) conseguem identificar padrões sem a necessidade de se utilizar as duas variáveis. Os períodos de 2010 a 2012, por exemplo, mostraram resultados baixos para tal parâmetro.

A segunda melhor técnica envolve a substituição da Temperatura do Ar (°C) pela Umidade Relativa do Ar (%), ainda utilizando a RNA com três parâmetros. Essa abordagem apresenta um desempenho muito próximo ao da primeira combinação mencionada. Portanto, pode-se considerar que esse modelo é equivalente ao anterior, mesmo ao analisar períodos de baixa umidade relativa, como em casos de secas prolongadas.

A terceira combinação mais eficiente é a técnica de RNA utilizando os parâmetros GSR + T + RH. Embora seu desempenho seja ligeiramente inferior às duas primeiras combinações, vale

ressaltar que esse modelo dispensa a inclusão da **WS** em suas previsões. A análise da tabela 4.1 ainda revelou que o modelo sem a variável **WS** possui um valor pouco distinto de quando a mesma é incluída, o que indica que essa técnica pode fornecer resultados satisfatórios sem a necessidade de um anemômetro, resultando assim em uma economia na montagem da estação de medição de **ETo**.

É esperado que todas as combinações incluam o parâmetro **GSR**, considerando as características da região, com variações anuais de temperatura entre 31 °C e 34 °C. Além da possibilidade de economia com o sistema de coleta de dados, a utilização de um modelo confiável com um menor número de parâmetros pode aumentar a precisão das previsões, reduzindo o erro em relação às medições.

É importante mencionar que a Temperatura do Ar (°C) e a Umidade Relativa do Ar (%) podem ser medidas com o mesmo equipamento, um Termo-Higrômetro, sem a necessidade de custos adicionais. A inclusão do parâmetro de Velocidade do Vento (m/s) é independente dos outros parâmetros. Portanto, a técnica de **RNA** utilizando os parâmetros **GSR + T + RH** se destaca ao gerar economia na configuração da estação de medição, enquanto fornece bons resultados, com eficiência de aproximadamente 93% em comparação com o método **PM-FAO**, fazendo uso de apenas três variáveis.

A metodologia proposta também permite compreender o fluxo de energia da região e o comportamento da vegetação futura, o que possibilita o planejamento e a execução de projetos com menos equipamentos, apresentando resultados próximos ao método **PM-FAO** [99].

No entanto, é importante reconhecer que a escolha da técnica de modelagem mais adequada depende do problema em questão e da disponibilidade de dados. Embora as Redes Neurais Artificiais (**RNA**) ofereçam várias vantagens para a modelagem de sistemas, elas também possuem algumas limitações. Elas podem ser consideradas “caixas-pretas”, tornando difícil entender como chegaram a uma determinada previsão ou decisão. Além disso, se houver dados insuficientes, a **RNA** pode não apresentar resultados precisos devido à limitação no treinamento e teste.

Um grande volume de dados para um treinamento adequado é necessário, o que pode ser desafiador quando os dados são limitados ou caros para serem coletados. Além disso, sistemas com recursos limitados podem ter dificuldades computacionais ao utilizar esse método. Também é importante considerar o risco de *overfitting*, ou seja, o modelo se ajustar muito bem aos dados de treinamento, mas não generalizar bem para novos dados. Nessa pesquisa, esses aspectos não

foram problemáticos, pois havia uma grande quantidade de valores coletados ao longo de 10 (dez) anos.

Gráficos de correlação foram gerados para todas as combinações citadas na tabela 4.1, conforme mostrado na figura 4.7. O eixo horizontal representa a Evapotranspiração de Referência calculada usando o método PM-FAO e o eixo vertical representa a ET_0 prevista por cada técnica estudada. Quando um modelo prevê corretamente uma observação, o ponto cairá exatamente sobre uma linha reta que representa a igualdade entre os valores previstos e observados. Quanto mais distante o ponto estiver dessa linha, pior será a previsão do modelo para aquela observação.

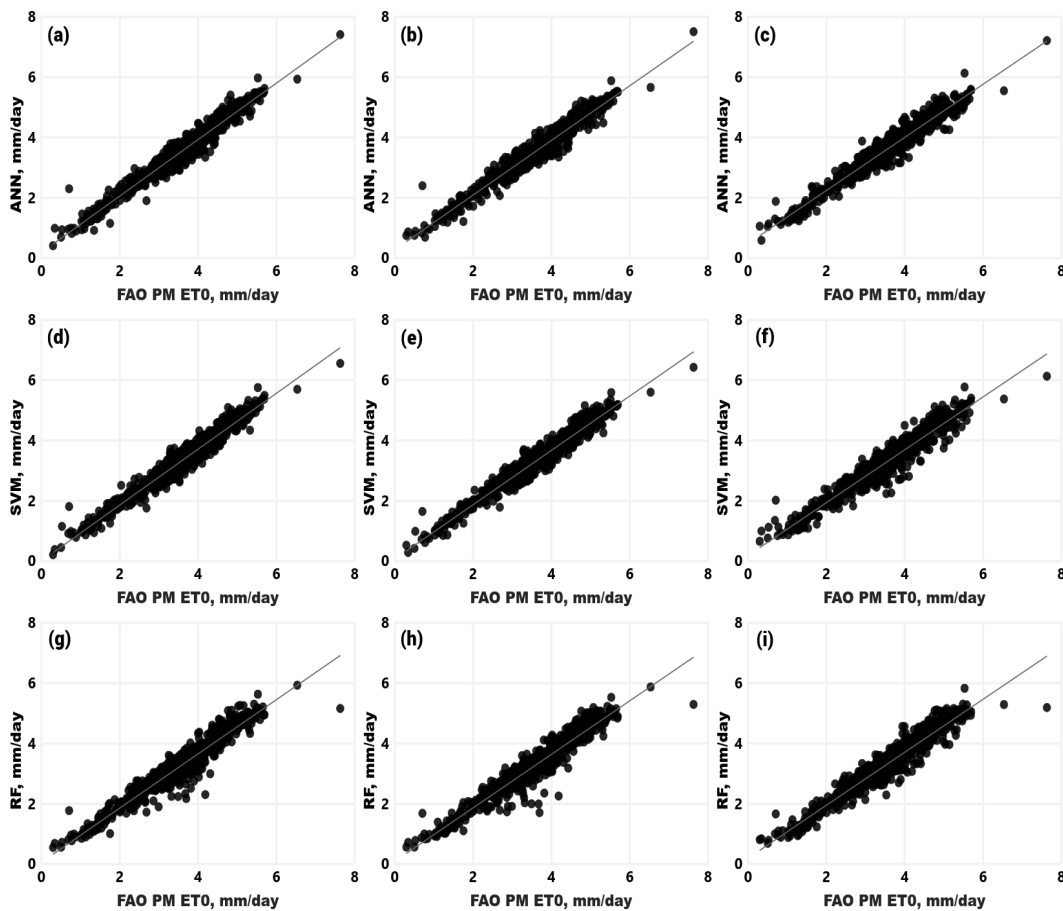


Figura 4.7: Correlação entre os parâmetros com base na técnica utilizada. (a), parâmetros GSR + T + WS, $y = 0,9356x + 0,1726$, $R^2 = 0,9450$; (b), parâmetros GSR + RH + WS, $y = 0,9059x + 0,2636$, $R^2 = 0,9405$; (c), parâmetros GSR + T + RH, $y = 0,8799x + 0,4695$, $R^2 = 0,9320$; (d), parâmetros GSR + T + WS, $y = 0,9263x - 0,0085$, $R^2 = 0,8794$; (e), parâmetros GSR + RH + WS, $y = 0,9004x + 0,0548$, $R^2 = 0,8642$; (f), parâmetros GSR + T + RH, $y = 0,8761x + 0,1715$, $R^2 = 0,8471$; (g), parâmetros GSR + T + WS, $y = 0,8987x + 0,0457$, $R^2 = 0,8261$; (h), parâmetros GSR + RH + WS, $y = 0,8872x + 0,0755$, $R^2 = 0,8190$; (i), parâmetros GSR + T + RH, $y = 0,8780x + 0,1816$, $R^2 = 0,8570$.

Diferentes pesquisas apresentam aplicações da Inteligência Artificial (IA) na agricultura, por exemplo, e indicam que essa tecnologia aliada ao campo, pode solucionar diversos problemas, inerentes a capacidade intelectual humana. Retratam os desafios e soluções no uso da

Inteligência Artificial (IA) mostrando a evolução em diferentes setores [99, 100, 101].

Capítulo 5

Conclusões

Este estudo foi fundamentado na busca por um método alternativo, eficaz e eficiente, com base em técnicas de Aprendizado de Máquina (AM), para descrever a Evapotranspiração de Referência (ET_o) de uma região de transição de uma savana brasileira, utilizando menos parâmetros em comparação com o método de Penman-Monteith (FAO) (PM-FAO). A análise das técnicas de Aprendizado de Máquina (AM) demonstrou que esses procedimentos são altamente produtivos na modelagem da Evapotranspiração de Referência (ET_o), pois conseguem processar um extenso banco de dados e encontrar a melhor interação entre os parâmetros envolvidos, com resultados superiores a 94% em comparação com o método padrão, utilizando menos variáveis.

A técnica de Redes Neurais Artificiais (RNA) se mostrou a mais eficiente entre as abordagens propostas. Consegue modelar a Evapotranspiração de Referência (ET_o) em uma região de transição do Cerrado com menos parâmetros, utilizando as variáveis microclimáticas Radiação Solar Global ($MJ/m^2.d$), Temperatura do Ar (°C) e Umidade Relativa do Ar (%) se destacando por gerar economia na configuração do sistema de medição e, ao mesmo tempo, fornecendo bons resultados, com uma precisão de aproximadamente 93% em comparação com o método PM-FAO.

A substituição do método de Penman-Monteith (FAO) (PM-FAO) se revela como uma alternativa muito eficaz, pois permite lidar com sistemas complexos e não lineares, que podem ser desafiantes de serem modelados com outros métodos. Adicionalmente, ela é capaz de lidar com ruídos e incertezas nos dados de entrada, gerando resultados razoáveis mesmo quando os dados são incompletos ou imprecisos. Além disso, ela identifica padrões que outras técnicas de modelagem matemática demorariam para encontrar.

Usar a Inteligência Artificial (IA) se apresenta como ferramenta de grande poder no tratamento de uma quantidade massiva de dados, apontando tendências que dificilmente seriam observadas por outras técnicas. Sua capacidade de processamento deve ser levada em consideração para que análises mais rápidas e precisas sejam executadas. A escolha e utilização das técnicas de *Machine Learning* (ML) deve estar diretamente alinhada com o banco de dados disponível, visto que um rol pode influenciar em possíveis falhas de processamento.

Portanto, ao admitir uma técnica de modelagem para uma aplicação específica, é essencial levar em consideração suas limitações. Pesquisas que analisem outros métodos podem ser aplicadas com o objetivo de mitigar os pontos críticos das Redes Neurais Artificiais (RNA), por meio da combinação de diferentes abordagens.

Referências Bibliográficas

- [1] D. M. P. Blank, “O contexto das mudanças climáticas e as suas vítimas,” *Mercator (Fortaleza)*, vol. 14, pp. 157–172, 2015.
- [2] J. d. R. de Souza and L. N. G. dos Reis, “Geoprocessing and landscape ecology for assessment fragmentation and connectivity of the habitats of the microregion of ceres, goiás (brazil),” *Periódico Eletrônico Fórum Ambiental da Alta Paulista*, vol. 17, no. 3, 2021.
- [3] I. A. Shiklomanov, A. I. Shiklomanov, R. B. Lammers, B. Peterson, and C. J. Vorosmarty, “The dynamics of river water inflow to the arctic ocean,” in *The freshwater budget of the Arctic Ocean*, pp. 281–296, Springer, 2000.
- [4] F. Bessat, “A mudança climática entre ciência, desafios e decisões: olhar geográfico,” *Terra Livre*, no. 20, pp. 11–26, 2003.
- [5] C. A. Nobre, G. Sampaio, and L. Salazar, “Mudanças climáticas e amazônia,” *Ciência e Cultura*, vol. 59, no. 3, pp. 22–27, 2007.
- [6] J. A. Marengo, L. Alves, M. Valverde, R. Rocha, and R. Laborbe, “Eventos extremos em cenários regionalizados de clima no brasil e américa do sul para o século xxi: Projeções de clima futuro usando três modelos regionais,” *Relatório*, vol. 5, pp. 495–516, 2007.
- [7] J. A. Marengo, C. A. Nobre, J. Tomasella, M. D. Oyama, G. S. De Oliveira, R. De Oliveira, H. Camargo, L. M. Alves, and I. F. Brown, “The drought of amazonia in 2005,” *Journal of climate*, vol. 21, no. 3, pp. 495–516, 2008.
- [8] R. G. Allen, L. S. Pereira, D. Raes, M. Smith, *et al.*, “Crop evapotranspiration-guidelines for computing crop water requirements-fao irrigation and drainage paper 56,” *Fao, Rome*, vol. 300, no. 9, p. D05109, 1998.

- [9] R. D. Jensen, M. E.; Burman and R. G. Allen, *Evapotranspiration and Irrigation Water Requirements*. ASCE Manuals and Reports on Engineering Practice No. 70, American Society of Civil Engineers, Reston, Virginia, 1990.
- [10] J. Doorenbos and W. O. Pruitt, *Guidelines for Predicting Crop Water Requirements*. FAO Irrigation and Drainage Paper 24, Food and Agriculture Organization of the United Nations, Rome, 1977.
- [11] F. d. O. Lemos, “Metodologia para seleção de métodos de previsão de demanda,” <http://hdl.handle.net/10183/5949>, 2006.
- [12] B. Tak, “A New Method for Forecasting Stock Prices Using Artificial Neural Network and Ondeleta Theory,” *Tese (Doutorado em Economia)*, 1995.
- [13] H. d. N. Camelo, P. S. Lucio, J. B. V. Leal, and P. C. M. de Carvalho, “Métodos de previsão de séries temporais e modelagem híbrida ambos aplicados em médias mensais de velocidade do vento para regiões do nordeste do brasil,” *Revista Brasileira de Meteorologia*, vol. 32, pp. 565–574, 2017.
- [14] C. F. de Angelo, R. Zwicker, N. M. M. D. Fouto, and M. R. Luppe, “Séries temporais e redes neurais: uma análise comparativa de técnicas na previsão de vendas do varejo brasileiro,” *BBR-Brazilian Business Review*, vol. 8, no. 2, pp. 1–21, 2011.
- [15] L. Vieira and F. J. Piculli, “Meteorologia e climatologia agrícola,” *Cidade Gaúcha—Pr: Universidade Estadual de Maringá*, 2009.
- [16] M. Monteith, J.L. e Unsworth, *Principles of Environmental Physics: Plants, Animals, and the Atmosphere*. Academic Press, 2008.
- [17] V. Singh, *Hydrology and Water Resources Engineering*. Elsevier Science, 2002.
- [18] R. Jensen, M.E.; Burman and R. Allen, *Evapotranspiration and Irrigation Water Requirements*. ASCE Manuals and Reports on Engineering Practice No. 70., 1990.
- [19] E. Taiz, L. e Zeiger, *Plant Physiology*. Sinauer Associates, 2010.
- [20] J. Kramer, P.J. e Boyer, *Water Relations of Plants and Soils*. Pearson, 1995.
- [21] C. Salisbury, F.B. e Ross, *Plant Physiology*. Wadsworth Publishing, 1991.

- [22] S. Gupta, H. V.; Sorooshian and P. O. Yapo, *Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration*. Journal of Hydrologic Engineering, 4(2), 135–143, 1999.
- [23] R. T. Mendonça, E. A.; Dantas, “Estimativa da evapotranspiração de referência no município de capim, pb,” *Revista Brasileira de Engenharia Agrícola e Ambiental, Campina Grande, v. 14, n. 2*, pp. 196– 202, 2008.
- [24] R. T. d. Faria, F. d. S. Campeche, and E. Y. Chibana, “Construção e calibração de lisímetros de alta precisão,” *Revista Brasileira de Engenharia Agrícola e Ambiental*, vol. 10, pp. 237–242, 2006.
- [25] L. Evett, S.R.; Heng and T. Howell, *Introduction to the Soil-Plant-Atmosphere System: Monitoring and Modeling*. CRC Press, 2012.
- [26] E. A. Mendonça and R. T. Dantas, “Estimativa da evapotranspiração de referência no município de capim, pb,” *Revista Brasileira de Engenharia Agrícola e Ambiental*, vol. 14, no. 2, pp. 196–202, 2010.
- [27] R. G. Bates and G. W. Henry, *Evaporation from Water Surfaces in Relation to Projected Water Supplies*. Transactions of the American Society of Civil Engineers, 91(1), 501–585, 1928.
- [28] P. C. Sentelhas, T. J. Gillespie, and E. A. Santos, “Evaluation of fao penman–monteith and alternative methods for estimating reference evapotranspiration with missing data in southern ontario, canada,” *Agricultural water management*, vol. 97, no. 5, pp. 635–644, 2010.
- [29] M. d. Jabloun and A. Sahli, “Evaluation of fao-56 methodology for estimating reference evapotranspiration using limited climatic data: Application to tunisia,” *Agricultural water management*, vol. 95, no. 6, pp. 707–715, 2008.
- [30] J. Cai, Y. Liu, T. Lei, and L. S. Pereira, “Estimating reference evapotranspiration with the fao penman–monteith equation using daily weather forecast messages,” *Agricultural and Forest Meteorology*, vol. 145, no. 1-2, pp. 22–35, 2007.
- [31] Z. Popova, M. Kercheva, and L. S. Pereira, “Validation of the fao methodology for computing eto with limited data. application to south bulgaria,” *Irrigation and Drainage*:

- The journal of the International Commission on Irrigation and Drainage*, vol. 55, no. 2, pp. 201–215, 2006.
- [32] R. L. C. d. Araujo, R. M. d. Silva, J. d. S. Carvalho, and I. B. Monteiro, “Influência ambiental sobre a estrutura geotermal rasa,” *Revista brasileira de Geofísica*, vol. 22, no. 1, pp. 33–44, 2004.
- [33] Z. A. Hargreaves, G. H.; & Samani, *Reference Crop Evapotranspiration from Temperature*. Applied Engineering in Agriculture, 1(2), 96–99, 1982.
- [34] C. W. Thornthwaite, *An Approach toward a Rational Classification of Climate*. Geographical Review, 38(1), 55–94, 1948.
- [35] A. S. Andrade Júnior, “Métodos de estimativa da evapotranspiração de referência diária para parnaíba e teresina, piauí,” *Revista Brasileira de Agrometeorologia, Santa Maria*, pp. 63–68, 2003.
- [36] R. C. Bassanezi, *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. São Paulo: Contexto, 2002.
- [37] A. G. Varriale, M. C.; Gomes, “Modelagem de ecossistemas - uma introdução,” *Santa Maria, RS: UFSM*, 2001.
- [38] L. Von Bertalanffy, *Teoria geral dos sistemas*. Vozes Petrópolis, 1975.
- [39] A. Christofletti, *Modelagem de sistemas ambientais*. Editora Blucher, 1999.
- [40] T. A. Spontoni and E. V. P. Spreafico, “A modelagem matemática no ensino médio,” *Dissertação PROFMAT - Campo Grande, MS: UFMS*, 2018.
- [41] R. G. Golledge, “Chorley, r.j. and haggett, p., editors 1967: Models in geography. london: Methuen,” *Progress in Human Geography*, vol. 30, no. 1, pp. 107–113, 2006.
- [42] U. Sodré, *Modelos matemáticos*. Departamento de Matemática. Universidade de Londrina. Londrina-PR., 2002.
- [43] L. F. N. VIANNA, *Métodos determinísticos ou probabilísticos de representação e análise espacial de dados para seleção de sítios em sistemas de informação geográficas: O exemplo da maricultura em Santa Catarina. Anais XIII Simpósio Brasileiro de Sensoriamento Remoto*. Florianópolis, Brasil, 21-26 abril, INPE., 2007.

- [44] S. Openshaw, Christine & Openshaw, *Artificial Intelligence in Geography*. England., 1997.
- [45] A. M. d. R. Fernandes, “Agentes inteligentes,” *Inteligência Artificial, Noções Gerais. Florianópolis: Visual Books*, pp. 85–113, 2003.
- [46] D. d. S. Gomes, “Inteligência artificial: conceitos e aplicações,” *Olhar Científico. v1*, no. 2, pp. 234–246, 2010.
- [47] K.-F. Lee, *Inteligência artificial*. Globo Livros, 2019.
- [48] G. F. Luger, *Artificial intelligence: structures and strategies for complex problem solving*. Pearson education, 2005.
- [49] E. Costa and A. Simões, *Inteligência artificial: fundamentos e aplicações*. FCA Editora, 2004.
- [50] S. Russell and P. Norvig, “Artificial intelligence: a modern approach,” *BSN: 0-13-103805-2*, 2002.
- [51] S. Russel and P. Norvig, “Inteligência artificial. 2^a,” *Edição. São Paulo: Campus*, 2004.
- [52] M. C. Pinto *et al.*, “Sistema de manutenção preditiva de falhas em válvulas em um processo industrial utilizando inteligência artificial.,” *repositório.ufu.br*, 2021.
- [53] M. A. Tolba, A.; & Sastry, *Artificial Intelligence Techniques in Environmental Monitoring: A Review*. Procedia Computer Science, 37, 388-394, 2014.
- [54] J. R. Saravanan, V.; Kim and A. Samal, *Applications of Artificial Intelligence in Biodiversity Conservation*. Environmental Science and Pollution Research, 28(27), 34962-34973, 2021.
- [55] Y. Wang, Y.; Sun and H. Wang, *Application of Artificial Intelligence Technology in Water Resources and Environmental Management: A Review*. Environmental Earth Sciences, 73(9), 4993-5006, 2015.
- [56] T. S. F. Sousa, J. A.; Silva and C. A. Silva, *Artificial Intelligence Applied to Satellite Remote Sensing for Detecting Deforestation and Forest Degradation: A Review*. Remote Sensing, 12(7), 1132, 2020.

- [57] B. M. da Silva and M. Vanderlinde, “Inteligência artificial, aprendizado de máquina,” *Rio de Janeiro: LTC*, 2010.
- [58] E. M. de Andrade, H. d. A. Q. Palácio, L. de Araújo Crisóstomo, I. H. de Souza, and A. dos Santos Teixeira, “Índice de qualidade de água, uma proposta para o vale do rio trussu, ceará,” *Revista Ciência Agronômica*, vol. 36, no. 2, pp. 135–142, 2005.
- [59] X.-D. Zhang, “Machine learning,” in *A Matrix Algebra Approach to Artificial Intelligence*, pp. 223–440, Springer, 2020.
- [60] I. El Naqa and M. J. Murphy, “What is machine learning?,” in *machine learning in radiation oncology*, pp. 3–11, Springer, 2015.
- [61] S. J. Whitty and H. Maylor, “And then came complex project management (revised),” *International Journal of Project Management*, vol. 27, no. 3, pp. 304–310, 2009.
- [62] X. Che, H.; Zhou and X. Zhang, *Machine Learning and Remote Sensing for Forest Biomass Estimation*. IEEE Geoscience and Remote Sensing Letters, 13(7), 950-954, 2016.
- [63] F. Acevedo-Whitehouse, K.; Gulland and D. Greig, *Diseases of Marine Mammals: From the Individual to the Population*. Journal of Comparative Pathology, 142(Suppl 1), S1-S2, 2010.
- [64] E. Pahikkala, T.; Tsivtsivadze and T. Salakoski, *Toward more Transparent Models of Gene Regulation*. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 9(6), 1596-1607, 2012.
- [65] D. W. Singh, M.; Lamb and E. Micijevic, *Unmanned Aerial Vehicles in Remote Sensing and Photogrammetry*. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3), 167-175, 2019.
- [66] L. FLECK, M. H. F. Tavares, E. Eyng, A. Helmann, and M. d. M. Andrade, “Redes neurais artificiais: Princípios básicos,” *Revista Eletrônica Científica Inovação e Tecnologia*, vol. 1, no. 13, pp. 47–57, 2016.
- [67] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

- [68] M. Van Otterlo and M. Wiering, “Reinforcement learning and markov decision processes,” in *Reinforcement learning*, pp. 3–42, Springer, 2012.
- [69] T. W. Rauber, “Redes neurais artificiais,” *Universidade Federal do Espírito Santo*, vol. 29, 2005.
- [70] E. P. Miguel, A. V. Rezende, F. A. Leal, E. A. T. Matricardi, A. T. d. Vale, and R. S. Pereira, “Redes neurais artificiais para a modelagem do volume de madeira e biomassa do cerrado com dados de satélite,” *Pesquisa Agropecuária Brasileira*, vol. 50, pp. 829–839, 2015.
- [71] J. S. d. Medeiros, *Bancos de dados geográficos e redes neurais artificiais: tecnologias de apoio à gestão do território*. PhD thesis, Universidade de São Paulo, 1999.
- [72] C. Spörl, E. Castro, and A. Luchiari, “Aplicação de redes neurais artificiais na construção de modelos de fragilidade ambiental,” *Revista do Departamento de Geografia*, vol. 21, pp. 113–135, 2011.
- [73] A. Cutler, D. R. Cutler, and J. R. Stevens, “Random forests,” in *Ensemble machine learning*, pp. 157–175, Springer, 2012.
- [74] L. Breiman, “Random forests,” *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [75] I. A. de Souza, T. B. Barreto, and K. P. S. O. R. Esquerre, “Modelo de interpretabilidade local pautado em random forest para classificação de usos domésticos de água,” *Revista do Seminário Internacional de Estatística com R*, vol. 4, no. 1, pp. 12–12, 2019.
- [76] S. S. d. Oliveira, “Uso de machine learning na mineração: revisão de literatura e aplicação do algoritmo random forest para otimização da recuperação mássica durante o beneficiamento de ferro,” *Repositório Institucional - Universidade Federal de Uberlândia*, 2022.
- [77] R. H. d. S. Ferreira and N. P. Anegelo, “Detecção de mudanças em dados de imagens multitemporais de sensoriamento remoto empregando support vector machines com o uso de kernel polinomial e kernel rbf (kernel função de base radial),” *Pesquisas em Geociências*, v. 45 (2018): e0674, 2018.
- [78] A. C. Lorena and A. C. de Carvalho, “Uma introdução às support vector machines,” *Revista de Informática Teórica e Aplicada*, vol. 14, no. 2, pp. 43–67, 2007.

- [79] I. Steinwart and A. Christmann, *Support vector machines*. Springer Science & Business Media, 2008.
- [80] A. Christmann and I. Steinwart, *Support vector machines*. Springer, 2008.
- [81] J. d. A. L. Vourlitis, G.L.; de Souza Nogueira, “Energy balance and canopy conductance of a tropical semi-deciduous forest of the southern amazon basin.,” *Water Resour.*, vol. 44, pp. 1–14, 2008.
- [82] T. R. Rodrigues, G. L. Vourlitis, F. d. A. Lobo, F. B. Santanna, P. H. de Arruda, and J. d. S. Nogueira, “Modeling canopy conductance under contrasting seasonal conditions for a tropical savanna ecosystem of south central mato grosso, brazil,” *Agricultural and forest meteorology*, vol. 218, pp. 218–229, 2016.
- [83] P. Alvares, C.A.; Sentelhas, “Köppen’s climate classification map for brazil.,” *Meteorol*, vol. 22, pp. 711–728, 2013.
- [84] C. B. de Barros, “Validação de métodos analíticos,” *Biológico, São Paulo*, vol. 64, no. 2, pp. 175–177, 2002.
- [85] R. S. Ehlers, “Análise de séries temporais,” *Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná*, vol. 1, pp. 1–118, 2007.
- [86] M. I. S. Bezerra, “Apostila de análise de séries temporais,” *UNESP: Curso de estatística. São Paulo, SP*, 2006.
- [87] E. C. e. a. Landau, *Geoespacialização da amplitude térmica no Brasil*. III Simpósio Internacional de Climatologia, 2009.
- [88] R. Haag, “Estimativa da distribuição espectral da radiação solar sobre o território brasileiro através de análise multiinstrumental,” *Tese*, 2012.
- [89] O. Mendes Jr and M. O. Domingues, “Introdução à eletrodinâmica atmosférica,” *Revista Brasileira de Ensino de Física*, vol. 24, pp. 3–19, 2002.
- [90] J. A. Kubby, *Adaptive optics for biological imaging*. CRC press, 2013.
- [91] F. A. M. da Silva, E. D. Assad, E. T. Steinke, and A. G. Müller, “Clima do bioma cerrado,” *Agricultura tropical: quatro décadas de inovações tecnológicas, institucionais e políticas. ALBUQUERQUE, ACS*, pp. 93–148, 2008.

- [92] A. L. Ruhoff, C. B. Saldanha, and W. Collischonn, “Análise multivariada do processo de evapotranspiração em áreas de cerrado e cana-de-açúcar,” *Repositório Institucional - FURG*, 2009.
- [93] M. E. Mata, H. L. Zanini, *et al.*, “Calor específico e densidade da polpa de cajá (*spondias lutea* l.) com diferentes concentrações de sólidos solúveis sob baixas temperaturas,” *Engenharia agrícola*, vol. 25, pp. 488–498, 2005.
- [94] A. Silva, R. S. Juvanhol, and J. S. F. Junior, “Relação entre a umidade e o poder calorífico de vegetação arbustiva do cerrado brasileiro,” in *9° Congresso Florestal Brasileiro*, vol. 1, pp. 265–269, 2022.
- [95] E. D. L. Alves and S. T. Silva, “Direção e velocidade do vento em uma floresta de transição amazônia-cerrado no norte de mato grosso, brasil,” *Boletim Goiano de Geografia*, vol. 31, no. 1, pp. 77–88, 2011.
- [96] L. C. Furquim, M. P. dos Santos, C. A. O. de Andrade, L. A. de Oliveira, and A. W. P. Evangelista, “Relação entre plantas nativas do cerrado e água,” *Científic@-Multidisciplinary Journal*, vol. 5, no. 2, pp. 146–156, 2018.
- [97] J. E. F. W. L. Lima and E. M. d. SILVA, *Análise da situação dos recursos hídricos do Cerrado com base na importância econômica e socioambiental de suas águas*. Embrapa Cerrados Planaltina, 2008.
- [98] M. C. Monard and J. A. Baranauskas, “Conceitos sobre aprendizado de máquina,” *Sistemas inteligentes-Fundamentos e aplicações*, vol. 1, no. 1, p. 32, 2003.
- [99] T. A. Spontoni, T. M. Ventura, R. S. Palácios, L. F. A. Curado, W. A. Fernandes, V. B. Capistrano, C. L. Fritzen, H. G. Pavão, and T. R. Rodrigues, “Evaluation and modelling of reference evapotranspiration using different machine learning techniques for a brazilian tropical savanna,” *Agronomy*, vol. 13, no. 8, 2023.
- [100] R. de Mesquita Pinheiro, G. I. Gadotti, R. d. C. M. Monteiro, and R. Bernardy, “Inteligência artificial na agricultura com aplicabilidade no setor sementeiro,” *Diversitas Journal*, vol. 6, no. 3, pp. 2996–3012, 2021.
- [101] M. A. Sellitto, “Inteligência artificial: uma aplicação em uma indústria de processo contínuo,” *Gestão & Produção*, vol. 9, pp. 363–376, 2002.