UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

ESTIMAÇÃO DA PRODUTIVIDADE DE SOJA A PARTIR DE MODELO AGROMETEOROLÓGICO COM BASE EM INTELIGÊNCIA ARTIFICIAL

Tatiana da Silva Santos

Engenheira Agrônoma

2023

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

ESTIMAÇÃO DA PRODUTIVIDADE DE SOJA A PARTIR DE MODELO AGROMETEOROLÓGICO COM BASE EM INTELIGÊNCIA ARTIFICIAL

Tatiana da Silva Santos

Orientador: Prof. Dr. Glauco de Souza Rolim

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutora em Agronomia (Produção vegetal).

2023

S237e

Santos, Tatiana da Silva
Estimação da produtivadade de soja a partir de modelo agrometeorológico com base em inteligência artificial / Tatiana da Silva Santos. -- Jaboticabal, 2023
89 p.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Glauco de Souza Rolim

1. Análise de cluster. 2. Aprendizado de máquina. 3. Correlação de Pearson. 4. Estádio fenológico. 5. Redes neurais. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Impacto potencial desta pesquisa

O presente trabalho utilizou de ferramentas de inteligência artificial para construir modelos matemáticos que possibilitam a estimação da produtividade de soja no Brasil, com base em uma série temporal de 30 anos de dados climáticos como temperatura, umidade de relativa do ar, velocidade do vento, radiação solar, entre outros; e também dados de produtividade de soja. A identificação da correlação da influência dos elementos climáticos nos estádios fenológicos da planta de soja, assim como a estimação da produtividade pode contribuir com o manejo estratégico da cultura bem como a prevenção de perdas de produtividade nas lavouras de soja relacionadas a elementos climáticos.

Potential impact of this research

The present work used artificial intelligence tools to build mathematical models that enable the estimation of soybean productivity in Brazil, based on a 30-year time series of climate data such as temperature, relative humidity, wind speed, radiation solar, among others; and also soybean productivity data. The identification of the correlation of the influence of climatic elements on the phenological stages of the soybean plant, as well as the estimation of productivity, can contribute to the strategic management of the crop as well as the prevention of productivity losses in soybean crops related to climatic elements.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: ESTIMAÇÃO DA PRODUTIVIDADE DE SOJA A PARTIR DE MODELO AGROMETEOROLÓGICO COM BASE EM INTELIGÊNCIA ARTIFICIAL

AUTORA: TATIANA DA SILVA SANTOS
ORIENTADOR: GLAUCO DE SOUZA ROLIM

Aprovada como parte das exigências para obtenção do Título de Doutora em Agronomia (Produção Vegetal), pela Comissão Examinadora:

Prof. Dr. GLAUCO DE SOUZA ROLIM (Participaçao Virtual)
Departamento de Engenharia e Ciencias Exatas DECEx / FCAV UNESP Jaboticabal

Dra. KAMILA CUNHA DE MENESES (Participaçao Virtual)
Autônoma / Agrodimensão Soluções Agropecuárias - Chapadinha/MA

gov.br Documento assinado digitalmente
KAMILA CUNHA DE MENESES
Data: 27/06/2023 08:18:56-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. RICARDO ANTONIO FERREIRA RODRIGUES (Participaçao Virtual)
Departamento de Fitossanidade, Engenharia Rural e Solos / FEIS UNESP Ilha Solt

gov.br Documento assinado digitalmente
RICARDO ANTONIO FERREIRA RODRI
Data: 27/06/2023 16:19:13-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. FRANCIELE MORLIN CARNEIRO (Participaçao Virtual)
Universidade Tecnológica Federal do Paraná / UTFPR - Curitiba/PR

gov.br Documento assinado digitalmente
FRANCIELE MORLIN CARNEIRO
Data: 05/07/2023 09:37:47-0300
Verifique em https://validar.iti.gov.br

Profa. Dra. TATIANA FERNANDA CANATA (Participaçao Virtual)
Departamento de Engenharia e Ciências Exatas / FCAV - UNESP Jaboticabal

Jaboticabal, 23 de junho de 2023

DADOS CURRICULARES DA AUTORA

Tatiana da Silva Santos – Filha de Josenir da Silva Santos e Antonio Alberto dos Santos, nascida em Açailândia no Estado do Maranhão, em 05 de outubro de 1993. Cursou Agronomia pela Universidade Federal do Maranhão – Campus IV, Chapadinha – MA, no período de 2011 a 2016. Durante a graduação participou do Grupo de Pesquisa e Estudos de Forragicultura e Pastagens no Maranhão - FOPAMA (2011 - 2012), também participou do Grupo de Pesquisa em Agricultura de Precisão (2015 - 2016), e foi monitora da disciplina de Química Orgânica. Foi bolsista da Fundação de Amparo à Pesquisa no Maranhão (FAPEMA), na modalidade de Iniciação Científica (2012 – 2013) e na Universidade Federal Rural do Pernambuco (outubro a dezembro de 2015). Realizou mestrado na Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), no Campus de Ilha Solteira – SP, no período de 2016 a 2018, na qual obteve o título de Mestra em Agronomia (Sistemas de Produção), sob orientação do prof. Dr. Ricardo Antonio Ferreira Rodrigues. Em agosto de 2019, ingressou no curso de Doutorado do Programa de Pós-graduação em Produção Vegetal da UNESP/FCAV, Campus de Jaboticabal – SP, sob orientação do Prof. Dr. Glauco de Souza Rolim, atuando em pesquisas sobre Agrometeorologia e Modelagem Agrícola. É integrante do Group of Agrometeorological Studies (GAS). Durante o doutorado realizou estágio docência nas disciplinas de Física e Agrometeorologia, nos cursos de graduação em Biologia e Agronomia, respectivamente, da UNESP/FCAV. Em dezembro de 2021, ingressou no mercado privado, atuando como Analista de Sinistros Agrícolas, cargo em que atua até o momento.

"Importante não é ver o que ninguém nunca viu, mas sim, pensar o que ninguém nunca pensou sobre algo que todo mundo vê."

Arthur Schopenhauer

Aos meus amados pais que sempre fizeram de tudo para que eu tivesse acesso aos estudos.

AGRADECIMENTOS

À Deus, por sempre estar ao meu lado, guiando os meus passos e me protegendo de todo o mal.

Aos meus pais Josenir da Silva Santos e Antonio Alberto dos Santos, que são exemplos de sabedoria, esforço e generosidade. Por sempre me incentivarem nos estudos e pelo grande esforço que fizeram para que eu pudesse ter acesso à educação.

À minha irmã Vanessa da Silva Santos, pelo carinho, amor e apoio em todos os momentos.

Ao meu noivo João Francisco Silva Trentin, pelo incentivo, compreensão, amor e companheirismo. Por acreditar em mim e ser o meu porto seguro.

Ao meu orientador Prof. Dr. Glauco de Souza Rolim, pelos ensinamentos, atenção e paciência.

A todos os professores responsáveis pela minha formação.

Aos membros do Group of Agrometeorological Studies (GAS), pela troca de conhecimento e trabalhos desenvolvidos.

À Universidade Estadual Paulista “Júlio de Mesquita Filho” – Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal pela oportunidade de realizar o Doutorado.

Ao Programa de Pós-graduação em Agronomia (Produção Vegetal) e a todos os funcionários do departamento de Ciências Exatas e Engenharia.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

SUMÁRIO

Página

RESUMO iii
ABSTRACT iv
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS 1
1.1 O cultivo de soja no Brasil 1
1.2 Ecofisiologia da soja e classificação de cultivares por região edafoclimática 4
1.3 Elementos climáticos e a produtividade da soja 7
1.4 Aprendizado de máquina aplicado a agricultura 9
1.5 Objetivo geral 10
Referências 11
CAPÍTULO 2 – INFLUÊNCIA DOS ATRIBUTOS CLIMÁTICOS NA PRODUTIVIDADE DA SOJA NO BRASIL 17
Resumo 17
Abstract 18
2.1 Introdução 19
2.2 Material e Métodos 21
2.2.1 Área de estudo e dados de produtividade 21
2.2.2 Fonte de dados meteorológicos 22
2.2.3 Determinação do período de semeadura e colheita 24
2.2.4 Classificação das regiões 25
2.2.5 Análise estatística 27
2.3 Resultados e Discussão 27
2.4 Conclusões 42
Referências 44
CAPÍTULO 3 – VARIÁVEIS CLIMÁTICAS PARA PREDIÇÃO DE PRODUTIVIDADE DE SOJA UTILIZANDO APRENDIZADO DE MÁQUINA 49
Resumo 49
Abstract 50
3.1 Introdução 51
3.2 Material e Métodos 53
3.2.1 Áreas de estudo e produtividade 53
3.2.2 Coleta de dados climáticos 58
3.2.3 Evapotranspiração potencial 58
3.2.4 Balanço hídrico climatológico 59
3.2.5 Agrupamento dos locais 60

3.2.6 Parâmetros do cultivo 63
3.2.7 Calibração dos modelos 63
3.2.8 Avaliação de desempenho dos modelos 66
3.3 Resultados e Discussão 67
3.4 Conclusão 80
Referências 82
CAPITULO 4. CONSIDERAÇÕES FINAIS 88

ESTIMAÇÃO DA PRODUTIVIDADE DE SOJA A PARTIR DE MODELO AGROMETEOROLÓGICO COM BASE EM INTELIGÊNCIA ARTIFICIAL

RESUMO – O Brasil é o maior produtor de soja do mundo e o clima é um dos principais fatores que influenciam a produtividade da soja. Identificar os elementos meteorológicos que influenciam na produtividade da soja, assim como a utilização de modelos de predição de produtividade podem auxiliar os produtores nas tomadas de decisões e estratégias de gerenciamento do campo. Nesse contexto, objetivo deste trabalho foi: a) identificar a influência dos fatores meteorológicos em cada estádio fenológico da soja, b) identificar qual melhor algoritmo de aprendizagem de máquina (ML) para estimação e predição de produtividade de soja no Brasil com base em dados climáticos. Tanto os dados de produtividade quanto os dados meteorológicos são de uma série histórica de 30 anos (1988 a 2018). Os dados metrológicos são provenientes do NASA-POWER, foram coletados: precipitação (mm), temperatura média, máxima e mínima do ar (ºC), velocidade do vento média, máxima e mínima (m s^{-1}), umidade relativa (%), radiação solar no topo da atmosfera (MJ m^2 dia^{-1}) e radiação solar global (MJ m^2 dia^{-1}). Para o objetivo “a” foi realizado analise de correlação de Pearson, resultando em 4 grupos determinados de acordo com a sua similaridade climática, e para o objetivo “b” foi utilizado o método de modelagem tradicional de regressão linear múltipla (RLM) e três métodos de aprendizado de máquina (ML): Random Forest (RF) Support vector machine (SVM) e Multilayer perceptron. Para avaliação dos modelos foi utilizado o R^2 ajustado (R^2 adj), erro quadrático médio (RMSE) e erro médio absoluto (MAE). Os dados foram separados em 70% para treino e 30% para teste. Todas as análises foram realizadas no software Python 3.8. Os fatores meteorológicos influenciaram a produtividade de formas diferente de acordo com o estádio fenológico da cultura. O RF apresentou o melhor desempenho para o grupo 4, com R^2 adj de 0,76 e RMSE de 389,41 kg ha^{-1}. O RLM apresentou o menor desempenho para grupo 4, com R^2 adj de 0,01 e RMSE de 3993,8kg ha^{-1}. Portanto, com os resultados do estudo concluímos que é possível estimar e predizer a produtividade de soja utilizando algoritmos de aprendizado de máquina. O RF possui precisão e acurácia para predizer a produtividade de soja no Brasil. Os fatores meteorológicos influenciam a produtividade da soja de diferentes formas em cada estádio fenológico da cultura.

Palavras-chave: Análise de cluster, aprendizado de máquina, correlação de Pearson, redes neurais, estádio fenológico

ESTIMATION OF SOYBEAN PRODUCTIVITY FROM AN AGROMETEOROLOGICAL MODEL BASED ON ARTIFICIAL INTELLIGENCE

ABSTRACT – Brazil is the largest soybean producer in the world and climate is one of the main factors that influence soybean productivity. Identifying the meteorological elements that influence soybean productivity, as well as the use of productivity prediction models can help producers in decision-making and field management strategies. In this context, the objective of this work was: a) to identify the influence of meteorological factors on each soybean phenological stage, b) to identify the best machine learning (ML) algorithm for estimation and predicting soybean productivity in Brazil based on climate data. Both productivity data and meteorological data are from a 30-year historical series (1988 to 2018). The metrological data comes from NASA-POWER and was collected: precipitation (mm), average, maximum and minimum air temperature (ºC), average, maximum and minimum wind speed (m s^{-1}), relative humidity (%), solar radiation at the top of the atmosphere (MJ m^2 day^{-1}) and global solar radiation (MJ m^2 day^{-1}). For objective “a” a Pearson correlation analysis was carried out, resulting in 4 groups determined according to their climatic similarity, and for objective “b” the traditional modeling method of multiple linear regression (RLM) was used and three machine learning (ML) methods: Random Forest (RF) Support vector machine (SVM) and Multilayer perceptron. To evaluate the models, adjusted R^2 (R^2 adj), mean squared error (RMSE) and mean absolute error (MAE) were used. The data was separated into 70% for training and 30% for testing. All analyzes were performed using Python 3.8 software. Meteorological factors influenced productivity differently according to the phenological stage of the crop. RF presented the best performance for group 4, with R^2 adj of 0.76 and RMSE of 389.41 kg ha^{-1}. The RLM presented the lowest performance for group 4, with R^2 adj of 0.01 and RMSE of 3993.8kg ha^{-1}. Therefore, with the results of the study we conclude that it is possible to estimation and predict soybean productivity using machine learning algorithms. The RF has precision and accuracy to predict soybean productivity in Brazil. Meteorological factors influence soybean productivity in different ways at each phenological stage of the crop.

Keywords: Cluster analysis, machine learning, Pearson correlation, neural networks, phenological stage

CAPITULO 1. CONSIDERAÇÕES GERAIS

1. INTRODUÇÃO E JUSTIFICATIVA

1.1 O cultivo de soja no Brasil

A soja (*Glycine max* L.) é uma leguminosa (Fabacea), originária do leste da Ásia e domesticada no nordeste da Ásia, a sua disseminação do Oriente para o Ocidente ocorreu através das navegações. A recomendação mais antiga de cultivo de soja data o ano de 207 a.C., entretanto, há relatos na literatura chinesa sobre a sua utilização como alimento antes mesmos dos registros serem feitos, indicando que talvez seja uma das espécies mais antigas cultivadas pelo homem (Morse, 1950; Hymowitz, 1970).

No Brasil, a soja foi introduzida em 1882, no estado da Bahia, onde foram realizados os primeiros testes de variedades trazidas dos Estados Unidos (D'utra, 1882). A partir de então, foram realizados diversos estudos em várias regiões do Brasil, sendo estes essenciais para o estabelecimento da cultura no país. Em 1892 foi realizado o primeiro estudo da soja no estado de São Paulo, no Instituto Agronômico de Campinas (IAC), entretanto, somente em 1945 foi identificado o primeiro registro de produção da cultura nas estatísticas oficias do Estado (Daffe, 1892; Miyasaka e Medina, 1981).

Posteriormente, em 1901 foi relatado o primeiro resultado de estudo da soja no Rio Grande do Sul, no município de Dom Pedrito e somente em 1941 foi registrada a primeira produção da cultura nas estatísticas oficiais do Estado (produção de 450 toneladas e rendimento de 700 Kg ha^{-1}), sendo este o Estado onde melhor se adaptaram as variedades trazidas dos Estados Unidos devido as condições climáticas serem similares as da região de origem dos materiais (Minsen, 1901; Vernetti, 1977; Bonetti, 1981).

Até 1950 a produção de soja no Brasil era pequena, destinada a alimentação animal como forragem ou grão. Foi a partir do final de 1960 que a soja começou a ganhar espaço como um produto comercial, na época a soja passou a ser uma opção de cultivo de verão subsequente ao trigo na região Sul

do país, além disso, havia a necessidade de suprir a demanda interna do país por farelo de soja para alimentação de aves e suínos (Federizzi, 2005).

Na década de 1970 a valorização do preço do produto no mercado mundial aumentou o interesse dos produtores na cultura e impulsionou os incentivos do governo brasileiro no desenvolvimento de pesquisas e melhoramento genético do grão. Foi nesse período que a cultura mais evolui, consolidando-se no país como uma das principais culturas do agronegócio (Federizzi, 2005; EMBRAPA, 2007).

Até o final de 1970 aproximadamente mais de 80% da produção de soja ainda se concentrava na região sul do Brasil. A partir de 1980 houve a expansão do cultivo para regiões do Centro-Oeste, Nordeste e Norte do país (Figura 1). Um dos principais fatores que proporcionaram essa expansão foi o desenvolvimento de cultivares adequadas as condições edafoclimáticas de regiões de clima tropical, sendo assim possível abrir novas fronteiras agrícolas no Brasil (Sanches et al., 2007).

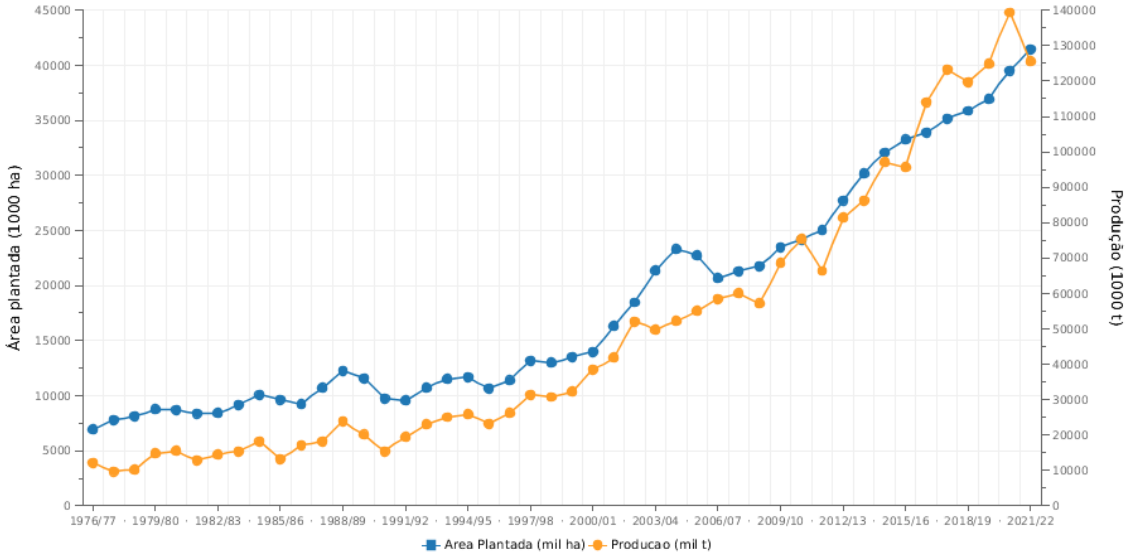

Figura 1. Área cultivada e produção de soja no Brasil.

Fonte: CONAB, 2023a.

Atualmente o Brasil é o maior produtor de soja do mundo (USDA, 2023). Na safra 2021/2022 a área cultivada de soja no Brasil correspondeu a 41.492 milhões de hectares, com uma produção de 125.549,8 milhões de toneladas e produtividade de 3.026 Kg ha^{-1}. Já na safra 2022/2023 houve um incremento de

4,9% na área cultivada em relação à safra anterior, totalizando uma área cultivada de 43.529,9 milhões de hectares, com estimativa de produção de 151.419,4 milhões de toneladas e produtividade de 3.479 Kg ha^{-1}. Destaque nesta safra para a região Centro-Oeste com a maior área cultivada 20.202,6 milhões de hectares, seguida da região Sul de 13.080,7 milhões de hectares (Tabela 1) (CONAB, 2023b).

Tabela 1. Área cultivada de soja no Brasil safras 2021/22 e 22/23, e estimativa de produção para a safra 2022/2023.

Região/UF	Área (Em mil ha)		Produção (Em mil t)	
	Safra 2021/2022	Safra 2022/2023	Safra 2021/2022	Safra 2022/2023
Norte	**2.577,0**	**2.850,9**	**8.379,9**	**9.410,0**
RR	95,0	120,0	285,0	360,0
RO	491,7	497,1	1.668,8	1.688,2
AC	6,1	11,4	20,4	37,8
AM	4,5	6,9	13,5	24,8
AP	6,5	6,5	17,2	17,6
PA	828,5	939,5	2.497,9	2.877,7
TO	1.144,7	1.269,5	3.877,1	4.403,9
Nordeste	**3.821,3**	**3.962,6**	**13.876,9**	**14.559,1**
MA	1.075,1	1.111,7	3.573,6	3.744,2
PI	850,7	926,4	3.014,0	3.310,0
CE	-	2,1	-	7,1
AL	2,3	2,7	6,2	9,1
BA	1.893,2	1.919,7	7.283,1	7.488,7
Centro-Oeste	**19.140,9**	**20.202,6**	**68.126,0**	**75.266,4**
MT	11.108,5	11.808,3	41.490,2	43.903,3
MS	3.554,6	3.760,8	8.932,7	13.764,5
GO	4.393,6	4.547,4	17.389,9	17.280,1
DF	84,2	86,1	313,2	318,5
Sudeste	**3.198,4**	**3.433,1**	**11.767,0**	**13.113,5**
MG	1.982,9	2.165,3	7.590,5	8.321,2
SP	1.215,5	1.267,8	4.176,5	4.792,3
Sul	**12.754,4**	**13.080,7**	**23.400,0**	**39.070,4**
PR	5.668,8	5.787,8	12.250,3	21.293,3
SC	727,6	737,8	2.038,7	2.588,9
RS	6.358,0	6.555,1	9.111,0	15.188,2
Brasil	**41.492,0**	**43.529,9**	**125.549,8**	**151.419,4**

Fonte: CONAB, estimativa março 2023.

A soja possui grande importância para o mercado brasileiro de grãos, sua produção representa 46% da produção brasileira de grãos (CONAB, 2023b). Pode ser comercializada na forma de grão, farelo e óleo. A soja é responsável por aproximadamente 29% da oferta mundial de óleo vegetal de consumo

(SoyStats, 2021). O complexo da soja gera retorno financeiro para o país por meio da exportação, fornecimento de matéria-prima para as indústrias e produção de alimentos (Costa e Santana, 2018).

A demanda de consumo interno do grão no país para a safra 2022/2023 é de 56.635,4 milhões de toneladas, no âmbito externo, a demanda para exportação é de 92.990 milhões de toneladas (CONAB, 2023b). A crescente demanda do grão tanto de forma interna quanto externa evidencia a sua importância em diversos setores como o social e financeiro. Dessa forma, é essencial que seja realizado estudos para melhor compreensão do comportamento da planta e as suas respostas a produtividade.

1.2 Ecofisiologia da soja e classificação de cultivares por região edafoclimática

A soja é uma planta dicotiledônea, da família Fabaceae, possui metabolismo fotossintético tipo C_3. Possui ciclo anual e suas características morfológicas sofrem influência das condições edafoclimáticas, por exemplo, o ciclo vegetativo da cultura pode variar entre 70 (cultivares precoces) a 200 (cultivares tardias) dias, assim como a altura que também pode variar de 30 a 200 cm (Sediyama et al., 2015; Ferrari et al., 2015).

O sistema radicular é constituído de raiz principal e secundária, o caule é herbáceo, ereto e o seu desenvolvimento é iniciado a partir do eixo embrionário. As cultivares podem apresentar três tipos de crescimento, sendo eles: determinado (com racemo terminal), semideterminado (intermediário) ou indeterminado (sem racemo terminal) (Silva, 2018).

As folhas são trifoliadas, possui flores de fecundação autógama e podem ser na cor branca ou púrpura, se desenvolvem em racemos terminais ou axilares. Seus frutos são vagens (legumes) que podem ser na cor verde, marrom ou cinza, apresentam de 2 a 7 cm de comprimento e 1 a 2 cm de largura, podem conter de um a cinco grãos por vagem. A soja é uma planta sensível ao fotoperíodo, conhecida por ser uma planta de dias curtos (EMBRAPA, 2020).

Em razão da sensibilidade da soja ao fotoperíodo e a adaptabilidade das cultivares a latitude, foi desenvolvido pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) - Soja e pesquisadores de diversas instituições uma

classificação de grupos de maturidade da soja que considera a diversidade de ecossistemas, tipos de solo e clima (latitude e altitude). Dessa forma, foram estabelecidas cinco macrorregiões sojícolas e vinte regiões edafoclimáticas (Tabela 2).

Tabela 2. Relação das regiões fisiográficas componentes das regiões edafoclimáticas de adaptação das cultivares de soja no Brasil.

Macrorregião sojícola	Região edafoclimática	UF	Região fisiográfica
Macrorregião 1 Sul	Região 101	RS	Campanha
			Depressão Central
			Baixo Vale do Uruguai
			Litoral
			Serra do Sudeste
	Região 102	RS	Missões
			Planalto Médio
			Alto Vale do Uruguai - Leste
			Alto Vale do Uruguai - Oeste
		SC	Oeste
			Meio-Oeste
			Nordeste
		PR	Sudoeste
	Região 103	RS	Serra do Nordeste
			Planalto Superior
		SC	Centro-Norte
			Serra Geral
		PR	Centro-Sul
		SP	Sul
	Região 104	SC	Litoral e Vale do Itajaí
Macrorregião 2 Centro-Sul	Região 201	PR	Oeste
			Norte

		SP	Médio Paranapanema
	Região 202	PR	Noroeste
		SP	Sudoeste
		MS	Sul
	Região 203	SP	Centro-Sul
			Oeste
	Região 204	MS	Centro-Sul
			Sudoeste
Macrorregião 3 Sudeste	Região 301	MS	Centro-Norte
		GO	Sudoeste
	Região 302	SP	Norte
		MG	Vale do Rio Grande
		GO	Sul
	Região 303	MG	Triângulo e Alto Paranaíba
		GO	Sudeste
	Região 304	MG	Noroeste
		GO	Leste
		DF	----------
Macrorregião 4 Centro-Oeste	Região 401	GO	Centro
		MT	Sul
	Região 402	MT	Centro-Norte
			Oeste
		RO	Sul
	Região 403	MT	Leste
	Região 404	GO	Norte
		TO	Sul (Gurupi)
	Região 405	BA	Oeste
Macrorregião 5	Região 501	PI	Sudoeste
		MA	Sul
		TO	Centro-Norte (Pedro Afonso)
		PA	Sudeste (Redenção)
	Região 502	MA	Leste (Chapadinha)
		PA	Nordeste (Paragominas)

		PA	Oeste (Santarém)
	Região 503	RR	Cerrado

Fonte: Kaster e Farias (2012).

1.3 Elementos climáticos e a produtividade da soja

As condições climáticas têm relação direta com a capacidade produtiva da soja. A planta é extremamente sensível ao estresse climático, por isso, condições de clima adequadas são essenciais para atingir alta produtividade da cultura. Os principais fatores relacionados a produtividade são: disponibilidade hídrica, temperatura e altitude (Sentelhas et al., 2015).

A alta variabilidade climática entre as regiões do Brasil impacta nos resultados de produtividade da cultura. Em regiões com altitude de até 400 m a baixa distribuição hídrica somada a temperaturas elevadas pode reduzir a produtividade da soja em até 37%. Já em regiões onde a altitude é acima de 800 m, as temperaturas tendem a ser amenas e a distribuição de chuvas mais regulares, nesse caso esses fatores podem reduzir a produtividade da soja em até 3% (Gonçalves, 2019).

O déficit hídrico combinado com altas temperaturas afeta negativamente o metabolismo das plantas de soja (Vital et al., 2022). O déficit hídrico é responsável por 74% das perdas de produtividade da cultura no Brasil e entre regiões essas perdas representam em média 46% (Sentelhas, 2015; Battisti et al., 2018). Os danos causados a cultura podem incluir o abortamento de flores, vagens e grãos pequenos, além da antecipação do ciclo da cultura, o que pode resultar em uma segunda floração normalmente infértil (Schapaugh e Roozeboom, 2012; Battisti, 2017).

Na região Sul do país, a severa redução das precipitações ocorridas entre os meses de novembro e dezembro de 2021 resultou na redução de 9,9% da produção nacional de grãos de soja na safra 2021/2022, em relação à safra anterior (CONAB, 2022c). O déficit hídrico na região Sul do Brasil é o principal limitante da produtividade de soja (Zanon, 2016). No contexto global, quebras de produção nas principais regiões produtoras podem afetar o preço da commoditie, além de causar instabilidade sociopolítica regional (Puma, 2019).

Em contra partida o excesso de chuvas também pode interferir em diversos processos metabólicos das plantas de soja (Kisman, 2023; Razhika e Cahyoadi, 2023). Assim como proporcionar o aumento da disseminação de doenças e consequentemente perdas de produtividade principalmente quando o evento ocorre próximo a colheita (Hill et al., 2013).

A temperatura do solo e do ar também são fatores importantes para o desenvolvimento da cultura (Latef et al., 2016). Baixas temperaturas podem limitar os processos fisiológicos, na soja esses efeitos refletem principalmente no metabolismo de proteínas e membranas (Nouri et al., 2011). A semeadura em solo com temperaturas menores que 20°C podem prejudicar a germinação e emergência das plântulas (França-Neto et al., 2016).

A ocorrência de baixas temperaturas no final do estádio reprodutivo associado a alta umidade do ar pode provocar atraso da colheita (Schapaugh e Roozeboom, 2012). Temperaturas abaixo de 13°C podem inibir ou retardar o florescimento, já em temperaturas elevadas acontece o inverso, ocorre a indução do florescimento precoce (Farias et al., 2007).

A eficiência do uso de radiação pode ser usada para melhorar o potencial produtivo de grãos de soja e absorção de nutrientes (Lopez, 2022; Zhou et al., 2019). Características como arquitetura da planta e dossel influenciam na absorção de luz (Pelech et al., 2023), durante o período reprodutivo a maior interceptação de luz contribui para o aumento da produtividade (Koester et al., 2014).

O fotoperíodo é um dos fatores que mais influenciam a adaptação local das variedades e consequentemente a produtividade (Lin et al., 2021). Esse fator desempenha papel importante no controle de floração da soja, o atraso no tempo de floração devido o fotoperíodo mais longo pode atrasar a maturidade fisiológica das plantas expondo a cultura a intemperes climáticos (Ort et al., 2022).

A ação dos fatores climáticos pode levar a danos irreversíveis na planta (Chaves e Oliveira, 2004). Staniak et al., (2023) destacam que condições climáticas adversas são responsáveis pela redução de aproximadamente 30% da produtividade da soja. Dessa forma, estudar as condições climáticas e seus

impactos sob a produtividade das culturas é essencial em tempos onde há cada vez mais instabilidade climática (Sentelhas et al., 2016; Edreia et al., 2017).

1.4 Aprendizado de máquina aplicado a agricultura

A agricultura está em constante evolução, novas tecnologias surgem a cada dia e diversas são as suas aplicações na agricultura. Entre os avanços ocorridos nos últimos anos destaca-se a utilização de técnicas de Aprendizado de Máquina (ML) como uma poderosa forma de compreender dados complexos dos processos agrícolas (Benos et al., 2021). O ML pode ser descrito como processos matemáticos interativos autônomos, e podem ser utilizados para classificação e previsão (Chlingaryan et al., 2018).

A estimação da produtividade é um dos principais desafios da agricultura moderna. Entender a relação entre os fatores climáticos, genéticos e manejo exige a demanda de algoritmos poderosos, como técnicas de ML (Van et al., 2020). Na literatura, diversos estudos comprovam a eficácia da utilização de ML para estimação e predição de produtividade de culturas agrícolas (Charoen e Mittrapiyanuruk, 2019; Desloires et al., 2023; Ren et al., 2023).

Nesse âmbito, há diversos algoritmos de ML de regressão, por exemplo, Random Forest (RF), Support Vector Machine (SVM) e Multilayer Perceptron (MLP). De acordo com Mohammed et al., (2022) modelos de MLs são capazes de prever secas agrícolas com maior precisão. Elbeltagi et al., (2023) observaram que o modelo RF prevê com maior precisão a complexidade de um conjunto de dados climáticos.

O SVM quando aplicado a dados de séries temporais possui alta capacidade de previsão (Kuwata e Shibasaki, 2016). Esse método é amplamente utilizado na pesquisa agrícola para classificação (Pisner e Schnyer, 2020). Além disso, Tian et al., (2018) concluíram que o SVM quando aplicado a índices climáticos pode melhorar a precisão na previsão da seca, mostrando-se eficaz nessa aplicação.

Os algoritmos RF e SVM tem apresentado resultados de alta precisão para predição de produtividade das culturas, esses algoritmos não sofrem influência da colinearidade e da distribuição não normal das variáveis, além disso

podem lidar com overfitting (Kayad et al., 2019). Os mesmos autores relatam que o RF foi o algoritmo com maior precisão para prever a variabilidade da produtividade do milho.

Morelli et al., (2021) na busca para identificar o melhor algoritmo de previsão de produtividade utilizando técnicas de MLs, constataram que os algoritmos MLP e RF apresentaram os melhores desempenhos, possibilitando seu uso como uma técnica que auxilia na previsão da produtividade do algodão e tomadas de decisão que podem auxiliar o agricultor nos processos agrícolas. Gniewko et al., (2022) também usaram uma rede neural artificial (ANN), graus-dia de crescimento e precipitação total para estimar a produtividade da soja, e obtiveram bons resultados.

Considerando a importância da cultura da soja no Brasil, seu impacto no mercado mundial e a influência dos fatores climáticos na produtividade, torna-se essencial a aplicação de técnicas que auxiliem o agricultor de forma estratégica nas tomadas de decisões. Nesse sentido, o estudo da aplicação de técnicas de ML para estimativa de produtividade de soja no Brasil com base em dados climáticos pode oferecer resultados mais precisos sobre a perspectiva de produtividade principalmente para mercados futuros. Além disso, essa abordagem pode otimizar o tempo relacionado ao emprego de estratégias de gerenciamento de manejo da cultura, bem como na comercialização.

1.5 Objetivo geral

Estimar a produtividade de soja no Brasil, utilizando modelos de Aprendizado de Máquina com base em dados climáticos.

Referências

Battisti R, Sentelhas PC, Boote KJ, Câmara GMS, Farias JRB, Basso CJ (2017). Assessment of soybean yield with altered water-related genetic improvement traits under climate change in Southern Brazil. **European Journal of Agronomy**, 83:1–14. DOI: https://doi.org/10.1016/j.eja.2016.11.004

Battisti R, Sentelhas PC, Pascoalino JAL, Sako H, Dantas JPS, Moraes MF (2018) Soybean Yield Gap in the Areas of Yield Contest in Brazil. **International Journal of Plant Production** 12: 159–168. DOI: https://doi.org/10.1007/s42106-018-0016-0

Benos L, Tagarakis AC, Dolias G, Berruto R, Kateris D, Bochtis D (2021) Machine Learning in Agriculture: A Comprehensive Updated Review. **Sensors** 21:3758. DOI: https://doi.org/10.3390/s21113758

Bonetti LP (1981) Distribuição da soja no mundo: origem, história e distribuição. In: Miyasaka S, Medina JC (Eds.). **A soja no Brasil**. Campinas: ITAL, p. 1-6.

Charoen-Ung P, Mittrapiyanuruk P (2019). Sugarcane Yield Grade Prediction Using Random Forest with Forward Feature Selection and Hyper-parameter Tuning. In: Unger H, Sodsee S, Meesad, P. (Eds.) **Recent Advances in Information and Communication Technology 2018**. IC2IT 2018. New York: Springer Cham, p. 33-42. DOI: https://doi.org/10.1007/978-3-319-93692-5_4

Chaves MM, Oliveira MM (2004) Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. **Journal of Experimental Botany** 55:2365–2379. DOI: https://doi.org/10.1093/jxb/erh269

Chlingaryan A, Sukkarieh S, Whelan B (2018) Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. **Computers and Electronics in Agriculture** 151: 61-69. DOI: https://doi.org/10.1016/j.compag.2018.05.012

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2023) **Acompanhamento da Safra Brasileira de Grão 2022/2023:** relatório técnico n.6, sexto levantamento. Disponível em: https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 12/03/2023.

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2022) **Acompanhamento da Safra Brasileira de Grão 2021/2022:** relatório técnico n.12, décimo segundo levantamento. Disponível em: https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 10/03/2023.

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2023) **Portal de informações agropecuárias**. Disponível em: https://portaldeinformacoes.conab.gov.br/safra-serie-historica-graos.html. Acesso em: 12/03/2023.

Costa NL, Santana AC (2018) Análise do mercado da soja: aspectos conjunturais da formação do preço pago ao produtor brasileiro. **Revista Plantio Direto** 28: 20-39.

Daffe RT, F.W (1982) **Relatório anual do Instituto Agronômico do Estado de São Paulo**. Campinas: IAC, 26 p.

Desloires J, Lenco D, Botrel A (2023) Out-of-Year Corn Yield Prediction at Field-Scale Using Sentinel-2 Satellite Imagery and Machine Learning Methods. **Preprint.** DOI: http://dx.doi.org/10.2139/ssrn.4329505

Edreira JIR, Mourtzinis S, et al. (2017). Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. **Agricultural and Forest Meteorology**, 247:170–180. DOI: https://doi.org/10.1016/j.agrformet.2017.07.010

Elbeltagi A, Pande CB, Kumar M, Tolche DA, Singh SK, Kumar A, Vishwakarma KD (2023) Prediction of meteorological drought and standardized precipitation index based on the random forest (RF), random tree (RT), and Gaussian process regression (GPR) models. **Environmetal Science and Pollution Research** 30:43183-43202. DOI: https://doi.org/10.1007/s11356-023-25221-3

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. (2007) **O complexo agroindustrial da soja brasileira:** circular técnica 43. Londrina, 11 p.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária (2020) **Tecnologias de produção de soja**: Sistemas de Produção. Londrina: Embrapa Soja, 347p.

Farias JRB, Neumaier N, Nepomuceno AL (2007) **Ecofisiologia da soja.** Londrina: Embrapa Soja, 9 p. (Embrapa Soja, Circular Técnica, 48).

Federizzi LC (2005) A soja como fator de competitividade no Mercosul: histórico, produção e perspectivas futuras. In: III ENCONTRO CEPAN: VANTAGENS COMPETITIVAS DOS AGRONEGÓCIOS NO MERCOSUL. **Anais …** Porto Alegre: CEPAN/UFRGS, p. 1-16.

Ferrari E, Paz A, Silva AC (2015) Deficit hídrico no metabolismo da soja em semeaduras antecipadas no Mato Grosso. **Nativa** 3:67-77.

França-Neto JB, Krzyzannowski FC, Henning AA, Padua GP, Lorini I, Henning, FA (2016) **Tecnologia da produção de semente de soja de alta qualidade.** Londrina: Embrapa Soja, 82 p. (Embrapa Soja, Documentos, 380).

França-Neto JB, Pádua GP, Krzyzannowski FC, Carvalho MLM, Henning AA, Lorini I (2012) **Semente esverdeada de soja: causas e efeitos sobre o desempenho fisiológico – Série Sementes.** Londrina: Embrapa Soja, 15 p. (Embrapa Soja, Circular técnica, 91).

Gniewko N, Popowska KD, Piekutowska M, Wojciechowski T, Kwiatek M, Nawracała J (2022) Application of Artificial Neural Network Sensitivity Analysis to Identify Key Determinants of Harvesting Date and Yield of Soybean (Glycine max

[L.] Merrill) Cultivar Augusta. **Agriculture** 12:754. DOI: https://doi.org/10.3390/agriculture12060754

Gonçalves SL (2019) **Impactos da deficiência hídrica e de altas temperaturas na produtividade da soja no estado do Paraná, safra 2018/2019.** Londrina: Embrapa Soja, 23 p. (Embrapa Soja, Boletim de pesquisa e desenvolvimento, 22).

Hill, CB, Bowen CR, Hartmann GL (2013) Effect of fungicide application and cultivar on soybean green stem disorder. **Plant Disease**, 97:1212-1220. DOI: https://doi.org/10.1094/PDIS-12-12-1191-RE

Hymowitz T (1970) On the domestication of the soybean. **Economic Botany**, 24:408-21. DOI: https://doi.org/10.1007/BF02860745

Kaster M, Farias JRB (2012) **Regionalização dos testes de Valor de Cultivo e Uso e da indicação de cultivares de soja – terceira Aproximação**. Londrina: Embrapa Soja, 69 p.

Kayad A, Sozzi M, Gatto S, Marinello F, Pirotti F (2019) Monitoring Within-Field Variavility of Corn Yield using Sentinel-2 and Machine Learning Techniques. **Remote Sensing** 11:2873. DOI: https://doi.org/10.3390/rs11232873

Kisman, Hemon AF, Sumarjan, Suprayanti MD (2023) Physiological response of three large-seeded soybean genotypes under drought and waterlogged stress conditions. **AIP Conference Proceedings** 2583:020027. DOI: https://doi.org/10.1063/5.0116156

Koester RP, Skoneczka JA, Cary TR, Diers WB, Ainsworth AE (2014) Historical gains in soybean (Glycine max Merr.) seed yield are driven by linear increases in light interception, energy conversion, and partitioning efficiencies. **Journal of Experimental Botany** 65:3311–3321. DOI: https://doi.org/10.1093/jxb/eru187

Kuwata K, Shibasaki R (2016) Estimating corn yield in the United States with MODIS EVI and machine learning methods. In: ISPRS ANNALS OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. **Anais …** Praga, República Checa: ISPRS, p. 131-136.

Latef AAHA, Jan S, Abd-Allah EF, Rashid B, John R, Ahmad P (2016). Soybean under abiotic stress. In.: Azooz MM, Ahmad P (Eds.) **Plant-Environment Interaction: Resposnses and Approaches to Mitigate Stress.** New Jersey: John Wiley e Sons, p.28-42

Lin X, Liu B, Weller JL, Abe J, Kong F (2021). Molecular mechanisms for the photoperiodic regulation of flowering in soybean. **Journal of Integrative Plant Biology** 63: 981– 994. DOI: https://doi.org/10.1111/jipb.13021

Lopez MA, Moreira FF, Hearst A, Cherkauer K, Rainey KM (2022) Physiological breeding for yield improvement in soybean: solar radiation interception-conversion, and harvest index. **Theoretical and Applied Genetics** 135:1477-1491. DOI: https://doi.org/10.1007/s00122-022-04048-5

Minsen G (1901) A soja. **Revista Agrícola do Rio Grande do Sul** 5:2-4.

Miyasaka S, Medina JC (1981) Introdução e evolução da soja no Brasil. 5. No Estado de São Paulo. In: Miyasaka S, Medina JC (Eds.) **A soja no Brasil.** Campinas: ITAL, p.24-5.

Mohammed S, Elbeltagi A, Bashir B, Alsafadi K, Alsilibe F, Alsalman A, Zeraatpisheh M, Széles A, Harsányi E (2022) A comparative analysis of data mining techniques for agricultural and hydrological drought prediction in the eastern Mediterranean. **Computers and Eletronics in Agriculture** 197:106925. DOI: https://doi.org/10.1016/j.compag.2022.106925

Morelli-Ferreira F, Maia NJC, Tedesco D, Kazama EH, Morlin Carneiro F, Santos LB, Seben Junior GF, Rolim GS, Shiratsuchi LS, Silva RP (2021) Comparison of Machine Learning Techniques in Cotton Yield Prediction Using Satellite Remote Sensing. **Preprints**, DOI: https://doi.org/10.20944/preprints202112.0138.v2.

Morse WJ (1950) History of soybean production. In: Markley KS (Ed.) **Soybeans and soybean products.** New York: Interscience, p. 3-59.

Nouri MZ, Toorchi M, Komatsu S (2011) Proteomics approach for identifying abiotic stress responsive proteins in soybean. In.: Sudaric A (Ed.) **Soybean: Molecular Aspects of Breeding.** Rijeka, Croatia: Intech, p.187–214.

Ort NWW, Morrison MJ, Cober ER, Samanfar B, Lawley Y (2022) Photoperiod Affects Node Appearance Rate and Flowering in Early Maturing Soybean. **Plants** 11:871. DOI: https://doi.org/10.3390/plants11070871

Pelech EA, Evers JB, Pederson TL, Drag DW, Fu P, Bernacchi CJ (2023) Leaf, plant, to canopy: A mechanistic study on aboveground plasticity and plant density within a maize–soybean intercrop system for the Midwest, USA. **Plant, Cell e Environment** 46:405–421. DOI: https://doi.org/10.1111/pce.14487

Pisner DA, Schnyer DM (2020) Support Vector Machine. In.: Mechelli A, Vieira SBT (Eds.) **Machine Learning – Methods and Applications to Brain Disorders.** Cambridge: Academic Press, p. 101–121. DOI: https://doi.org/10.1016/C2017-0-03724-2

Puma MJ (2019) Resilience of the global food system. **Nature Sustainability** 2: 260–261. DOI: https://doi.org/10.1038/s41893-019-0274-6

Razhika F, Cahyoadi B (2023) Length of the wet seasons and its effect on food crop productivity (rice, corn, and soybean). **AIP Conference Proceedings** 2583:020016. DOI: https://doi.org/10.1063/5.0115799

Ren Y, Li Q, Du X, Zhang Y, Wang H, Shi G, Wei M (2023) Analysis of Corn Yield Prediction Potential at Various Growth Phases Using a Process-Based Model and Deep Learning. **Plants** 12:446. DOI: https://doi.org/10.3390/plants12030446

Sanches AC, Michellon E, Roessing AC (2007) Os limites de expansão da soja. **Informe GEPEC** 9:1-21. DOI: https://doi.org/10.48075/igepec.v9i1.201

Schapaugh B, Roozeboom K (2012) K-State Agronomy Extension e-Updates. Manhattan: Arkansas State University, 2012. 20 p. n. 371. Disponível em: http://www.agronomy.k-state.edu/documents/eupdates/eupdate100512.pdf Acesso em: 21 setembro. 2020.

Sediyama T, Silva F, Borém A. (2015) **Soja: do plantio a colheita.** Viçosa, MG: UFV, p. 333.

Sentelhas PC, Battisti R, Monteiro LA, Duarte YCN, Visses FA (2016) Yield gap: conceitos, definições e exemplos. **Informações Agronômicas** 155:9-12.

Sentelhas PC, Battisti R, Câmara GMS, Farias JRB, Hampf AC, Nendel, C (2015) The soybean yield gap in Brazil – magnitude, causes and possible solutions for sustainable production. **The Journal of Agricultural Science** 153:1394-1411 DOI: https://doi.org/10.1017/S0021859615000313

Silva RA (2018) **Impacto das mudanças climáticas sobre a produtividade e pegada hídrica da soja cultivada na região do Matopiba.** Tese (Doutorado em Meteorologia) – Universidade Federal de Campina Grande, Centro de Tecnologia e Recursos Naturais.

SoyStats 2021. Disponível online: http://www.soystats.com (acessado em 23 de novembro de 2022)

Staniak M, Szpunar-Krok E, Kocira A (2023) Responses of Soybean to Selected Abiotic Stresses—Photoperiod, Temperature and Water. **Agriculture** 13:146. DOI: https://doi.org/10.3390/agriculture13010146

USDA - United States Department of Agriculture Foreign. Oilseeds: World Markets and Trade. PSD 3/2023. Disponível em: https://usda.library.cornell.edu/concern/publications/tx31qh68h?locale=en. Acesso em: 12/03/2023.

Van Klompenburg T, Kassahun A, Catal C (2020) Crop yield prediction using machine learning: A systematic literature review. **Computers and Electronics in Agriculture** 177:105709. DOI: https://doi.org/10.1016/j.compag.2020.105709

Vernetti FJ (1977) História e importância da soja no Brasil. **A Lavoura** 81: 21-4.

Vital RG, Müller C, Freire FBS et al. (2022) Metabolic, physiological and anatomical responses of soybean plants under water deficit and high temperature condition. **Scientific Reports** 12:16467. DOI: https://doi.org/10.1038/s41598-022-21035-4

Zanon AJ, Streck NA, Grassini P (2016) Climate and management factors influence soybean yield potential in a subtropical environment. **Agronomy Journal** 108: 1447-1454. DOI: https://doi.org/10.2134/agronj2015.0535

Zhou T, Wang L, Yang H, Yang G, Liu W, Yang W (2019) Ameliorated light conditions increase the P uptake capability of soybean in a relay-strip intercropping system by altering root morphology and physiology in the areas with

low solar radiation. **Science of The Total Environment** 688: 1069-1080. DOI: https://doi.org/10.1016/j.scitotenv.2019.06.344

CAPÍTULO 2 – INFLUÊNCIA DOS ELEMENTOS CLIMÁTICOS NA PRODUTIVIDADE DA SOJA NO BRASIL

RESUMO – A capacidade produtiva da soja é extremamente sensível aos elementos climáticos. Compreender os impactos desses elementos é essencial para quantificar sua influência na produtividade da soja. Dessa forma, o objetivo deste estudo foi avaliar a influência dos elementos climáticos na produtividade da soja, de acordo com o estádio fenológico da cultura. O estudo foi realizado em 154 localidades produtoras de soja no Brasil. Os dados de produtividade de soja assim como os dados meteorológicos são de uma série histórica de 30 anos (1988 – 2018). Os dados meteorológicos são provenientes do sistema NASA-POWER, os dados utilizados foram: precipitação (mm), temperatura média, máxima e mínima do ar (ºC), velocidade do vento média, máxima e mínima (m s^{-1}), umidade relativa do ar (%), radiação solar no topo da atmosfera (MJ m^2 dia^{-1}) e radiação solar global (MJ m^2 dia^{-1}). Para cada localidade foi calculada a evapotranspiração potencial pelo método padrão da FAO e o balanço hídrico por Thornthwaite e Mather (1955). Todos os dados climáticos foram agregados em escala decendial. Foi realizado análise de cluster para classificação das localidades de acordo com as características climáticas de cada local, resultando em quatro grupos, sendo esses os grupos 1, 2, 3 e 4. A influência dos elementos meteorológicos na produtividade da soja foi verificada pela análise de correlação de Pearson. As análises foram realizadas utilizando o software Python 3.8. Observamos que as variáveis que mais influenciaram a produtividade da soja foram: no grupo 1, o excedente hídrico (EXC) teve correlação negativa de 12% e 22%, nos estádios de estabelecimento e desenvolvimento, respectivamente. A temperatura mínima (Tmin) apresentou correlação positiva de 20% no estádio de florescimento. E o armazenamento de água no solo (STO) teve correlação positiva de 17% no estádio de frutificação. No grupo 2, a umidade relativa do ar (RU) apresentou correlação positiva de 17% no estádio de estabelecimento. O saldo de radiação (RN) teve correlação positiva de 21% e 27%, nos estádios de desenvolvimento e florescimento, respectivamente. A umidade relativa do ar (RU) teve correlação positiva de 22% no estádio de frutificação. E a temperatura máxima (Tmax) apresentou correlação negativa de 16% no estádio de maturação. No grupo 3, a umidade relativa do ar (RU) apresentou correlação positiva de 13% no estádio de estabelecimento. A radiação global (Qg) apresentou correlação positiva de 16% no estádio de desenvolvimento. O saldo de radiação (RN) apresentou correlação positiva de 16% no estádio de florescimento. A umidade relativa do ar (RU) teve correlação positiva de 22% e 25% nos estádios de frutificação e maturação, respectivamente. No grupo 4, a velocidade mínima do vento (WSMIN) mostrou correlação negativa de 18% e 20% nos estádios de desenvolvimento e estabelecimento. O saldo de radiação (RN) teve correlação positiva de 33% no estádio de florescimento. E a umidade relativa do ar (RU) apresentou correlação positiva de 35% nos estádios de frutificação e maturação. Os elementos climáticos influenciaram a produtividade da soja de forma diferente em cada estádio fenológico da cultura.

Palavras-chave: Agrometeorologia, balanço hídrico, correlação de Pearson, estádio fenológico, Nasa – Power

INFLUENCE OF CLIMATE ELEMENTS ON SOYBEAN PRODUCTIVITY IN BRAZIL

ABSTRACT – The productive capacity of soybeans is extremely sensitive to climatic elements. Understanding the impacts of these elements is essential to quantify their influence on soybean productivity. Therefore, the objective of this study was to evaluate the influence of climatic elements on soybean productivity, according to the phenological stage of the crop. The study was carried out in 154 soybean-producing locations in Brazil. Soybean productivity data as well as meteorological data are from a 30-year historical series (1988 – 2018). The meteorological data come from the NASA-POWER system, the data used were: occurrences (mm), average, maximum and minimum air temperature (ºC), average, maximum and minimum wind speed (m s^{-1}), relative humidity relative air content (%), solar radiation at the top of the atmosphere (MJ m^2 day^{-1}) and global solar radiation (MJ m^2 day^{-1}). For each location, potential evapotranspiration was calculated using the FAO standard method and the water balance of Thornthwaite and Mather (1955). All climate data was aggregated on a decennial scale. A cluster analysis was carried out to classify the locations according to the climatic characteristics of each location, resulting in four groups, these being groups 1, 2, 3 and 4. The influence of meteorological elements on soybean productivity was verified by the analysis Pearson brightness. Analyzes were performed using Python 3.8 software. We observed that the variables that most influenced soybean productivity were: in group 1, water surplus (EXC) had a negative correlation of 12% and 22%, in the establishment and development stages, respectively. The minimum temperature (Tmin) showed a positive correlation of 20% at the flowering stage. And soil water storage (STO) had a positive correlation of 17% at the fruiting stage. In group 2, relative air humidity (RU) showed a positive correlation of 17% at the establishment stage. The radiation balance (RN) had a positive correlation of 21% and 27%, in the development and flowering stages, respectively. Relative air humidity (RU) had a positive correlation of 22% at the fruiting stage. And the maximum temperature (Tmax) showed a negative correlation of 16% at the maturation stage. In group 3, relative air humidity (RU) showed a positive correlation of 13% at the establishment stage. Global radiation (Qg) showed a positive correlation of 16% in the developmental stage. The radiation balance (RN) showed a positive correlation of 16% at the flowering stage. Relative air humidity (RU) had a positive correlation of 22% and 25% at the fruiting and maturation stages, respectively. In group 4, the minimum wind speed (WSMIN) showed a negative correlation of 18% and 20% in the development and establishment stages. The radiation balance (RN) had a positive correlation of 33% at the flowering stage. And relative air humidity (RU) showed a positive correlation of 35% in the fruiting and maturation stages. Climatic elements influenced soybean productivity differently at each phenological stage of the crop.

Keywords: Agrometeorology, water balance, Pearson correlation, phenological stage, NASA – Power

2.1 Introdução

A soja é um dos grãos mais produzidos no Brasil, sua produção representa 46% da produção brasileira de grãos (CONAB, 2022). Desempenha papel importante na geração de emprego no setor agrícola, na alimentação humana e de animais, e movimenta a máquina financeira do país (Costa e Santana, 2018). O clima é uma das variáveis determinantes na produtividade da soja, elementos meteorológicos como temperatura do ar, radiação solar e precipitação estão diretamente correlacionados com a capacidade produtiva da soja (Battisti et al., 2017), em condições climáticas adversas a produtividade da soja pode reduzir em até 30% (Staniak et al., 2023).

Os elementos climáticos são responsáveis por regular diversos processos morfológicos e fisiológicos das plantas de soja, podendo a ação desses elementos causar diversos danos às plantas (Elmerich et al., 2023). Estima-se que até 2080, a maioria das áreas de cultivo do mundo estarão expostas a altas temperaturas (Battisti e Naylor, 2009). A exposição da soja a altas temperaturas causa impactos negativos a produtividade, pode induzir a floração precoce e encurtar o período de enchimento de grãos, quanto menor a duração da cultura menor é a produtividade (Ramteke et al., 2015). No cerrado brasileiro, estima-se a redução de 6% da produtividade de soja para o aumento de cada 1ºC (Silva et al., 2023).

No Brasil, o déficit hídrico é responsável por 74% das perdas de produtividade de soja (Battisti et al., 2018). Durante a seca de 2015, a produtividade ficou 33% abaixo da média do ano anterior, sendo o maior impacto observado nas regiões do cerrado e Matopiba (Silva et al., 2023). Dentre os principais danos ocasionados pelo déficit hídrico está o abortamento de flores,

vagens e antecipação do ciclo da cultura (Battisti et al., 2017). O período de maior demanda de água ocorre durante o florescimento e enchimento de grãos (7 a 8 mm dia^{-1}). A necessidade total de água exigida na cultura da soja para garantir alta produtividade varia entre 450 e 800 mm/ciclo (EMBRAPA, 2013).

No entanto, o excesso de água no solo também pode provocar diversos danos a cultura levando a queda de produtividade (Dash et al., 2022; Wu et al., 2022). As perdas de produtividade podem variar de 40 a 80%, dependendo da fenologia da cultura (Phunkan et al., 2016). Nos EUA, por exemplo, anualmente as perdas econômicas podem chegar a US$ 1,5 bilhão (Wu et al., 2020).

A radiação solar é outro elemento importante, sendo necessário para realização da fotossíntese (Taiz e Zeiger, 2017). Na cultura da soja está associado a arquitetura da planta (Pelech et al., 2023), expansão foliar, pegamento de grãos e peso de vagens (Bianculli et al., 2016). Dessa forma, a eficiência do uso de radiação pode ser usada para melhorar o potencial produtivo da soja (Lopez et al., 2022).

Muitos estudos abordam a relação dos elementos climáticos com a produtividade da soja (Zilli et al., 2020; Zhang et al., 2022; Moreira et al., 2023). No entanto, nenhum deles quantifica o impacto da correlação desses elementos em cada estádio fenológico da soja no Brasil. O Brasil é o maior produtor de soja do mundo (USDA, 2023), é sabido que os elementos climáticos são uma das principais causas de perda de produtividade na soja. Nesse sentido, o foco deste estudo é identificar a correlação dos elementos climáticos com a produtividade de soja em cada estádio fenológico da cultura. A compreensão do impacto

desses elementos pode auxiliar em estratégias que minimizem as perdas de produtividade causadas pelos elementos climáticos.

2.2 Material e métodos

2.2.1 Área de estudo e dados de produtividade

O estudo foi realizado em 154 localidades produtoras de soja no Brasil. Os dados de produtividade de soja foram obtidos por meio do banco de dados do PAM (Produção Agrícola Municipal), disponibilizados pelo Sistema IBGE de Recuperação Automática – SIDRA (IBGE, 2020), referentes ao período de 1988 a 2018. A seleção dos locais estudados foi realizada a partir da elaboração de um ranking proveniente destes dados, onde selecionamos os locais com maior área semeada e maior produtividade de soja no Brasil (Figura 1).

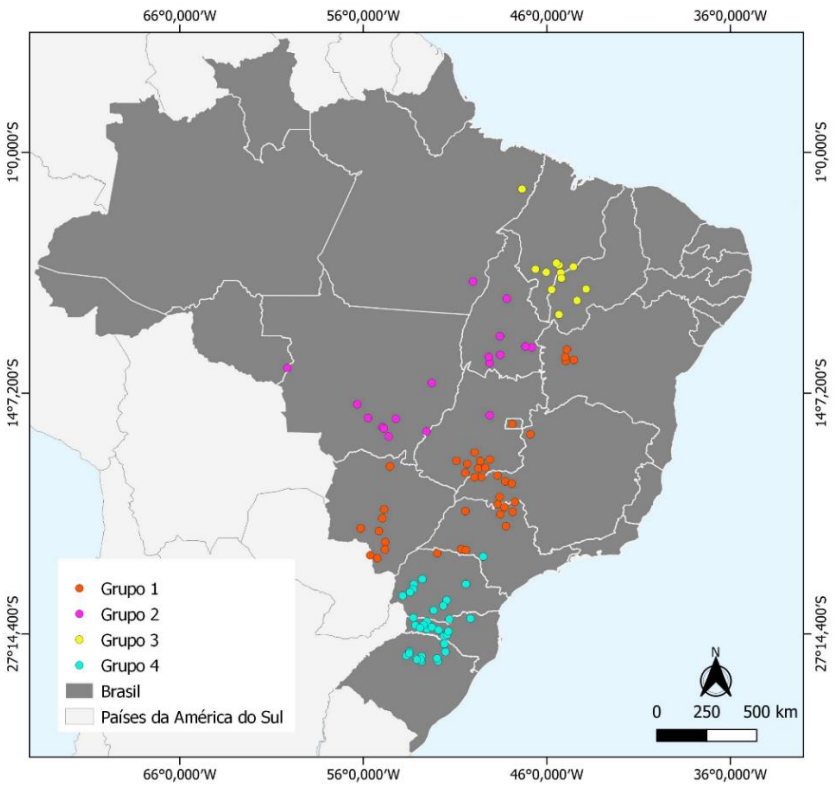

Figura 1. Localização geográfica das localidades estudadas.

2.2.2 Fonte de dados meteorológicos

Os dados meteorológicos foram coletados para cada local por meio da plataforma *National Aeronautics and Space Administration/Prediction of Worldwide Energy Resources* (NASA-POWER) referente ao período de 1988 a 2018. Estes dados possuem resolução espacial de 0,5º x 0,5º (latitude-longitude, equivalente a aproximadamente 56 km) em uma resolução temporal diária. Os dados meteorológicos coletados foram: precipitação (mm), temperatura média, máxima e mínima do ar (ºC), velocidade do vento média, máxima e mínima (m s^{-1}), umidade relativa do ar (%), radiação solar no topo da atmosfera (MJ m^2 dia^{-1}) e radiação solar global (MJ m^2 dia^{-1}). Em seguida, os dados foram organizados em períodos decendiais (10 dias). Uma explicação das etapas desenvolvidas

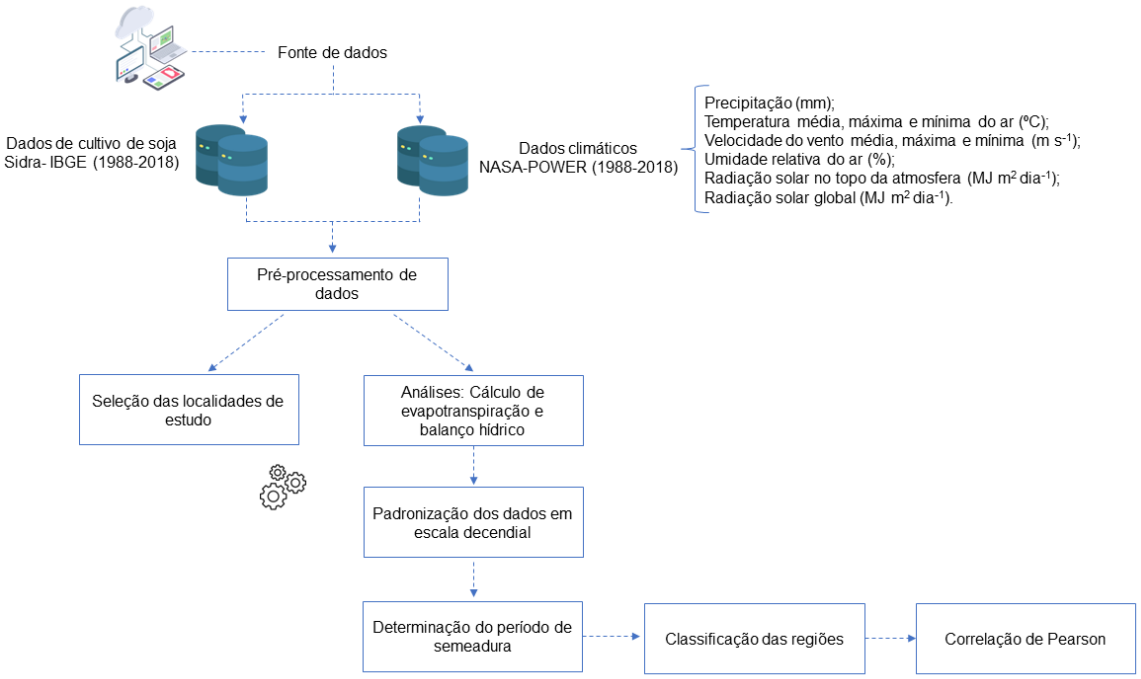

Figura 2. Fluxograma das etapas desenvolvidas neste trabalho.

neste trabalho pode ser observada na figura 2.

Os dados do NASA – POWER são confiáveis e precisos, fornecem informações meteorológicas de regiões onde não há medições de estações de superfície ou lacunas (Monteiro et al., 2018; Duarte e Sentelhas, 2020), diversos estudos utilizaram esses dados para modelagem de culturas agrícolas (Menezes et al., 2022; Oliveira et al., 2022; Lorençone et al., 2022).

A evapotranspiração potencial (PET) foi calculada para todas as localidades estudadas conforme o método de Penman e Monteith (Allen et al., 1998) em escala diária e em seguida estratificado em escala decendial, conforme descrito pelas equações 1 a 4.

$$PET = \frac{0.408 \times s \times (Rn - G) + \frac{\gamma \times 900 \times u_2 \times (e_s - e_a)}{T+273}}{s + \gamma \times (1 + 0.34 \times u_2)} \quad (1)$$

$$s = \frac{4098 \times e_s}{(T+273)^2} \quad (2)$$

$$e_a = \frac{RH \times e_s}{100} \quad (3)$$

$$e_s = 0.6108 \times e^{\frac{17.27 \times T}{237.3+T}} \quad (4)$$

em que, s é a declividade da curva de pressão de vapor em relação à temperatura (kPa °C^{-1}), Rn é o saldo de radiação diária (MJ m^{-2} dia^{-1}), G é o fluxo total diário de calor no solo (MJ m^{-2} dia^{-1}), T é a temperatura média do ar (°C), γ é a constante psicométrica (0.063 kPa ºC^{-1}), RH é a umidade relativa (%), u_2 é a velocidade do vento a 2 m de altura (m s^{-1}), e_a é a pressão atual de vapor (kPa), e_s é a pressão de saturação de vapor (kPa).

O balanço hídrico foi calculado para todas as localidades estudadas conforme o método de Thornthwaite e Mather (1955) (Equações 5 a 10) em

escala decendial. Foi utilizado para todas as localidades a capacidade de água disponível (CAD) de 100 mm, seguindo as sugestões de Duarte e Sentelhas (2020) para a cultura da soja.

$$se\ (P-PET)_i < 0 \begin{Bmatrix} NAC_i = NAC_{i-1} + (P-PET)_i \\ STO_i = WC_e^{\left(\frac{NAC_i}{WC}\right)} \end{Bmatrix} \quad (5)$$

$$se\ (P-PET)_i \geq 0 \begin{Bmatrix} STO_i = (P-PET)_i + STO_{i-1} \\ NAC_i = WCl_n^{\left(\frac{STO_i}{WC}\right)} \end{Bmatrix} \quad (6)$$

$$ALT_i = STO_i - STO_{i-1} \quad (7)$$

$$AET_i = \begin{Bmatrix} P + |ALT_i|\ ,\ se\ ALT < 0 \\ PET_i\ ,\ se\ ALT \geq 0 \end{Bmatrix} \quad (8)$$

$$DEF = PET - AET \quad (9)$$

$$EXC_i = \begin{Bmatrix} 0\ ,\ se\ WC < 0 \\ P - PET_i - AL, se\ ALT_i\ ,\ se\ WC = 0 \end{Bmatrix} \quad (10)$$

onde, P é a precipitação (mm), PET é a evapotranspiração potencial (mm), NAC o valor de P-PET, i é o mês atual, STO é o armazenamento de água no solo (mm), WC é a capacidade de água disponível (mm), ALT é a alteração do armazenamento de água no solo entre o mês anterior e o atual (mm), AET é a evapotranspiração real (mm), DEF é o déficit hídrico no solo (mm), e EXC é o excedente hídrico no solo (mm).

2.2.3 Determinação do período de semeadura e colheita

O Brasil é um país de escala continental com alta variabilidade climática entre as regiões. A soja por sua vez, é extremamente sensível aos elementos climáticos, dessa forma, a data de semeadura é um critério importante a ser considerado para que o ciclo da cultura evite passar por períodos de condições

climáticas desfavoráveis que venham prejudicar o potencial produtivo da cultura. Portanto, para este estudo determinamos que o decêndio 31 é o período de semeadura que melhor representa as localidades estudadas, sendo a finalização do ciclo da cultura/colheita estabelecida no decêndio 9 (Figura 3).

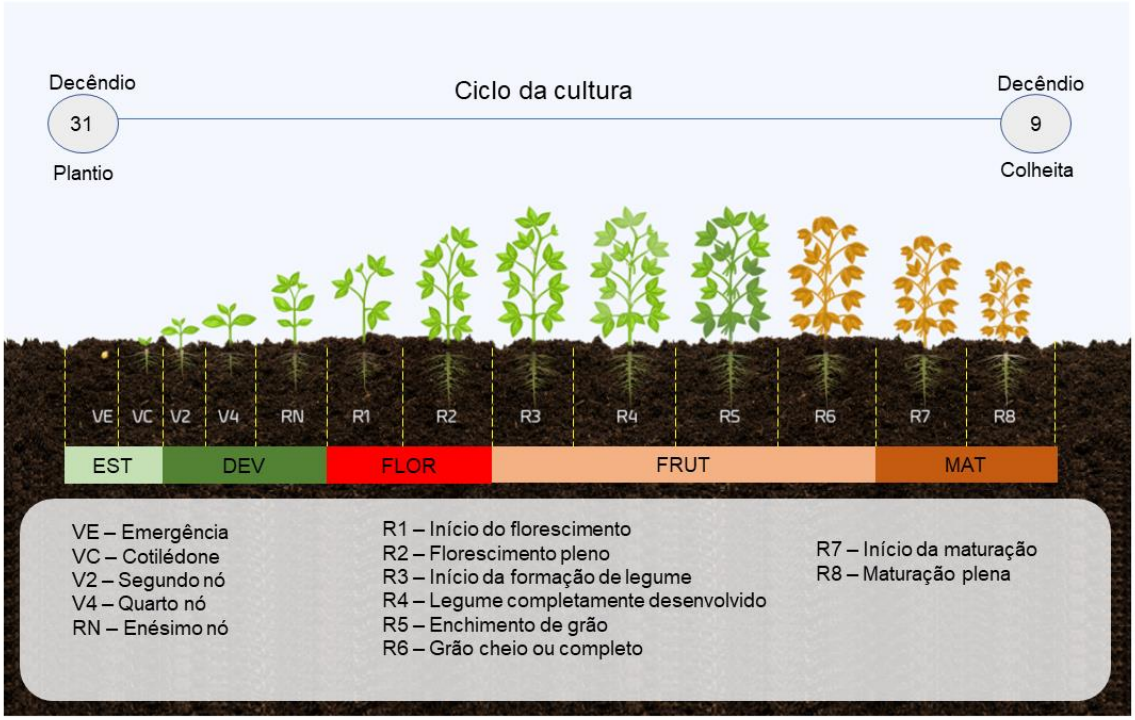

Figura 3. Esquema fenológico da soja. EST: Estabelecimento, DEV: Desenvolvimento, FLOR: Florescimento, FRUT: Frutificação, MAT: Maturação.

2.2.4 Classificação das regiões

A classificação das localidades foi realizada através da análise de cluster, levando em consideração os dados climáticos, resultando na divisão das localidades em quatro grupos de similaridade climática (Figura 4).

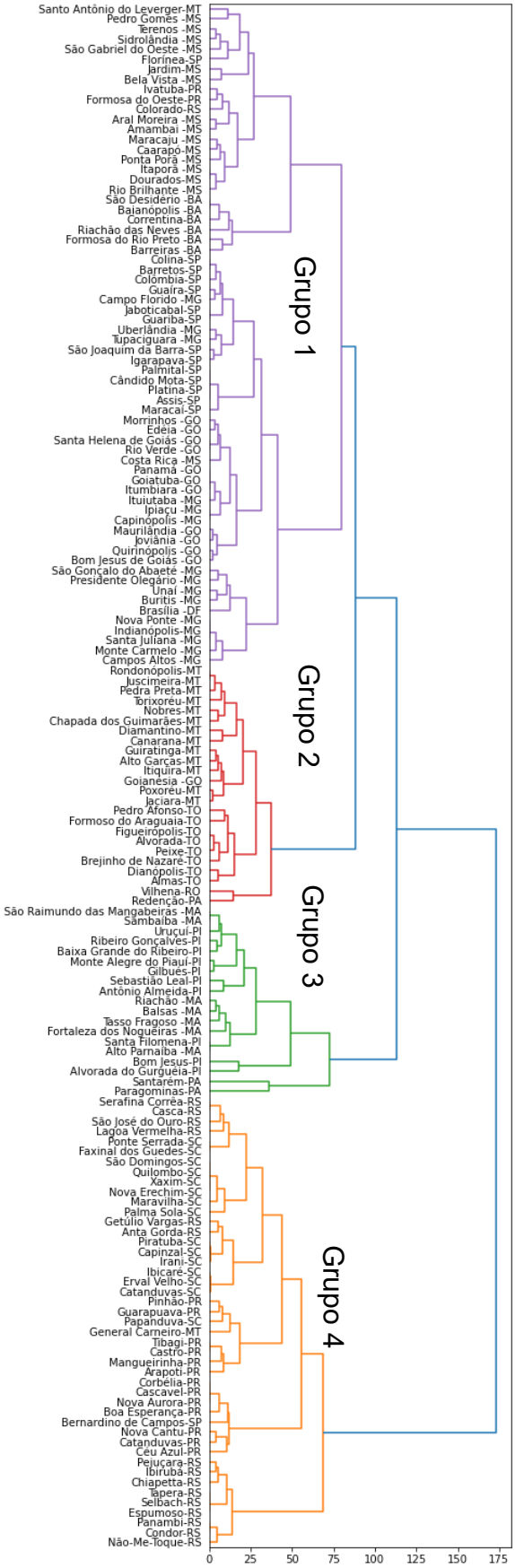

Figura 4. Dendrograma das localidades produtoras de soja no Brasil em função das varáveis climáticas.

2.2.5 Análise estatística

A influência dos elementos climáticos na produtividade da soja considerando cada estádio fenológico da planta foi quantificada por meio da correlação de Pearson. Esta correlação nos permite quantificar a relação entre duas variáveis, dessa forma, é possível identificar a importância de cada elemento climático em relação à produtividade da cultura. Todas as análises foram feitas com a linguagem de programação python versão 3.9 (https://www.python.org/).

2.3 Resultados e discussão

De modo geral constatamos que os elementos climáticos estudados influenciaram na produtividade das plantas de soja de forma diferente em cada estádio fenológico da planta, de acordo com as suas necessidades fisiológicas. Em todos os grupos, a velocidade do vento (WS) no estádio inicial do estádio de desenvolvimento e florescimento possui correlação negativa de 14% com a produtividade, assim como na metade do estádio de maturação, ou seja, nesses estádios da cultura, ventos com velocidades mais altas (>5,43 m $^{s-1}$) afetam negativamente a produtividade final da planta (Figura 4a, 5a, 6a e 7a).

A velocidade do vento influência aspectos morfológicos e fisiológicos das plantas, por exemplo, o crescimento das plantas, alongamento do caule, tamanho das folhas, distribuição de assimilados, evapotranspiração, déficit de pressão de vapor e consequentemente a fotossíntese (Zhang et al., 2022). Além disso, altas velocidades do vento podem provocar danos mecânicos as folhas, abrasão em plantas jovens e acamamento. Danos como abrasão e rasgo de

folhas podem danificar a cutícula e resultar em perda de água, além de facilitar a entrada de patógenos (Gardiner et al., 2016; Zhang et al., 2021; Li et al., 2022).

Estudos demonstraram que plantas expostas a ventos fortes fecham os estômatos diminuindo assim a condutância estomática, resultando na diminuição da transpiração estomática e consequentemente a redução da concentração de CO_2 intercelular e temperatura da folha, levando a diminuição da taxa fotossintética da planta (Rosenberg, 1966; Brandle et al., 2004; Avecilla et al., 2015). Dessa forma, a limitação dos tecidos fotossintéticos inibe a assimilação do carbono, reduzindo assim a eficiência do uso da água e acúmulo de biomassa (Lin et al., 2023).

No estádio de florescimento, foi observado correlação positiva de 18% entre o saldo de radiação (RN) e a produtividade. O RN é constituído pelo balanço de ondas curtas e o balanço de ondas longas, é importante na descrição do ambiente físico da cultura, pois representa a energia disponível para o crescimento e desenvolvimento das plantas, e está relacionado diretamente com a radiação e evapotranspiração (Pereira, 2002; Sivakumar, 2023). Pierdoná et al., (2019) observaram que a maior parte do saldo de radiação foi consumida no processo de evaporação da água.

Em seguida também foi observado correlação negativa de 7% do déficit hídrico (DEF) principalmente no estádio de frutificação (Figura 2b, 3b, 4b e 5b). A maior exigência hídrica da soja ocorre principalmente no estádio de enchimento de grãos (Mandić, 2021). Conforme Villa et al., (2022) o déficit hídrico pode causar queda de produtividade, dificuldades de emergência e

estabilidade do stand da cultura, redução da área foliar, de raízes, da floração e senescência foliar.

Nas plantas, o processo mais afetado pelo DEF é a expansão celular. Nas plantas pode ocorrer inibição do crescimento e da fotossíntese, enquanto que as concentrações de solutos no citoplasma e de ácido abscísico (ABA) no vacúolo das células aumentam, permitindo que seja mantida a pressão de turgor a baixos potenciais hídricos (Taiz e Zeiger, 2017; Mukherjee et al., 2022). A pressão de turgor é essencial nos processos vegetais, pois o crescimento de células vegetais requer pressão de turgor para expansão das paredes celulares, além disso, aumenta a rigidez mecânica de células e tecidos (Taiz e Zeiger, 2017).

Resultados encontrados por Cotrim et al., (2021) confirmam que plantas sob estresse hídrico reduzem a condutância estomática e a transpiração, porém, aumentam a eficiência de uso da água quando expostas a condições de umidade do solo abaixo do ideal como um sistema de defesa da planta.

Nos grupos 1, 2 e 4 a umidade relativa do ar (RU) é um dos principais elementos correlacionados de forma positiva com a produtividade, com uma correlação de 15% no estádio de frutificação, entretanto, para o grupo 3 esta correlação ocorre no estádio de maturação das plantas 16%. O mesmo ocorre com a evapotranspiração (ETC), no qual teve correlação positiva de 6% para o grupo 1, grupo 2 e grupo 4 no estádio de frutificação, entretanto, para o grupo 3 esta correlação positiva de RU ocorreu no estádio de florescimento 9%. Os resultados do grupo 3 podem ser explicados devido as regiões desse grupo se caracterizarem por apresentarem temperaturas mais altas, períodos de seca e

chuva com volumes bem concentrados, e definidos em duas épocas do ano: seca de maio a setembro e chuvoso de outubro a abril (Reis et al., 2020).

No grupo 1 (Figura 5) observamos que no estádio de estabelecimento especificamente no primeiro decêndio um dos elementos mais importantes correlacionados de forma positiva com a produtividade foi a umidade relativa (RU) com 11%. Portanto, durante esse estádio valores de RU mais elevadas refletem de forma positiva na produtividade. Já no final do estabelecimento essa correlação positiva é observada na radiação global (Qg) 9,8%, porém, há correlação negativa com o excedente hídrico (EXC) de 12%, indicando que altos valores de EXC podem resultar em dificuldades durante o estabelecimento da cultura e consequentemente a redução da produtividade. Monteiro (2009), afirma que na germinação-emergência o conteúdo de água no solo não deve exceder 85% da água total máxima disponível.

No estádio de desenvolvimento, no primeiro decêndio há correlação positiva de 16% para a temperatura mínima (Tmin) e no terceiro decêndio 17% para Qg, porém, há correlação negativa para EXC de 22%. Resultados semelhantes foram observados no estádio de florescimento onde o primeiro decêndio teve correlação positiva de 20% para a Tmin e correlação negativa no ultimo decêndio de 6% para EXC. Board e Kahlon (2011) observaram que altas temperaturas (>40ºC) provocam distúrbios na floração e abortamento de flores.

Wu et al., (2022) destacaram que em condições de excesso de água no solo sinais do sistema radicular são transferidos para o solo, fazendo com o que os estômatos se fechem, reduzindo a taxa fotossintética e a absorção de CO_2, resultando em impactos negativos na produtividade.

No estádio de frutificação, no segundo decêndio há correlação negativa de 15% para DEF e positiva de 11% no terceiro decêndio para evapotranspiração real (AET), conforme vai chegando ao final do estádio há correlação positiva de 18% para RU, assim como 13% para Tmin.

No estádio de maturação, no primeiro decêndio há correlação positiva de 14% para RU, porém conforme avança para o final do ciclo da cultura esse percentual vai diminuindo e chega a 3% no ultimo decêndio. No penúltimo decêndio há correlação positiva de 15% para Qg e no ultimo decêndio há correlação negativa de 14% para EXC.

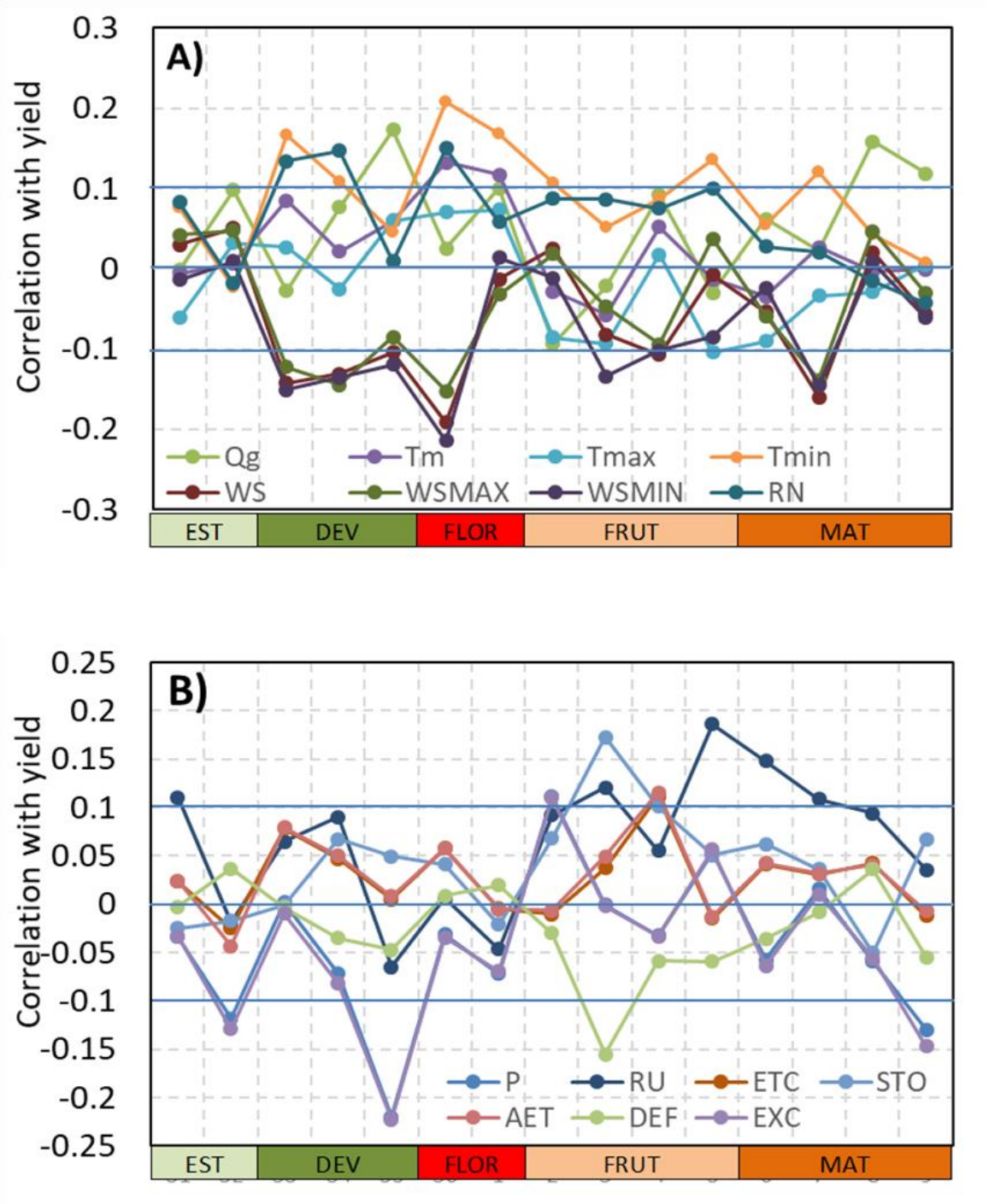

Figura 5. Correlação das variáveis climáticas x produtividade do Grupo 2. EST: Estabelecimento; DEV: Desenvolvimento vegetativo; FLOR: Florescimento; FRUT: Frutificação; MAT: Maturação. A) Qg: Radiação solar global; Tm: Temperatura média; Tmax: Temperatura máxima; Tmin: Temperatura mínima; WS: Velocidade do vento; WSMAX: Velocidade máxima do vento (WSMAX); WSMIN: Velocidade mínima do vento; RN: Saldo de radiação. B) P: Precipitação; RU: Umidade relativa; ETC: Evapotranspiração; STO: Armazenamento de água no solo; AET: Evapotranspiração real; DEF: Déficit hídrico; EXC: Excedente hídrico.

No grupo 2 (Figura 6) no estádio de estabelecimento constatamos que no primeiro decêndio há correlação positiva de 17% para RU, por outro lado, há correlação negativa de 11% para DEF, assim como para a temperatura máxima (Tmax). Em seguida, no segundo decêndio há correlação negativa de 7% para a precipitação (P), sendo este um elemento diretamente correlacionado com os resultados de EXC em todos os estádios fenológicos da cultura.

No estádio de desenvolvimento, no primeiro decêndio há correlação negativa de 15% para a velocidade mínima do vento WSMIN. Já no segundo decêndio há correlação positiva de 21% para RN, assim como 15% para Tmin. No último decêndio há correlação negativa de 13% para EXC.

No primeiro decêndio do florescimento o principal elemento correlacionado de forma positiva foi o RN com correlação de 27%, observamos ainda que esta correlação vai diminuindo conforme chega ao final do florescimento. No último decêndio há correlação negativa de 8% para EXC. No estádio de frutificação no primeiro decêndio também há correlação positiva de 17% para RN. O RN é importante no estádio de enchimento de grãos devido ao aumento da demanda fotossintética nesse estádio (Purcell et al., 2002). Em seguida, no segundo decêndio há correlação negativa de 16% para Tmax, no terceiro decêndio há correlação negativa de 9% para EXC, assim como correlação de 11% para DEF no ultimo decêndio, e correlação positiva de 22% para RU.

No estádio de maturação, no primeiro decêndio há correlação negativa de 16% para Tmax. No penúltimo decêndio há correlação positiva de 14% para Qg e no ultimo decêndio há correlação negativa de 13% para EXC. Tonello et al.,

(2022) destacam que em relação a qualidade dos grãos as oleaginosas requerem condições ótimas no estádio maturação/colheita, com temperaturas amenas e baixo índice de precipitação para evitar estresses bióticos e abióticos que podem levar a produção de grãos esverdeados, com qualidades fisiológicas e organolépticas reduzidas, esses grãos quando destinados a produção de óleo vegetal resultam num produto com maior acidez e custo elevado de refinação.

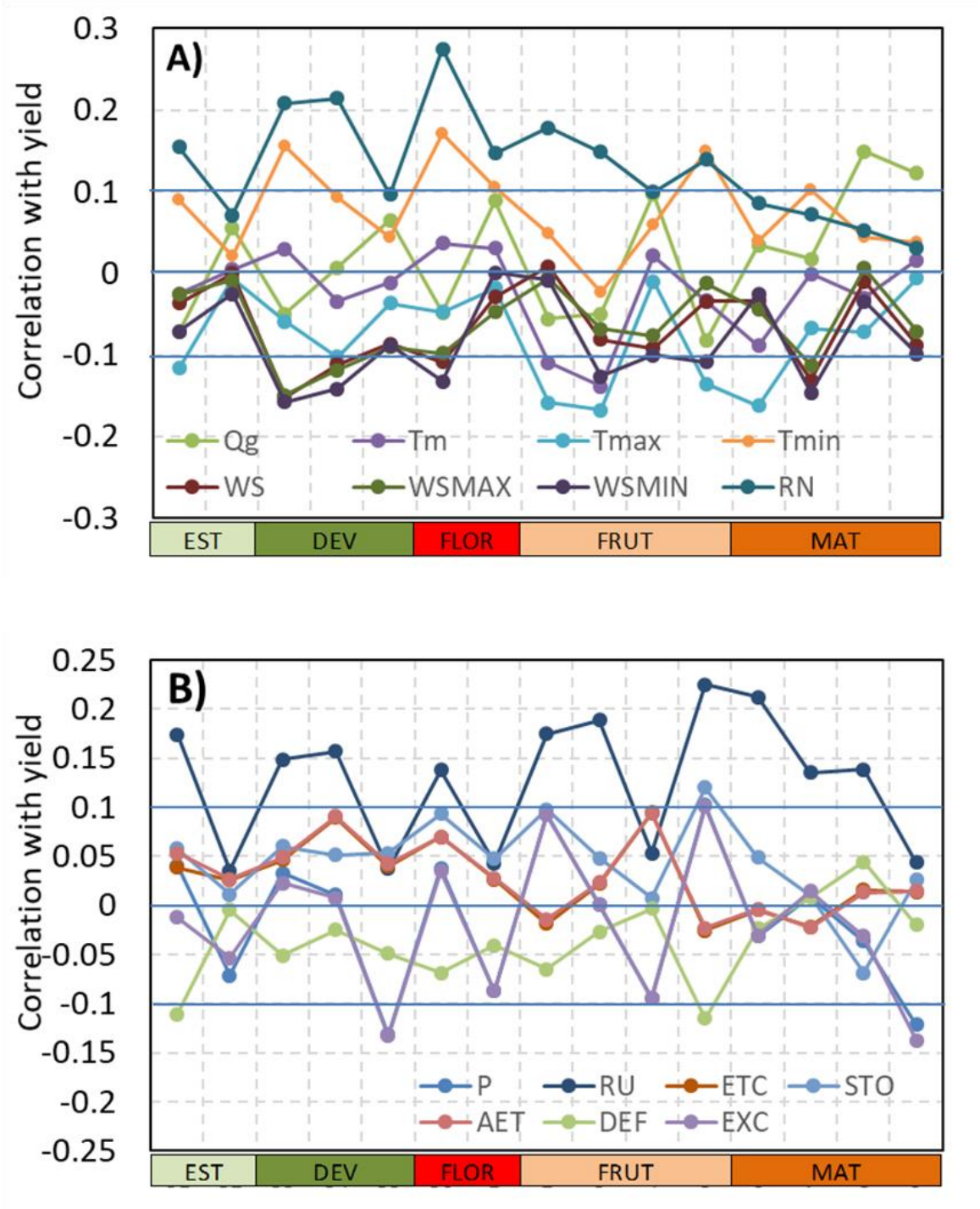

Figura 6. Correlação das variáveis climáticas x produtividade do Grupo 2. EST: Estabelecimento; DEV: Desenvolvimento vegetativo; FLOR: Florescimento; FRUT: Frutificação; MAT: Maturação. A) Qg: Radiação solar global; Tm: Temperatura média; Tmax: Temperatura máxima; Tmin: Temperatura mínima; WS: Velocidade do vento; WSMAX: Velocidade máxima do vento (WSMAX); WSMIN: Velocidade mínima do vento; RN: Saldo de radiação. B) P: Precipitação; RU: Umidade relativa; ETC: Evapotranspiração; STO: Armazenamento de água no solo; AET: Evapotranspiração real; DEF: Déficit hídrico; EXC: Excedente hídrico.

No grupo 3 (Figura 7) no estádio de estabelecimento no primeiro decêndio há correlação positiva de 13 % para RU, já no segundo decêndio essa correlação positiva é representada por Qg, aproximadamente 13%. No estádio de desenvolvimento no segundo decêndio há correlação positiva de 14% para RN, no último decêndio há correlação positiva de 16% para Qg e correlação negativa de 14% para EXC.

No florescimento no primeiro decêndio há correlação positiva de 16% para RN e 12% para AET, no último decêndio há correlação negativa de 11% para P e também há correlação negativa de 13% para Qg. Na frutificação no segundo decêndio há correlação negativa de 15% para DEF e no terceiro decêndio há correlação de 8% positiva para Qg. No ultimo decêndio da frutificação há correlação negativa de 16% para Tmax, e correlação positiva de 22% para RU, assim como correlação positiva de 12% para EXC. Os estádios de florescimento e enchimento de grãos são períodos críticos para soja, o aumento de temperatura pode levar à diminuição da altura da planta e aceleração da maturação (Farias et al., 2009), além da diminuição da fotossíntese em 22% e redução da formação de vagens em 11% (Djanaguiraman et al., 2019). Já o déficit hídrico diminui a interceptação solar devido a redução da área foliar, reduzindo a absorção de dióxido de carbono e a taxa fotossintética (Gilbeart et al., 2011).

Na maturação no primeiro decêndio há correlação positiva de 25% para RU e 20% para armazenamento de água (STO), porém a correlação para essas duas variáveis vai diminuindo conforme avança para o final do ciclo da cultura. No terceiro decêndio há correlação positiva de 15% para Qg e correlação negativa de 10% no último decêndio para EXC. De acordo com Sentelhas e

Batisti (2015) o excesso de água na época da colheita pode favorecer a disseminação de doenças e reduzir a qualidade de grãos.

Reis et al., (2020) constataram que os Estados Maranhão e Piauí estão cada vez mais expostas a temperaturas mais altas, período chuvoso com chuvas mais intensas e período seco cada vez mais seco, além da redução de dias frios, consequentemente, essas alterações causam impacto negativo na produtividade da soja, devido aos danos fisiológicos e ao encurtamento do ciclo da cultura.

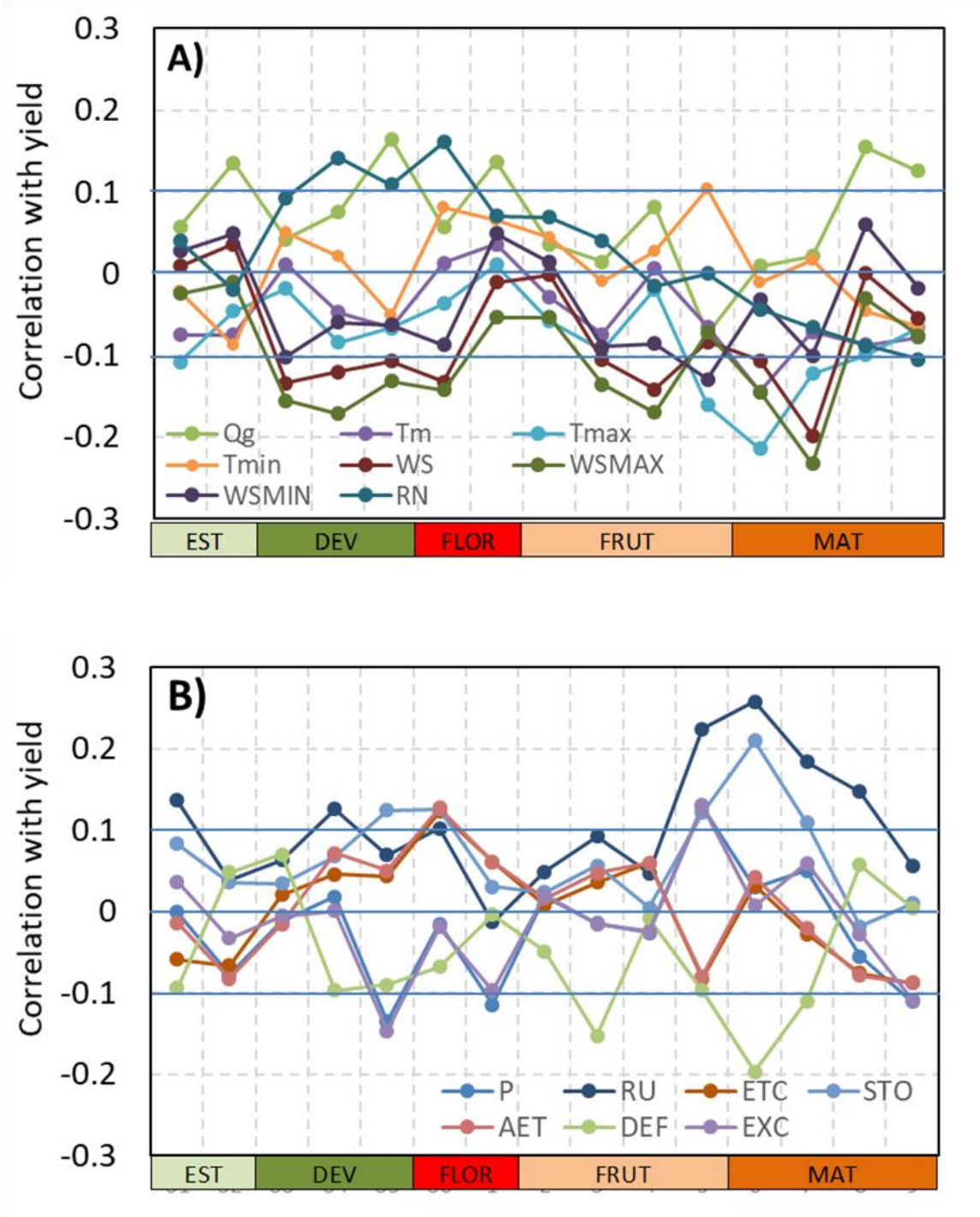

Figura 7. Correlação das variáveis climáticas x produtividade do Grupo 3. EST: Estabelecimento; DEV: Desenvolvimento vegetativo; FLOR: Florescimento; FRUT: Frutificação; MAT: Maturação. A) Qg: Radiação solar global; Tm: Temperatura média; Tmax: Temperatura máxima; Tmin: Temperatura mínima; WS: Velocidade do vento; WSMAX: Velocidade máxima do vento (WSMAX); WSMIN: Velocidade mínima do vento; RN: Saldo de radiação. B) P: Precipitação; RU: Umidade relativa; ETC: Evapotranspiração; STO: Armazenamento de água no solo; AET: Evapotranspiração real; DEF: Déficit hídrico; EXC: Excedente hídrico.

No grupo 4 (Figura 8) constatamos que há correlação negativa de 17% para WS em todos os estádios fenológicos, indicando que este é um elemento que possui forte influência na produtividade final da soja para as regiões que compõe esse grupo. No estádio de estabelecimento há correlação positiva de 11% no primeiro decêndio para RN e 10% no segundo decêndio para DEF. No estádio de desenvolvimento no primeiro decêndio há correlação negativa de 21% para Qg. Schirmbeck et al., (2019) observaram que nesse estádio inicial a maior disponibilidade de Qg seguido do déficit hídrico, resulta na utilização de uma maior porção do saldo de radiação para o aquecimento do ar e do solo.

No florescimento no primeiro decêndio há correlação positiva de 33% para RN e correlação negativa de 25% para DEF. No segundo decêndio há correlação positiva de 33% para RU e correlação negativa de 25% para Qg.

Na frutificação no segundo decêndio há correlação positiva de 12% para Qg, no terceiro decêndio há correlação positiva de 28% para AET e ETC que por estarem relacionadas possuem comportamento semelhante durante todo o ciclo da cultura. No último decêndio há correlação positiva de 35% para RU e correlação negativa de 24% para DEF.

Na maturação no primeiro decêndio há correlação positiva de 35% para RU, porém conforme avança para o final do ciclo esta correlação vai diminuindo. No segundo decêndio há correlação negativa de 6% para Qg, no terceiro decêndio há correlação positiva de 12% para EXC e no ultimo decêndio há correlação positiva de 12% para Tmin.

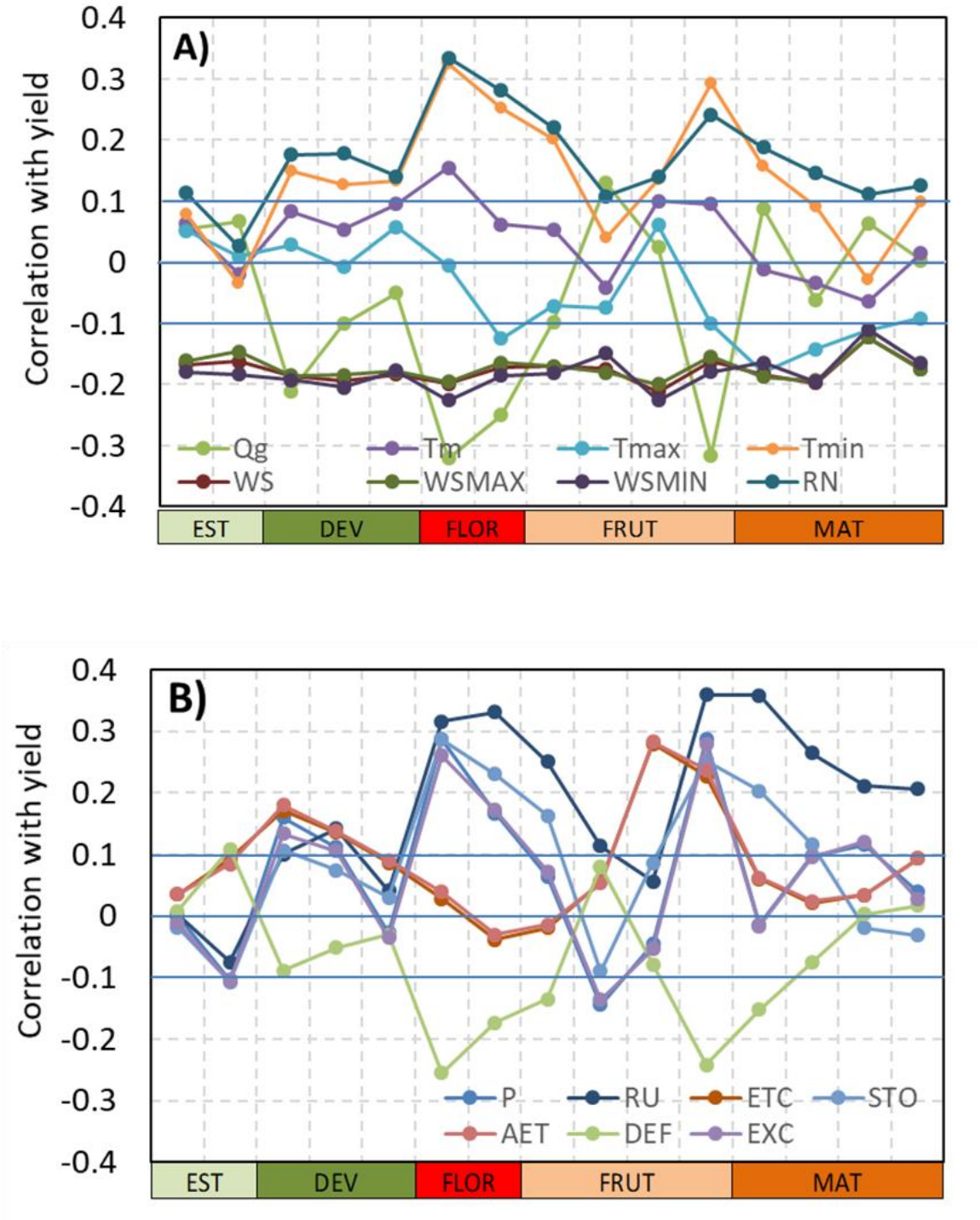

Figura 8. Correlação das variáveis climáticas x produtividade do Grupo 4. EST: Estabelecimento; DEV: Desenvolvimento vegetativo; FLOR: Florescimento; FRUT: Frutificação; MAT: Maturação. A) Qg: Radiação solar global; Tm: Temperatura média; Tmax: Temperatura máxima; Tmin: Temperatura mínima; WS: Velocidade do vento; WSMAX: Velocidade máxima do vento (WSMAX); WSMIN: Velocidade mínima do vento; RN: Saldo de radiação. B) P: Precipitação; RU: Umidade relativa; ETC: Evapotranspiração; STO: Armazenamento de água no solo; AET: Evapotranspiração real; DEF: Déficit hídrico; EXC: Excedente hídrico.

Tabela 1. Principais elementos climáticos correlacionados com a produtividade da soja.

Locais	Estádios fenológicos				
	Estabelecimento	**Desenvolvimento**	**Florescimento**	**Frutificação**	**Maturação**
Grupo 1	RU (+11%) e EXC (-12%)	Qg (+17%) e EXC (-22%)	Tmin (+20%) e EXC (-6%)	STO (+17%) e RU (+18%)	Qg (+15%) e EXC (-14%)
Grupo 2	RU (+17%) e DEF (-11%)	WSMIN (-15%) e RN (+21%)	RN (+27%) e EXC (-8%)	RN (+17%) e RU (+22%)	Tmax (-16%) e EXC (+13%)
Grupo 3	RU (+13%) e DEF (-9%)	Qg (+16%) e EXC (-14%)	RN (+16%) e P (-11%)	Tmax (-16%) e RU (+22)	RU (+25%) e EXC (-10%)
Grupo 4	RN (+11%) e WSMIN (-18%)	WSMIN (-20%) e RN (-17%)	RN (+33%) e DEF (-25%)	RU (+35%) e DEF (-24%)	RU (+35%) e WSMIN (-16%)

+ Correlação positiva. – Correlação negativa. RU: Umidade relativa do ar; EXC: Excedente hídrico; Qg: Radiação global; Tmin: Temperatura mínima; STO: Armazenamento de água no solo (STO), DEF: Déficit hídrico; WSMIN: Velocidade mínima do vento; RN: Saldo de radiação; Tmax: Temperatura máxima; P: precipitação.

2.4 Conclusões

Os elementos climáticos estudados possuem impacto diferente na produtividade da soja em cada estádio fenológico da cultura. Os resultados indicam que os seguintes elementos climáticos tiveram maior influência na produtividade da soja para cada grupo estudado:

Grupo 1, o EXC apresentou correlação negativa de 12 e 22%, no final dos estádios de estabelecimento e desenvolvimento, respectivamente. A Tmin apresentou correlação positiva de 20% no estádio de florescimento. O STO teve correlação positiva de 17% no estádio de frutificação. E o EXC apresentou correlação negativa de 14% no estádio de maturação.

Grupo 2, a RU mostrou correlação positiva de 17% na fase inicial do estádio de estabelecimento. O RN apresentou correlação positiva de 21 e 27%, nos estádios de desenvolvimento e florescimento, respectivamente. A RU apresentou correlação positiva de 22% no final do estádio de frutificação. E a Tmax apresentou correlação negativa de 16% no estádio de maturação.

Grupo 3, a RU apresentou correlação positiva de 13% no estádio de estabelecimento. A Qg teve correlação positiva de 16% no estádio de desenvolvimento. O RN apresentou correlação positiva de 16% no estádio de florescimento. A RU apresentou correlação positiva de 22% e 25%, no final do estádio de frutificação e no início do estádio de maturação, respectivamente.

Grupo 4, a WSMIN apresentou correlação negativa de 18% e 20%, no estádio de estabelecimento e desenvolvimento, respectivamente. O RN teve

correlação positiva de 33% no estádio de florescimento. A RU teve correlação positiva de 35% no estádio de frutificação e maturação.

Referências

Allen RG, S PL, Raes D, Martin S. (1998) Crop evapotranspiration: Guidelines for computing crop water requirements / by Richard G. Allen ... [et al.]. **FAO Irrigation and Drainage Pap** 56.:1–15

Avecilla F, Panebianco JE, Buschiazzo DE (2015) Variable effects of saltation and soil properties on wind erosion of different textured soils. **Aeolian Research**, 18:145-153. DOI: https://doi.org/10.1016/j.aeolia.2015.07.005

Battisti DS, Naylor RL (2009) Historical warnings of future food insecurity with unprecedented seasonal heat. **Science** 323:240–244. DOI: https://doi.org/10.1126/science.1164363

Battisti R, Sentelhas PC, Boote KJ (2017) Inter-comparison of performance of soybean crop simulation models and their ensemble in Southern Brazil. **Field Crops** Res 200:28–37. DOI: https://doi.org/10.1016/j.fcr.2016.10.004

Battisti R, Sentelhas PC, Pascoalino JAL, Sako H, Dantas JPS, Moraes MF (2018) Soybean Yield Gap in the Areas of Yield Contest in Brazil. **International Journal of Plant Production** 12: 159–168. DOI: https://doi.org/10.1007/s42106-018-0016-0

Bianculli ML, Aguirrezábal LAN, Irujo GAP, Echarte MM (2016) Contribution of incident solar radiation on leaves and pods to soybean seed weight and composition. **European Journal of Agronomy** 77:1-9. DOI: https://doi.org/10.1016/j.eja.2016.03.002

Board JE, Kahkin CS (2011) Soybean yield formation: what controls it and how it can be improved. In: El-Shemy HA (Ed.) **Soybean physiology and biochemistry.** Rijeka: InTech Open Access Publisher, p. 1-36.

Brandle JR, Hodges L, Zhou HX (2004). Windbreaks in North American Agricultural Systems. In: Nair PKR, Rao MR, Buck LE (Eds.) **New Vistas in Agroforestry.** Dordrecht:Springer p.65-78.

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2022) **Acompanhamento da Safra Brasileira de Grão 2021/2022:** relatório técnico n.12, décimo segundo levantamento. Disponível em: https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos. Acesso em: 10/03/2023.

Costa NL, Santana AC (2018) Análise do mercado da soja: aspectos conjunturais da formação do preço pago ao produtor brasileiro. **Revista Plantio Direto** 28: 20-39.

Cotrim MF, Gava R, Campos CNS, David CHO, Reis IA, Teodoro LPR, Teodoro PE (2021) Physiological performance of soybean genotypes grown under irrigated and rainfed conditions. **Journal of Agronomy and Crop Science** 207: 34– 43. DOI: https://doi.org/10.1111/jac.12448

Dash SSS, Lenka D, Sahoo JP, Tripathy SK, Samal KC, Lenka D, Panda RK (2022). Biochemical characterization of maize (Zea mays L.) hybrids under excessive soil moisture stress. **Cereal Research Communications**, 50-875-884. DOI: https://doi.org/10.1007/s42976-021-00241-2

de Lima RF, de Oliveira Aparecido LE, Lorençone JA, Lorençone PA, Torsoni GB, Moraes JRSC, Meneses KC (2022) Assessing life zone changes under climate change scenarios in Brazil. **Theoretical and Applied Climatology** 149: 1687–1703. DOI:https://doi.org/10.1007/s00704-022-04133-1

Djanaguiraman M, Schapaugh W, Fritschi F, Nguyen H, Prasad PVV (2019) Reproductive success of soybean (Glycine max L. Merril) cultivars and exotic lines under high daytime temperature. **Plant, Cell e Environment**. 42:321–336. DOI: https://doi.org/10.1111/pce.13421

Duarte YCN, Sentelhas PC (2020) NASA/POWER and daily gridded weather datasets - how good they are for estimating maize yields in Brazil? **International Journal of Biometeorology** 64:319–329. DOI: https://doi.org/10.1007/s00484-019-01810-1

Elmerich C, Boulch G, Faucon MP, Lakhal L, Lange B (2023) Identification of Eco-Climatic Factors Driving Yields and Genotype by Environment Interactions for Yield in Early Maturity Soybean Using Crop Simulation. **Agronomy** 13: 322. DOI: https://doi.org/10.3390/agronomy13020322

EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. (2013) **Tecnologia de produçao de soja: Região Central do Brasil 2014**. Londrina: Embrapa Soja, 266 p.

Farias JRB, Neumaier N, Nepomuceno AL (2009) Soja. In: Monteiro JEBA (Ed.) **Agrometeorologia dos cultivos: o fator meteorológico na produção agrícola.** Brasília: INMET, p. 261–278.

Gardiner B, Berry P, Moulia B (2016) Review: Wind impacts on plant growth, mechanics and damage. **Plant Science** 245:94–118. DOI: https://doi.org/10.1016/j.plantsci.2016.01.006

Gilbert ME, Holbrook NM, Zwieniecki MA, Sadok W, Sinclair TR (2011). Field confirmation of genetic variation in soybean transpiration response to vapor pressure deficit and photosynthetic compensation. **Field Crops Research** 124: 85–92. DOI: https://doi.org/10.1016/j.fcr.2011.06.011

IBGE (Instituto Brasileiro de Geografia e Estatística), (2020) Sistema IBGE de Recuperação Automática - SIDRA: Produção Agrícola Municipal. Disponível em: https://sidra.ibge.gov.br/pesquisa/ppm/quadros/brasil/2020. (Acesso em 10/09/2020).

Li Q, Fu C, Liang C, Ni X, Zhao X, Chen M, Ou L (2022) Crop Lodging and The Roles of Lignin, Cellulose, and Hemicellulose in Lodging Resistance. **Agronomy** 12:1795. DOI:https://doi.org/10.3390/agronomy12081795

Lin X, Wu B, Wang J, Wang G, Chen Z, Liang Y, Liu J, Wang H (2023) Effects of Geographical and Climatic Factors on the Intrinsic Water Use Efficiency of Tropical Plants: Evidence from Leaf ^{13}C. **Plants** 12:951. DOI: https://doi.org/10.3390/plants12040951

Lopez MA, Moreira FF, Hearst A, Cherkauer K, Rainey KM (2022) Physiological breeding for yield improvement in soybean: solar radiation interception-conversion, and harvest index. **Theoretical and Applied Genetics** 135:1477-1491. DOI: https://doi.org/10.1007/s00122-022-04048-5

Lorençone PA, Oliveira LEA, Lorençone JA, Torsoni GB, Lima RF (2022) Estimation of Air Temperature Using Climate Factors in Brazilian Sugarcane Regions. Revista Brasileira de Meteorologia 37:121-140. DOI: http://dx.doi.org/10.1590/0102-7786371000

Mandić V, Krnjaja V, Tomić Z, Bijelić Z, Simić A, Đorđević S, Stanojković A, Gogić M (2015) Effect of water stress on soybean production. In.: PROCEEDING OF THE 4TH INTERNATIONAL CONGRESS OF NEW PERSPECTIVES AND CHALLENGES OF SUSTAINABLE LIVESTOCH PRODUCTION. **Anais ...** Belgrado, Sérvia: p. 405-414.

Menezes CT, Casaroli D, Heinemann AB, Moschetti VC, Battisti R (2022) The impact of gridded weather database on soil water availability in rice crop modeling. **Theoretical and Applied Climatology** 147:1401–1414. DOI: https://doi.org/10.1007/s00704-021-03906-4

Monteiro LA, Sentelhas PC, Pedra GU (2018) Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. **Internation Journal of Climatology** 38:1571-1581. DOI: https://doi.org/10.1002/joc.5282

Monteiro JEBA (2009) Agrometeorologia dos cultivos: o fator meteorológico na produção agrícola. Brasília: **Instituto Nacional de Meteorologia (Brasil)**, p. 261–278.

Moreira VS, Candido AL, Mota CM, Webler G, Oliveira EP, Roberti DR (2023) Impacts of climate change on water fl uxes and soybean growth in southern Brazil. **Revista Ciência Agronômica** 54:1-12. DOI: https://doi.org/10.5935/1806-6690.20230014

Mukherjee S, Nandi R, Kundu A, Bandyopadhyay PK, Nalia A, Ghatak P, Nath R (2022) Soil water stress and physiological responses of chickpea (Cicer arietinum L.) subject to tillage and irrigation management in lower Gangetic plain. **Agricultural Water Management** 263:107443 DOI: https://doi.org/10.1016/j.agwat.2021.107443

Pelech EA, Evers JB, Pederson TL, Drag DW, Fu P, Bernacchi CJ (2023) Leaf, plant, to canopy: A mechanistic study on aboveground plasticity and plant density within a maize–soybean intercrop system for the Midwest, USA. **Plant, Cell e Environment** 46:405–421. DOI: https://doi.org/10.1111/pce.14487

Pereira AR, Angelocci LR, Sentelhas PC (2002) **Agrometeorologia: Fundamentos e Aplicações Práticas.** Porto Alegre: Guaíba Agropecuária p. 478.

Phukan UJ, Mishra S, Shukla RK (2016) Waterlogging and submergence stress: affects and acclimation. **Critical Reviews in Biotechnology** 36:956–966. DOI: https://doi.org/10.3109/07388551.2015.1064856

Pierdoná HL, Paula ACP, Silva CL, Warren MS, Rodrigues LN (2019) Comparação entre dois modelos usados na estimativa da evapotranspiração real da cultura da soja. **Revista Engenharia na Agricultura**. 27:431-439. DOI: https://doi.org/10.13083/reveng.v27i5.877

Purcell LC, Ball RA, Reaper JD, Vories ED (2002) Radiation use efficiency and biomass production in soybean at diferente plant population densities. **Crop Science** 42:172-177. DOI: https://doi.org/10.2135/cropsci2002.1720

Ramteke R, Gupta GK, Singh DV (2015) Growth and Yield Responses of Soybean to Climate Change. **Agricultural Research** 4: 319–323. DOI: https://doi.org/10.1007/s40003-015-0167-5

Reis L, Santos e Silva CM, Bezerra B, Mutti P, Spyrides MH, Silva P, Magalhães T, Ferreira R, Rodrigues D, Andrade L (2020) Influence of Climate Variability on Soybean Yield in MATOPIBA, Brazil. **Atmosphere** 11:1130. DOI: https://doi.org/10.3390/atmos11101130

Reis LC, Silva CMS, Bezerra BG, Mutti PR, Spryrides MHC, Silva PE (2020) Analysis of Climate Extreme Indices in the MATOPIBA Region, Brazil. **Pure Applied Geophysics** 177: 4457–4478. DOI: https://doi.org/10.1007/s00024-020-02474-4

Rosenberg NJ (1966) Microclimate, air mixing and physiological regulation of transpiration as influenced by wind shelter in an irrigated bean field. **Agricultural Meteorology** 3:197-224. DOI:https://doi.org/10.1016/0002-1571(66)90029-X

Schirmbeck LW, Fontana DC, Dalmago AG, Schirmbeck J, Vargas PR, Fernandes JMC (2019) Condições hídricas de lavoura de soja usando sensoriamento remoto terrestre. **Agrometeoros** 27:173-181. DOI: http://dx.doi.org/10.31062/agrom.v27i1.26565

Sentelhas PC, Battisti R, Câmara GMS, Farias JRB, Hampf AC, Nendel, C (2015) The soybean yield gap in Brazil – magnitude, causes and possible solutions for sustainable production. **The Journal of Agricultural Science** 153:1394-1411 DOI: https://doi.org/10.1017/S0021859615000313

Silva DS, Arima EY, dos Reis, TNP, Rattis L (2023) Temperature effect on Brazilian soybean yields, and farmers' responses. **International Journal of Agricultural Sustainability** 21:1-15. DOI: https://doi.org/10.1080/14735903.2023.2173370

Sivakumar M (2023) Importance of solar radiation and the need for improved respect to Sun by Agrometeorologists, **Journal of Agrometeorology** 25: 51–60. DOI: https://doi.org/10.54386/jam.v25i1.1971.

Staniak M, Szpunar-Krok E, Kocira A (2023) Responses of Soybean to Selected Abiotic Stresses—Photoperiod, Temperature and Water. **Agriculture** 13:146. DOI: https://doi.org/10.3390/agriculture13010146

Taiz L, Zeiger E (2017) **Fisiologia vegetal**. Porto Alegre: Artmed, p.888.

Thornthwaite CW, Mather JR (1955) The water balance. **Drexel Institute of Technology**, Laboratory of Climatology, Centerton, NJ, p 104

Tonello CZ, Machado BO, Maurer TR, Lângaro NC (2022) Physiological performance of green soybean seeds: a systematic review. **Conjecturas** 22: 749–758. DOI:https://doi.org/10.53660/CONJ-2056-2S67

USDA - United States Department of Agriculture Foreign. Oilseeds: World Markets and Trade. PSD 3/2023. Disponível em: https://usda.library.cornell.edu/concern/publications/tx31qh68h?locale=en. Acesso em: 12/03/2023.

Villa B, Petry MT et. al. (2022) Water Deficit and Excesss and the Main Physiological Disorders in Agricultural Crops. **Journal of Agricultural Science** 14:1-7. DOI: https://doi.org/10.5539/jas.v14n8p166

Wu C, Mozzoni LA, Moseley D, Hummer W, Ye H, Chen P, Nguyen H (2020) Genome-wide association mapping of flooding tolerance in soybean. **Molecular Breeding** 40:1-14. DOI: https://doi.org/10.1007/s11032-019-1086-0

Wu J, Wang J, Hui W, Zhao F, Wang P, Su C, Gong W (2022). Physiology of plant responses to water stress and related genes: A review. **Forests** 13(2):324. DOI: https://doi.org/10.3390/f13020324

Zhang S, Liu G, Cui Q, Huang Z, Ye X, Cornelissen JHC (2021) New field wind manipulation methodology reveals adaptive responses of steppe plants to increased and reduced wind speed. **Plant Methods** 17:1-16. DOI: https://doi.org/10.1186/s13007-020-00705-2

Zhang T, Xu X, Jiang H, Qiao S, Guan M, Huang Y, Gong R (2022) Widespread decline in winds promoted the growth of vegetation**. Science of The Total Environment**. 825:153682. DOI:https://doi.org/10.1016/j.scitotenv.2022.153682

Zhang Z, Li T, Guo E, Zhao C, Zhao J, Liu Z, Sun S, Zhang F, Guo S, Nie J, Yang X (2022) 20% of uncertainty in yield estimates could be caused by the radiation source. **Science of The Total Environment**. 838:156015. DOI: https://doi.org/10.1016/j.scitotenv.2022.156015

Zilli M, Scarabello M, Soterroni AC, Valin H, Mosnier A, Leclère D, Havlík P, Kraxner F, Antonio M, Ramos FM (2020) The impact of climate change on Brazil's agriculture. **Science of the Total Environment** 740:139384. DOI: https://doi.org/10.1016/j.scitotenv.2020.139384

CAPÍTULO 3 – VARIÁVEIS CLIMÁTICAS PARA ESTIMAÇÃO E PREDIÇÃO DE PRODUTIVIDADE DE SOJA UTILIZANDO APRENDIZADO DE MÁQUINA

RESUMO - A produtividade da soja é sensível à dinâmica dos elementos climáticos. Atualmente o Brasil é o maior produtor mundial de grãos de soja. Partindo deste princípio, técnicas de estimação e predição de produtividade utilizando algoritmos podem ser uma alternativa para apoio nas tomadas de decisão sobre o mercado futuro e planejamento agrícola. O objetivo deste estudo foi avaliar o desempenho de quatro algoritmos de estimação e predição de produtividade de soja em 154 locais com base em elementos climáticos. Para este estudo, utilizamos uma série histórica de 30 anos (1988-2018), que se iniciou de acordo com o ciclo da cultura, estabelecido entre os decêndios 31 a 9. Os dados de produtividade de soja foram coletados na base SIDRA/IBGE, e os dados meteorológicos são da plataforma NASA/POWER. Para a seleção dos locais estudados foi realizado uma seleção dos locais que apresentaram maior área cultiva e maior produtividade de soja no Brasil. Foi realizado análise de cluster para agrupar os locais de acordo com as suas similaridades climáticas, resultando em quatro grupos (grupo 1, 2, 3 e 4). Para a estimação e predição avaliamos quatro algoritmos: Regressão Linear Mútipla (RLM), Random Forest (RF), Support Vector Machine (SVM) e Multilayer Perceptron (MLP). Os dados foram separados em 70% para treino e 30% para teste. Para avaliação dos modelos foi utilizado R^2 ajustado (R^2 adj), erro quadrático médio (RMSE) e erro médio absolute (MAE). Todas as análises foram realizadas no software Python 3.8. O RF apresentou o melhor desempenho para o grupo 4, com R^2 adj de 0,76 e RMSE de 389,41 kg ha^{-1}. O RLM apresentou o menor desempenho para grupo 4, com R^2 adj de 0,01 e RMSE de 3993,8 kg ha^{-1}. Portanto, com os resultados do estudo concluímos que é possível estimar e predizer a produtividade de soja utilizando algoritmos de aprendizado de máquina. O RF foi o modelo que apresentou o melhor ajuste de precisão e acurácia para estimar e predizer a produtividade de soja no Brasil.

Palavras-chave: Agrometeorologia, Random forest, Regressão linear múltipla, Multilayer perceptron, Suport vector machine

CLIMATE VARIABLES FOR ESTIMATION AND PREDICTION SOYBEAN PRODUCTIVITY USING MACHINE LEARNING

ABSTRACT - Soybean productivity is sensitive to the dynamics of climatic elements. Brazil is currently the world's largest producer of soybeans. Based on this principle, productivity estimation and prediction techniques using algorithms can be an alternative to support decision-making about the future market and agricultural planning. The objective of this study was to evaluate the performance of four soybean productivity prediction algorithms in 154 locations based on climatic elements. For this study, we used a 30-year historical series (1988-2018), which began according to the crop cycle, established between decades 31 and 9. Soybean productivity data were collected in the SIDRA/IBGE database, and meteorological data are from the NASA/POWER platform. To select the sites studied, a selection was made of the sites that presented the largest cultivated area and highest soybean production in Brazil. Cluster analysis was performed to group the locations according to their climatic similarities, resulting in four groups (groups 1, 2, 3 and 4). For prediction, we evaluated four algorithms: Multiple Linear Regression (RLM), Random Forest (RF), Support Vector Machine (SVM) and Multilayer Perceptron (MLP). The data was separated into 70% for training and 30% for testing. To evaluate the models, adjusted R^2 (R^2 adj), mean squared error (RMSE) and mean absolute error (MAE) were used. All analyzes were performed using Python 3.8 software. RF presented the best performance for group 4, with R^2 adj of 0.76 and RMSE of 389.41 kg ha^{-1}. The RLM presented the lowest performance for group 4, with R^2 adj of 0.01 and RMSE of 3993.8 kg ha^{-1}. Therefore, with the results of the study we conclude that it is possible to estimation and predict soybean productivity using machine learning algorithms. The RF was the model that presented the best adjustment of precision and accuracy to estimation and predict soybean productivity in Brazil.

Keywords: Agrometeorology, Random forest, Multiple linear regression, Multilayer perceptron, Suport vector machine

3.1 Introdução

A soja (*Glycine max* L.) é a principal oleaginosa cultivada no mundo, devido ao seu alto valor protéico, podendo ser utilizada tanto para alimentação humana quanto animal (Medeiros, 2016). Possui grande importância econômica, atualmente o Brasil é o maior produtor mundial de grãos de soja, para a safra 2021/2022, prevê uma produção total de 122,43 milhões de toneladas, um aumento na produção de 11,4% em relação ao ciclo anterior (CONAB, 2022).

Dada a importância da soja em cenário nacional, atenção deve ser dada aos parâmetros climáticos, sendo este um fator que influência de forma direta e indireta a produtividade de soja. Os elementos que mais afetam o seu desenvolvimento e produtividade são: temperatura, fotoperíodo e disponibilidade hídrica (Farias, 2007; Sentelhas, 2015).

No Rio Grande do Sul houve perdas de produtividade de 41,7% na safra de 2019/2020, devido aos baixos índices pluviométricos (CONAB, 2020). Os estádios fenológicos da soja mais afetados pela disponibilidade de água são germinação e emergência, florescimento e enchimento de grãos (Tagliapietra et al., 2021; Farias, 2007). Portanto, algoritmos de predição com bases climatológicas podem ser utilizados para apoio à tomada de decisão do produtor (Moraes, 2020; Morelli, 2021).

Vários estudos tem demostrado o potencial de precisão e acurácia de algoritmos de predição de produtividade (Aparecido et al.,2020; Beisekenov et al., 2022; Santos et al., 2022). O conhecimento prévio da produtividade auxilia os produtores na tomada de decisão da safra futura, evitando o aumento de custos e perdas no processo de colheita (Aparecido et al., 2022).

Os modelos de regressão linear múltipla (RLM) são amplamente utilizados na agricultura para predição de safras agrícolas (Portugal et al., 2018; Aparecido et al., 2022). A RLM é uma técnica multivariada capaz de obter uma relação matemática que estima e prevê resultados por meio da regra estatística dos mínimos quadrados, além disso, permiti a detecção de padrões de similaridade e correlação entre as variáveis (Werkena e Aguiar, 1996; Branco e Pires, 2017).

O Random Forest (RF) é uma técnica de aprendizado de máquina que vem sendo muito utilizada para predição de produtividade agrícola, é um algoritmo rápido para treinar e prever, além de poder ser utilizado em problemas complexos, possui detecção de outliers e é uma técnica de aprendizagem não supervisionada (Cutler e Stevens, 2012).

O Suport Vector Machine (SVM) é um algoritmo avançado da técnica de aprendizado de máquina amplamente utilizado na agricultura (Santos et al., 2022). O modelo foi estabelecido por Cortes e Vapnik (1995), a ideia principal deste algoritmo é a minimização de risco estrutural, que fornece melhor generalização e reduz o erro na fase de treinamento (Tanveer et al., 2022).

A rede neural Multilayer Perceptron (MLP) é um algoritmo de aprendizado supervisionado, inspirado nas redes neurais biológicas (Haykin, 1999). Essa técnica altera os pesos das conexões entre os neurônios repetidas vezes até que o erro, seja o menor possível (Sudheer et al., 2002). Com base nos valores das variáveis preditoras o algoritmo é capaz de produzir um modelo preditivo para uma ou mais variáveis dependentes (Santos et al., 2022).

Nossos resultados poderão sugerir qual algoritmo possui melhor desempenho, precisão e acurácia para estimação e predição de produtividade de soja para as regiões brasileiras estudadas. Além disso, os resultados desse estudo são importantes para tomada de decisão do produtor, planejamento agrícola, mercado de *commodities* e mercado futuro. Neste estudo, o objetivo principal é avaliar o desempenho de quatro algoritmos de estimação e predição de produtividade de soja em 154 locais com base em parâmetros climáticos. Os algoritmos avaliados foram o RLM, RF, SVM e MLP.

3.2 Material e métodos

3.2.1 Áreas de estudo e produtividade

O estudo foi realizado em 154 localidades produtoras de soja no Brasil (Tabela 1). Os critérios para a seleção destas áreas foram: as de maiores áreas semeadas e produtividades. As localidades estão distribuídas entre todas as regiões do país (norte, sul, nordeste, sudeste e centro-oeste) (Figura 1).

Os dados de produtividade de soja são de uma série histórica de 30 anos (1988-2018), provenientes do PAM (Produção Agrícola Municipal) disponibilizados pelo IBGE (https://sidra.ibge.gov.br/pesquisa/pam/tabelas).

Tabela 1. Localidades produtoras de soja estudadas.

Município	Estado
Baianópolis	Bahia
Barreiras	Bahia
Correntina	Bahia
Formosa do Rio Preto	Bahia
Riachão das Neves	Bahia
São Desidério	Bahia
Brasília	Distrito Federal

Bom Jesus de Goiás	Goiás
Edéia	Goiás
Goianésia	Goiás
Goiatuba	Goiás
Itumbiara	Goiás
Joviânia	Goiás
Maurilândia	Goiás
Morrinhos	Goiás
Panamá	Goiás
Quirinópolis	Goiás
Rio Verde	Goiás
Santa Helena de Goiás	Goiás
Alto Parnaíba	Maranhão
Balsas	Maranhão
Fortaleza dos Nogueiras	Maranhão
Riachão	Maranhão
Sambaíba	Maranhão
São Raimundo das Mangabeiras	Maranhão
Tasso Fragoso	Maranhão
Alto Garças	Mato Grosso
Canarana	Mato Grosso
Chapada dos Guimarães	Mato Grosso
Diamantino	Mato Grosso
General Carneiro	Mato Grosso
Guiratinga	Mato Grosso
Itiquira	Mato Grosso
Jaciara	Mato Grosso
Juscimeira	Mato Grosso
Nobres	Mato Grosso
Poxoréu	Mato Grosso
Rondonópolis	Mato Grosso
Santo Antônio do Leverger	Mato Grosso
Torixoréu	Mato Grosso
Pedra Preta	Mato Grosso
Amambai	Mato Grosso do Sul
Aral Moreira	Mato Grosso do Sul
Bela Vista	Mato Grosso do Sul
Caarapó	Mato Grosso do Sul
Costa Rica	Mato Grosso do Sul
Dourados	Mato Grosso do Sul
Jardim	Mato Grosso do Sul
Maracaju	Mato Grosso do Sul
Pedro Gomes	Mato Grosso do Sul
Ponta Porã	Mato Grosso do Sul

Rio Brilhante	Mato Grosso do Sul
Sidrolândia	Mato Grosso do Sul
São Gabriel do Oeste	Mato Grosso do Sul
Terenos	Mato Grosso do Sul
Buritis	Minas Gerais
Campo Florido	Minas Gerais
Campos Altos	Minas Gerais
Capinópolis	Minas Gerais
Indianópolis	Minas Gerais
Ipiaçu	Minas Gerais
Itaporã	Minas Gerais
Ituiutaba	Minas Gerais
Monte Carmelo	Minas Gerais
Nova Ponte	Minas Gerais
Presidente Olegário	Minas Gerais
Santa Juliana	Minas Gerais
São Gonçalo do Abaeté	Minas Gerais
Tupaciguara	Minas Gerais
Uberlândia	Minas Gerais
Unaí	Minas Gerais
Paragominas	Para
Redenção	Para
Santarém	Para
Arapoti	Paraná
Boa Esperança	Paraná
Cascavel	Paraná
Castro	Paraná
Catanduvas	Paraná
Corbélia	Paraná
Céu Azul	Paraná
Formosa do Oeste	Paraná
Guarapuava	Paraná
Ivatuba	Paraná
Mangueirinha	Paraná
Nova Aurora	Paraná
Nova Cantu	Paraná
Pinhão	Paraná
Tibagi	Paraná
Alvorada do Gurguéia	Piauí
Antônio Almeida	Piauí
Baixa Grande do Ribeiro	Piauí
Bom Jesus	Piauí
Gilbués	Piauí
Monte Alegre do Piauí	Piauí

Ribeiro Gonçalves	Piauí
Santa Filomena	Piauí
Sebastião Leal	Piauí
Uruçuí	Piauí
Anta Gorda	Rio Grande do Sul
Casca	Rio Grande do Sul
Chiapetta	Rio Grande do Sul
Colorado	Rio Grande do Sul
Condor	Rio Grande do Sul
Espumoso	Rio Grande do Sul
Getúlio Vargas	Rio Grande do Sul
Ibirubá	Rio Grande do Sul
Lagoa Vermelha	Rio Grande do Sul
Não-Me-Toque	Rio Grande do Sul
Panambi	Rio Grande do Sul
Pejuçara	Rio Grande do Sul
Selbach	Rio Grande do Sul
Serafina Corrêa	Rio Grande do Sul
São José do Ouro	Rio Grande do Sul
Tapera	Rio Grande do Sul
Vilhena	Roraima
Capinzal	Santa Catarina
Catanduvas	Santa Catarina
Erval Velho	Santa Catarina
Faxinal dos Guedes	Santa Catarina
Ibicaré	Santa Catarina
Irani	Santa Catarina
Maravilha	Santa Catarina
Nova Erechim	Santa Catarina
Palma Sola	Santa Catarina
Papanduva	Santa Catarina
Piratuba	Santa Catarina
Ponte Serrada	Santa Catarina
Quilombo	Santa Catarina
São Domingos	Santa Catarina
Xaxim	Santa Catarina
Assis	São Paulo
Barretos	São Paulo
Bernardino de Campos	São Paulo
Colina	São Paulo
Colômbia	São Paulo
Cândido Mota	São Paulo
Florínea	São Paulo
Guariba	São Paulo

Guaíra	São Paulo
Igarapava	São Paulo
Jaboticabal	São Paulo
Maracaí	São Paulo
Palmital	São Paulo
Platina	São Paulo
São Joaquim da Barra	São Paulo
Almas	Tocantins
Alvorada	Tocantins
Brejinho de Nazaré	Tocantins
Dianópolis	Tocantins
Figueirópolis	Tocantins
Formoso do Araguaia	Tocantins
Pedro Afonso	Tocantins
Peixe	Tocantins

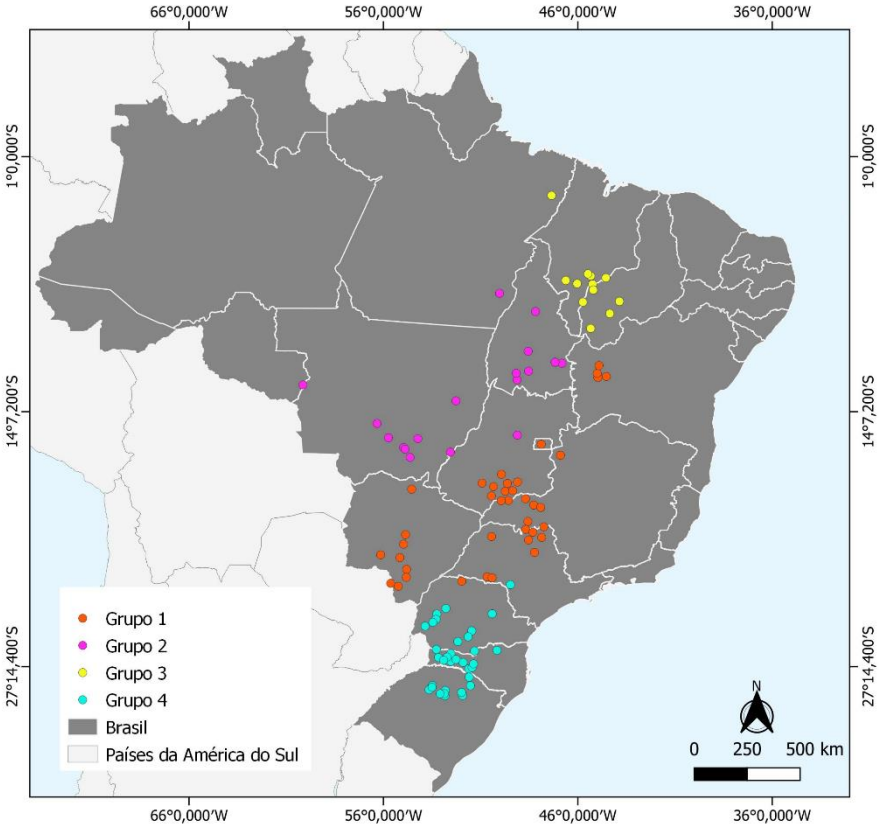

Figura 1. Localização espacial das localidades produtoras de soja estudadas.

3.2.2 Coleta de dados climáticos

Os dados climáticos são provenientes da plataforma *National Aeronautics and Space Administration/Prediction of Worldwide Energy Resources* (NASA/POWER) em escala diária durante o período de (1988–2018) e agregados em escala de 10 dias (Stackhouse et al., 2016). Os dados coletados foram: precipitação (mm), temperatura máxima e mínima do ar (°C), radiação solar global (MJ m^2 dia^{-1}), umidade relativa do ar (%) e velocidade do vento (m s^{-1}).

Os dados fornecidos pela plataforma NASA/POWER são baseados em modelos e informações de satélite suficientemente precisos para fornecer dados confiáveis de medições solares e meteorológicas para regiões onde as medições de superfície são escassas ou inexistentes (Monteiro et al., 2018).

3.2.3 Evapotranspiração potencial

A evapotranspiração de potencial (PET) foi calculada pelo método de Penman e Monteith (Allen et al., 1998), descrito pelas equações 1 a 4.

$$PET = \frac{0.408 \times s \times (Rn - G) + \frac{\gamma \times 900 \times u_2 \times (e_s - e_a)}{T+273}}{s + \gamma \times (1 + 0.34 \times u_2)} \quad (1)$$

$$s = \frac{4098 \times e_s}{(T+273)^2} \quad (2)$$

$$e_a = \frac{RH \times e_s}{100} \quad (3)$$

$$e_s = 0.6108 \times e^{\frac{17.27 \times T}{237.3+T}} \quad (4)$$

em que, s é a declividade da curva de pressão de vapor em relação à temperatura (kPa °C^{-1}), Rn é o saldo de radiação diária (MJm^{-2} dia^{-1}), G é o fluxo total diário de calor no solo (MJm^{-2} dia^{-1}), T é a temperatura média do ar (°C), γ é a constante psicométrica (0.063 kPa ºC^{-1}), RH é a umidade relativa (%), u_2 é a velocidade do vento a 2 m de altura (m s^{-1}), e_a é a pressão atual de vapor (kPa), e_s é a pressão de saturação de vapor (kPa).

3.2.4 Balanço hídrico climatológico

O balanço hídrico climatológico foi calculado pelo método de Thornthwaite e Mather (1955). Foi utilizado para todos os municípios a capacidade de água disponível (CAD) de 100 mm, seguindo as sugestões de Duarte e Sentelhas (2020) para soja (Figura 2).

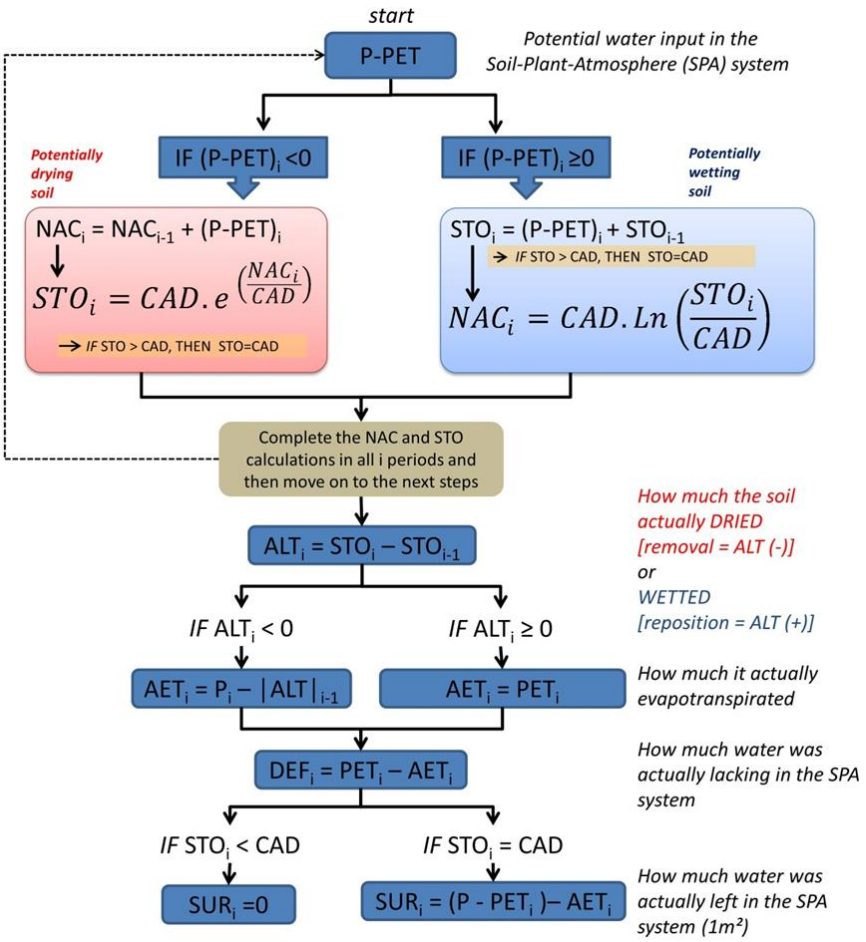

Figura 2. Fluxograma do modelo de balanço hídrico Thornthwaite e Mather (1955). Adaptado por Rolim et al. (2020).

Em que, P a precipitação (mm), PET a evapotranspiração potencial (mm), NAC o valor de P-PET, i é o mês atual, STO é o armazenamento de água no solo (mm), CAD é a capacidade de água disponível (mm), ALT é a alteração do armazenamento de água no solo entre o mês anterior e o atual (mm), AET é a evapotranspiração real (mm), DEF é o déficit hídrico no solo (mm^{-1}), e SUR é o excedente hídrico no solo (mm).

3.2.5 Agrupamento dos locais

As localidades foram agrupadas entre aquelas que possuem similaridade em relação as características climáticas estudas. A classificação

hierárquica dos grupos foi realizada pela análise de cluster e construído um dendrograma, onde foi possível identificar quatro grupos diferentes (Figura 3).

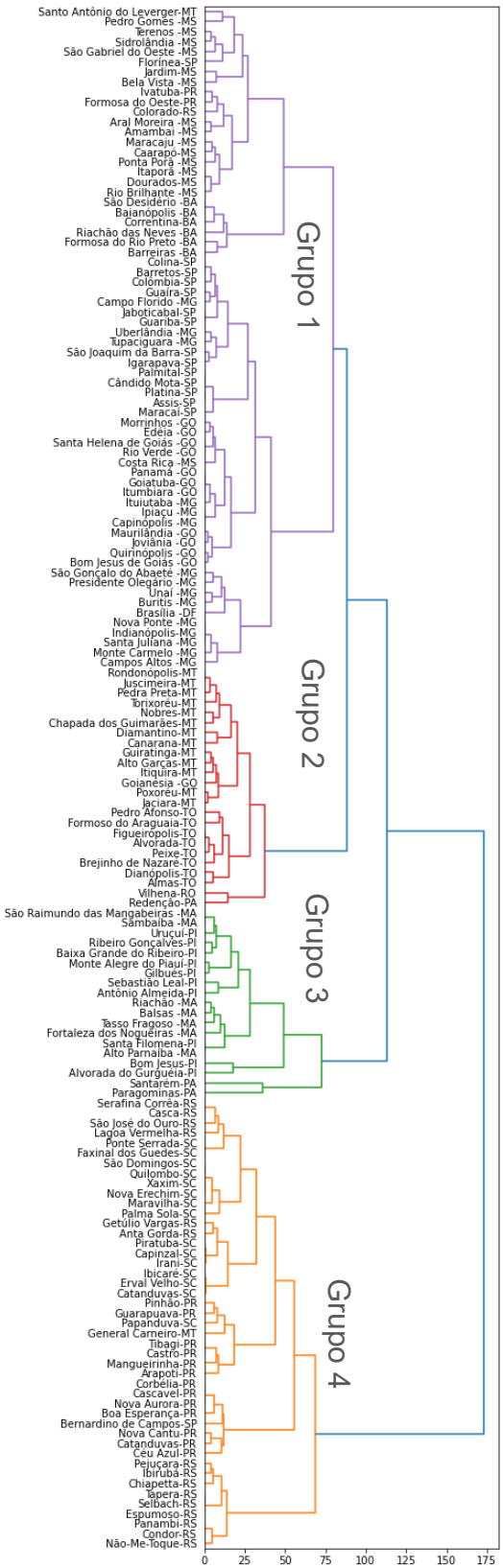

Figura3**.** Dendrograma para identificação dos grupos em relação as condições de similaridade climática dos locais.

3.2.6 Parâmetros do cultivo

Os dados de entrada dos modelos foram ajustados de acordo com o ciclo fenológico da soja. Para este estudo foi considerado a soja de ciclo médio, com o ciclo fenológico estabelecido entre os decêndios 31 e 9, foram utilizados os parâmetros: coeficiente de cultivo (kc), que expressa a relação entre a evapotranspiração máxima do cultivo e a evapotranspiração de referência; o fator de resposta à produtividade (ky), que diz respeito à resposta da vegetação à disponibilidade hídrica e descreve a diminuição relativa da produtividade, em função da diminuição da evapotranspiração (Doorenbos e Kassam, 1979).

3.2.7 Calibração dos modelos

Para estimativa e predição de produtividade da soja utilizamos quatro algoritmos: Regressão Linear Múltipla (RLM), Random Forest (RF), Suport Vector Machine (SVM) e Multilayer Perceptron (MLP). Na Figura 4 pode ser observado o fluxograma de desenvolvimento dos modelos.

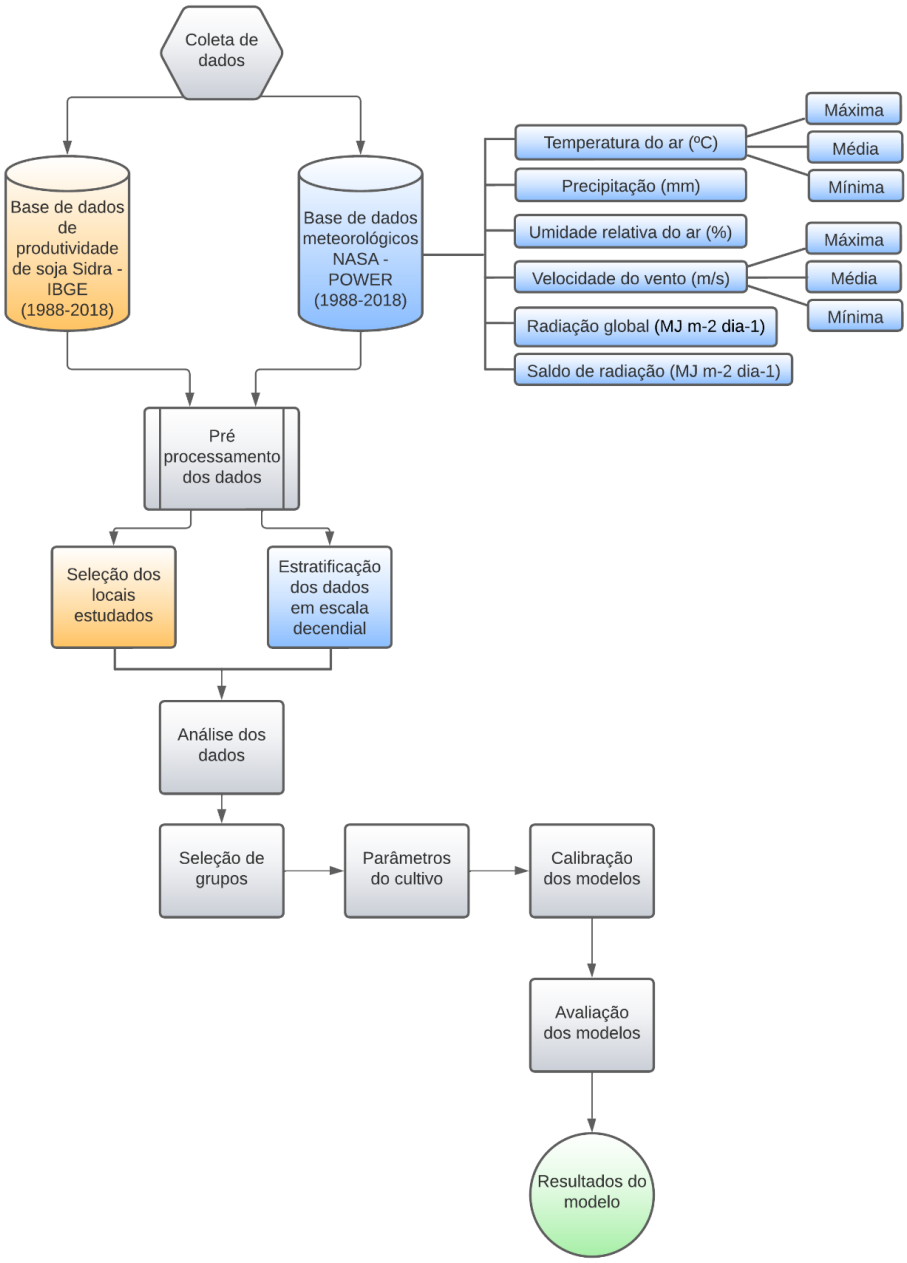

Figura 4. Fluxograma de desenvolvimento dos modelos estudados.

A RLM representa a relação linear entre muitas variáveis dependentes e independentes (Hoffmann, 2016). Por outro lado, utilizamos o modelo RF, que é capaz de lidar de forma mais robusta com não linearidades e contra o ajuste

excessivo (overfitting), além de possuir alta eficiência computacional (Gill et al., 2022).

O SVM busca definir uma margem de tolerância para que encontre o hiperplano ótimo para que o erro seja minimizado (Basak et al., 2007) com diferentes funções de kernel (Zhang et al., 2004). A rede neural MLP é capaz de distinguir dados que não são linearmente separáveis, o erro é distribuído na rede e ajusta os pesos entre os neurônios (Hykin et al., 1999).

Todos os parâmetros dos modelos foram ajustados para produzir os melhores resultados (Tabela 2). Para ambos os modelos testados os dados de entrada foram padronizados. Utilizamos também a separação do conjunto de dados em dois conjuntos: um conjunto de treinamento com 70% dos dados e um conjunto de teste com 30% dos dados (Kohavi, 1995).

Utilizamos o método stepwise forward para a seleção das melhores variáveis para o modelo de RLM. Este método seleciona as variáveis independentes eliminando a colinearidade entre elas (Chen e Xie, 2014).

Tabela 2. Parâmetros dos modelos.

Modelo	Parâmetro	Valores testados	Valores Ajustados			
			Grupo 1	Grupo 2	Grupo 3	Grupo 4
RLM	coeficientes angulares ou pesos (seleção de variáveis por stepwise)		*	*	*	*
RF	max_depth	50	50	50	50	50
	n_estimators	100	100	100	100	100
	max_features	315	315	315	315	315
SVM linear	gamma	auto; scale	auto	auto	auto	auto
	c	100,151,152	100	100	100	100
SVM Polinomial	degree	2;3;4	2	2	2	2
	gamma	auto; scale	auto	auto	auto	auto
	c	100,151,152	100	100	100	100
SVM-RBF	gamma	auto; scale	auto	auto	auto	auto
	c	100,151,152	100	100	100	100
MLP	camadas	12,8,22	8	8	8	8
	random_state	20	20	20	20	20
	max_iter	500	500	500	500	500

* Verificar no texto.

3.2.8 Avaliação de desempenho dos modelos

Para avaliar o desempenho dos modelos utilizamos índices estatísticos de precisão e acurácia. Para avaliar a precisão foi utilizado o coeficiente de determinação ajustado (R^2 ajustado) (Eq.1).

$$R^2\ ajustado = 1 - \left[\frac{(1-R^2)(n-1)}{n-k-1}\right] \quad (1)$$

em que, R^2 é o coeficiente de determinação ajustado, n é o número de elementos e k é o número de variáveis independentes da regressão.

Para avaliar a acurácia foi utilizado o RMSE (Raiz Quadrada do Erro Médio) (Eq. 2), e o MAE (Erro Médio Absoluto) (Eq. 3).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Y_{obs_i} - Y_{est_i}\right)^2}{N}} \quad (2)$$

$$MAE = \frac{\sum_{i=1}^{N}|Yobs_i - Yest_i|}{N} \quad (3)$$

em que, $Yest$ é o valor estimado de y; $Yobs$ é o valor observado de y; e N é o número de elementos. Todas as análises foram feitas com a linguagem de programação python versão 3.9 (https://www.python.org/).

3.3 Resultados e discussão

Os locais do grupo 2, apresentaram a maior média de precipitação 78,79 mm. O grupo 1 apresentou média de 62,97 mm, no grupo 3 foi 55,77 mm, e o grupo 4 apresentou a menor média 51,97 mm (Figura 5A). A Temperatura média do grupo 1 foi 24,75 ºC, a do grupo 2 foi 25,49 ºC, e no grupo 4 foi 21,97 ºC, os locais agrupados no grupo 3 apresentaram as maiores médias de temperatura (26,34 ºC), esses locais concentram-se maior parte na região nordeste do país, onde as temperaturas geralmente são mais elevadas (Figura 5B). Vale ressaltar que todos os grupos apresentaram temperatura média dentro dos limites ideias para o desenvolvimento da soja, entre 20ºC e 30ºC (Farias, 2007). Resultados semelhantes foram encontrados por Silva (2021), Reis et al., (2022) e Souza et al., (2022).

Os locais do grupo 1 e grupo 4 apresentaram as maiores médias de radiação solar global: 20,37 MJ m^{-2} dia^{-1} e 20,88 MJ m^{-2} dia^{-1}, respectivamente. O grupo 2 e 3 apresentaram as menores médias, 18,64 MJ m^{-2} dia^{-1} e 18,12 MJ m^{-2} dia^{-1}, respectivamente (Figura 5C). Resultados semelhantes foram encontrados por Oliveira et al., (2014) e Auler e Minuzzi (2022). De acordo com Tsukahara (2016), as maiores perdas de produtividade de soja ocorrem em ambientes que apresenta altos índices de precipitação, temperatura elevada e altos índices de radiação solar global.

O grupo 2 apresentou a maior média de evapotranspiração potencial 23,47 mm. O grupo 3 apresentou a segunda maior média 22,90 mm, e os grupos 1 e 4, apresentaram as menores médias 21,89 mm e 21,09 mm, respectivamente (Figura 5D). Os valores de evapotranspiração são maiores no estádio reprodutivo das plantas de soja, consequentemente demandam maior requerimento hídrico (Segal, 2004; Curto et al., 2019).

O grupo 1 apresentou a maior média de velocidade do vento para o 1,73 m s^{-1}. O grupo 3 apresentou a segunda maior média 1,62 m s^{-1}, e o grupo 2 e 4 apresentaram as menores médias, 1,27 m s^{-1} e 0,86 m s^{-1}, respectivamente (Figura 5E). A velocidade do vento pode causar danos as culturas agrícolas, na absorção de CO_2, estímulo excessivo da evapotranspiração, danos mecânicos de quebra de galhos e tombamento (Pereira et al., 2002). Resultados observados por Borges e Souza (2004), identificaram perdas de produtividade de 20 a 30% por danos causadas por ventos em bananicultura.

As maiores médias de umidade relativa do ar são apresentadas pelos locais do grupo 2 (79,79%) e grupo 1 (75,92%). O grupo 3 e 4 apresentaram as

menores médias, 72,91% e 78,33%, respectivamente (Figura 5F). A umidade relativa do ar é um elemento meteorológico que também influência os rendimentos das culturas agrícolas, pois regula o ciclo de pragas e doenças em culturas, assim como no desenvolvimento e crescimento das plantas (Ometto, 1981). Resultados observados por Roriz et al., (2014) concluíram que plantas de soja podem apresentar crescimento reduzido, devido ao volume de transpiração reduzido, levando a menor translocação de nutrientes.

O grupo 3 apresentou as maiores médias de déficit hídrico 0,365 mm, consequentemente os locais desse grupo podem apresentar um potencial período de seca. Por outro lado, os locais que apresentaram as menores médias foram: o grupo 1 apresentou 0,049 mm, grupo 2 (0,018 mm) e grupo 4 (0,027 mm) (Figura 5G). Resultados semelhantes foram observados por Aparecido et. al (2023), concluíram que as maiores DEF estão concentradas na região sul do estado do maranhão, sendo a região com maior produtividade de soja do Maranhão. Entretanto, Sentelhas et al., (2015), ao avaliar os efeitos integrados da deficiência hídrica e do manejo da cultura sobre as brechas produtivas da soja, concluíram que atenção especial deve ser dada ao sul do Brasil, por apresentar lacunas de rendimento total superior a 2000 kg ha^{-1}. O déficit hídrico é o principal fator para as lacunas de produtividade de soja (Sentelhas et al., 2015), pode influenciar diretamente o desenvolvimento dos estádios fenológicos da soja (Zanon et al., 2016, Tagliapietra et al., 2021).

O grupo 2 apresentou a maior média de excedente hídrico 62,84 mm, seguido do grupo 1 que apresentou 47,94 mm, estando esses resultados diretamente relacionado com os resultados de precipitação. E os grupos com as menores médias foram os grupos 4 e 3, que apresentaram 40 mm e 37,44 mm,

respectivamente (Figura 5H). Resultados semelhantes foram observados por Sentelhas (2008).

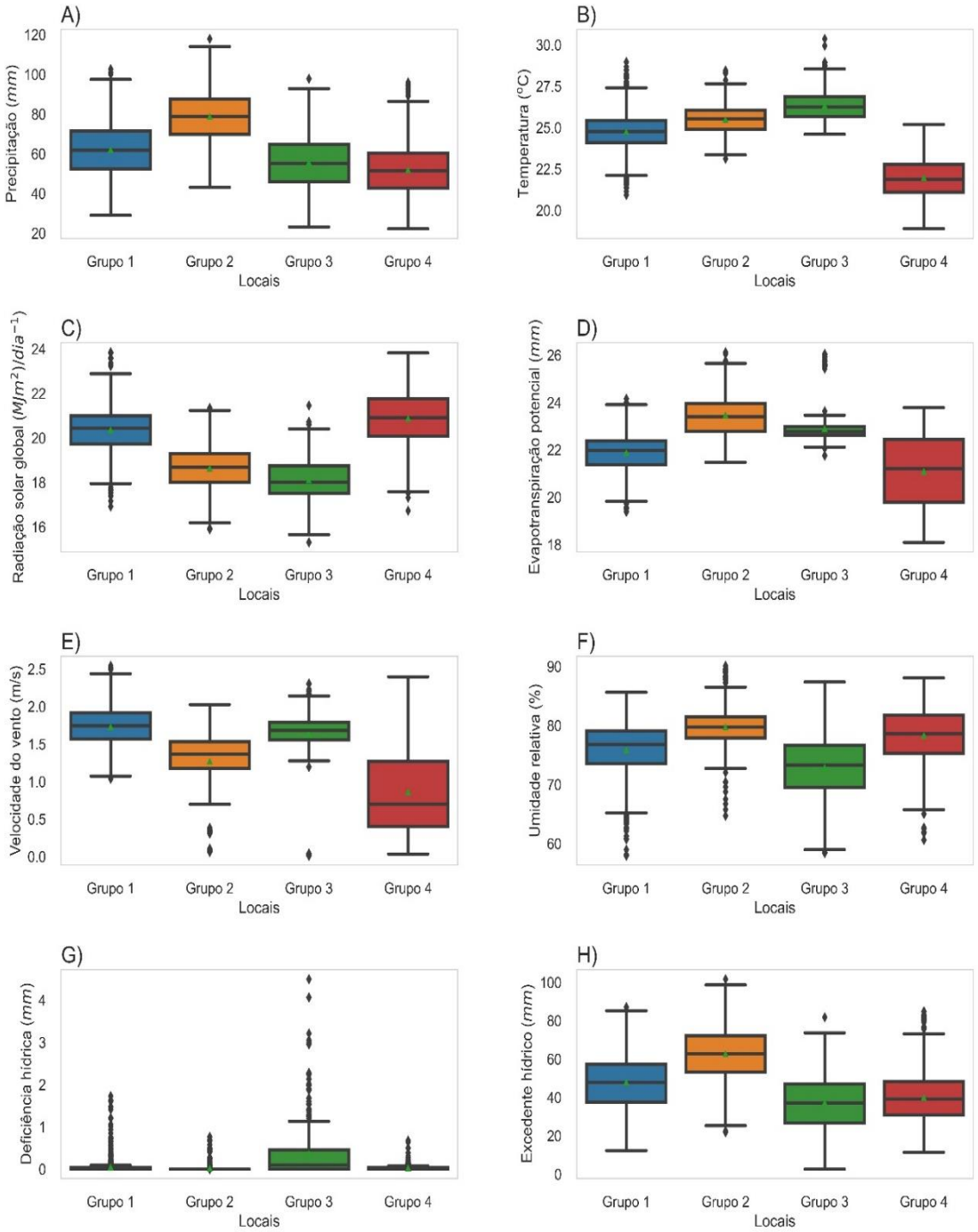

Figura 5. Box-plot dos elementos meteorológicos precipitação, temperatura média do ar (T_{Med}), radiação solar (Qg), evapotranspiração potencial (ETp), velocidade do vento (U2), umidade relativa (UR), déficit hídrico (EXC) e excedente hídrico (DEF).

Os algoritmos demonstraram diferentes acurácia e precisão na predição de produtividade de soja utilizando dados climáticos (Tabela 3 a 6).

Tabela 3. Desempenho da Regressão linear múltipla durante as etapas de treinamento e teste.

	Treinamento			Teste		
Locais	R^2 adj	RMSE	MAE	R^2 adj	RMSE	MAE
Grupo 1	0.53	407.5	313.45	0.01	3062.3	2513.6
Grupo 2	0.66	315.53	237.09	0.01	3496.2	3096.7
Grupo 3	0.84	279.07	216.22	0.01	3993.8	3358.4
Grupo 4	0.72	431.87	337.85	0.01	3786	3315.6

R^2 ajustado (R^2 adj), raiz do erro quadrático médio RMSE, Erro médio absoluto (MAE).

Tabela 4. Desempenho do Random Forest durante as etapas de treinamento e teste.

	Treinamento			Teste		
Locais	R^2 adj	RMSE	MAE	R^2 adj	RMSE	MAE
Grupo 1	0.95	129.09	94.79	0.64	355.62	263.31
Grupo 2	0.93	138.32	96.96	0.53	352.87	269.39
Grupo 3	0.97	128.4	91.77	0.66	399.93	304.99
Grupo 4	0.96	158.45	115.31	0.76	389.41	297.03

R^2 ajustado (R^2 adj), raiz do erro quadrático médio RMSE, Erro médio absoluto (MAE).

Tabela 5. Desempenho do Suport Vector Machine durante as etapas de treinamento e teste.

	Treinamento			Teste		
Locais	R^2 adj	RMSE	MAE	R^2 adj	RMSE	MAE
Grupo 1	0.75	297.81	194.62	0.59	380.11	276.66
Grupo 2	0.54	364.29	220.37	0.31	426.96	317.11
Grupo 3	0.73	360.69	215.48	0.49	492.63	358.13

Grupo 4	0.80	365.85	249.08	0.74	406.93	309.15

R^2 ajustado (R^2 adj), raiz do erro quadrático médio RMSE, Erro médio absoluto (MAE).

Tabela 6. Desempenho do Multilayer Perceptron nas etapas de treino e teste.

	Treinamento			Teste		
Locais	R^2 adj	RMSE	MAE	R^2 adj	RMSE	MAE
Grupo 1	0.61	371.46	287.59	0.45	439.92	333.82
Grupo 2	0.87	190.82	121.18	0.45	619.98	438.13
Grupo 3	0.96	135.6	78.01	0.18	626.61	426.99
Grupo 4	0.86	307.57	232.66	0.67	457.67	357.52

R^2 ajustado (R^2 adj), raiz do erro quadrático médio RMSE, Erro médio absoluto (MAE).

De modo geral na etapa de treinamento o modelo com o melhor desempenho foi apresentado pelo algoritmo RF (Tabela 4), e o menor desempenho foi apresentado pelo algoritmo RLM (Tabela 3). O RF demonstrou maior precisão ($R^2\ adj$) e acurácia ($RMSE$). os melhores ajustes foram observados para os locais do grupo 3, com $R^2\ adj$ de 0,97 e $RMSE$ de 128,4 Kg ha^{-1}. O RF explora diversas alternativas dentro do conjunto de dados variando os subconjuntos e entradas de dados para cada árvore e toma suas decisões com base na média de todas as árvores que o compõe, o que consequentemente define assim o RF um algoritmo robusto e eficaz contra o overfitting. De acordo com Gill (2022), o RF pode superar os modelos de aprendizado profundo em desempenho de previsões.

O menor desempenho para etapa de treinamento foi apresentado pelo algoritmo RLM para o grupo 1 com $R^2\ adj$ de 0,53 e $RMSE$ de 407,5 kg ha^{-1}. Assim como na etapa de teste, onde o algoritmo RLM apresentou o menor

desempenho. Nessa etapa de teste o menor desempenho foi demonstrado no grupo 4, com $R^2\ adj$ de 0,01 e $RMSE$ de 3993,8 kg ha^{-1} (Tabela 3). Resultados semelhantes foram observados por Aparecido et al., (2020), avaliando algoritmos para a previsão de algodão com base em parâmetros climáticos no Brasil, concluíram que o RLM teve o menor desempenho, $R^2\ adj$ de 0,17.

O algoritmo RF apresentou o melhor desempenho na etapa de teste dos modelos. O melhor ajuste de RF foi observado para o grupo 4, com $R^2\ adj$ de 0,76 e $RMSE$ de 389,41 kg ha^{-1} (Tabela 4).

De modo geral o SVM apresentou o segundo melhor desempenho entre os modelos estudados. Na etapa de treinamento o melhor ajuste foi apresentado pelo grupo 4 com $R^2\ adj$ de 0,80 e $RMSE$ 365,85 kg ha^{-1}. O mesmo foi observado na etapa de teste com $R^2\ adj$ de 0,74 e $RMSE$ 406,93 kg ha^{-1} (Tabela 5).

O MLP apesar de ter apresentado bom desempenho durante a etapa de treinamento, na etapa de teste o algoritmo não demostrou bons ajustes. Neste modelo o grupo 4 também apresentou os melhores ajustes. Na etapa de treino o $R^2\ adj$ foi de 0,86 e $RMSE$ de 307,57 kg ha^{-1}. Já na etapa de teste apresentou $R^2\ adj$ de 0,67 e $RMSE$ de 457,67 kg ha^{-1} (Tabela 6).

Nossos resultados corroboram com os resultados encontrados por Santos et al., (2022), em estudos visando avaliar o desempenho de algoritmos de aprendizado de máquina para a previsão de produtividade de soja no cerrado brasileiro, observaram que o RF demonstrou alta precisão e acurácia na predição da produtividade da soja no Cerrado brasileiro quando comparado ao SVM e

MLP. O algoritmo RF obteve maior desempenho, com $R^2\ adj$ de 0,81 e $RMSE$ de 176,93 kg ha^{-1}.

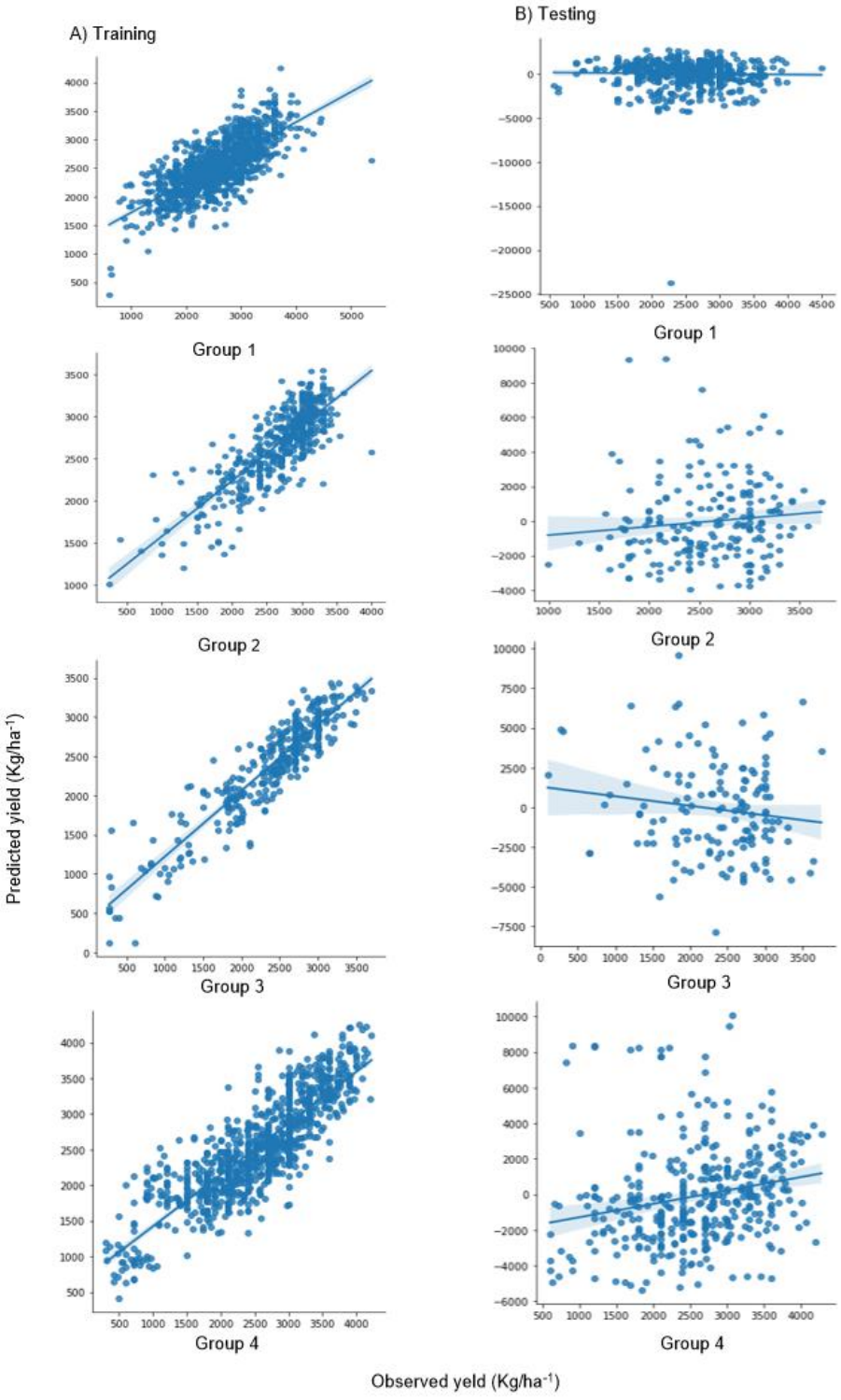

Figura 5. Treinamento e teste do modelo RLM dos grupos de localidades produtoras de soja estudadas.

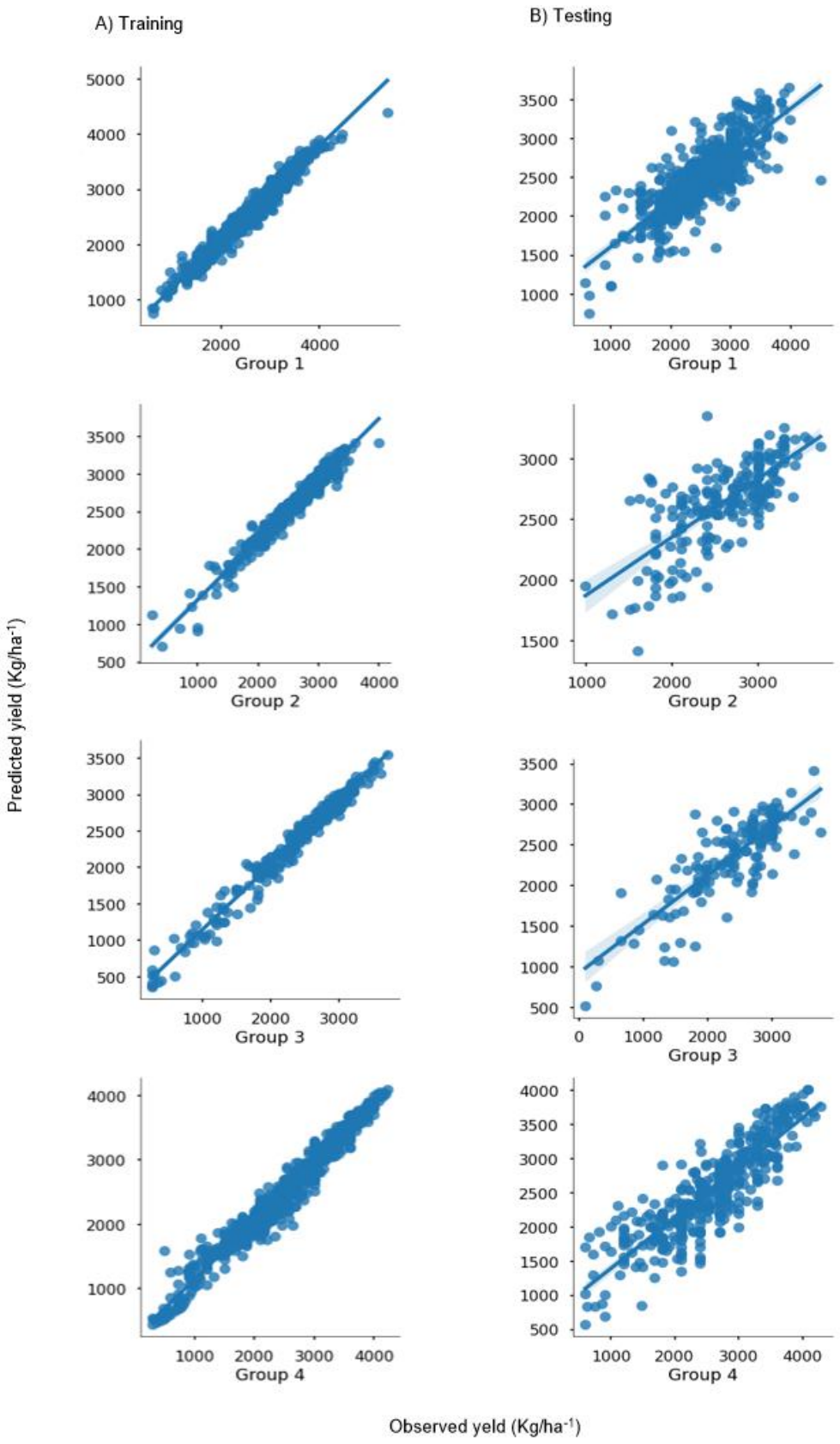

Figura 6. Treinamento e teste do modelo RF dos grupos de localidades produtoras de soja estudadas.

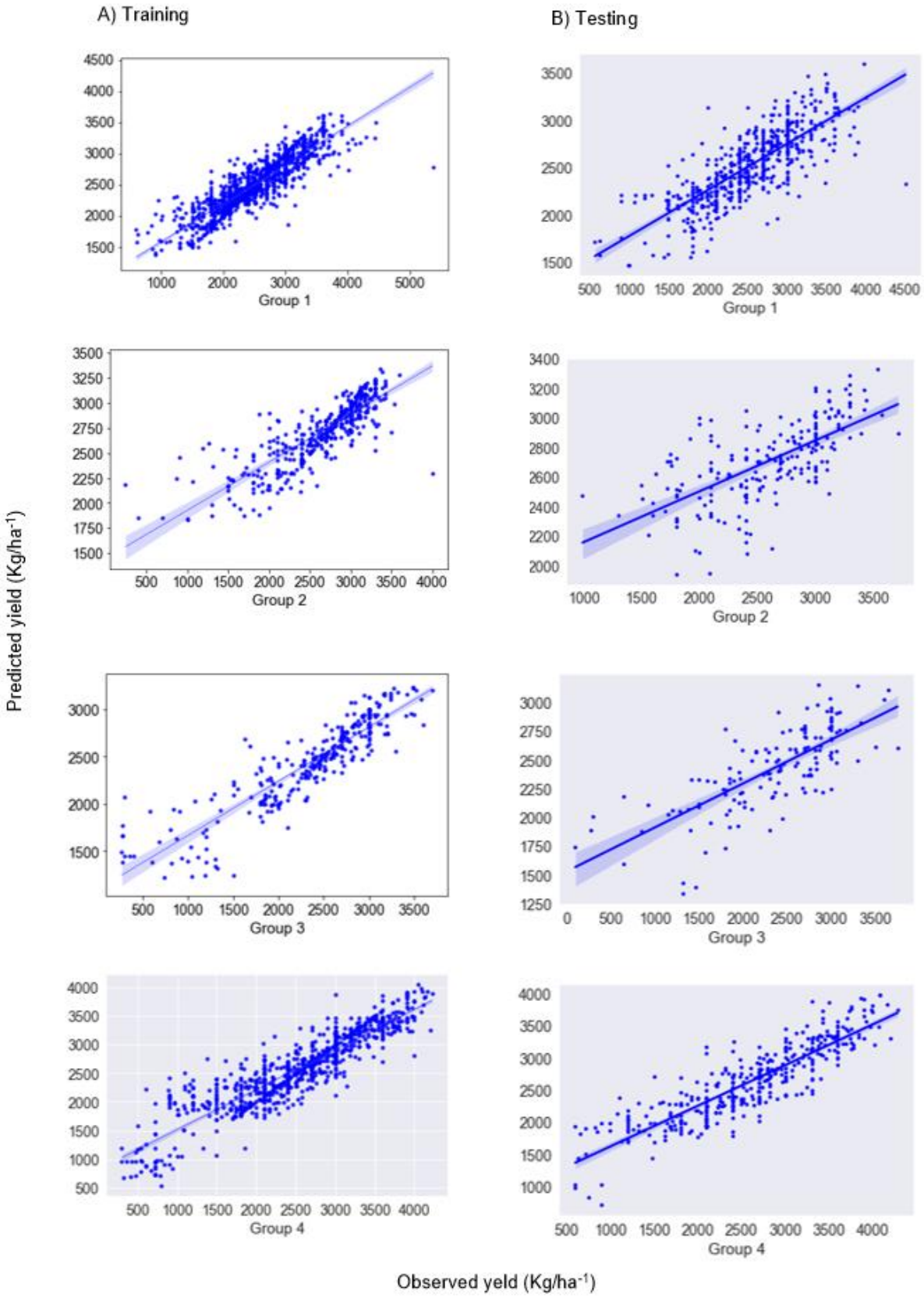

Figura 7. Treinamento e teste do modelo SVM dos grupos de localidades produtoras de soja estudadas.

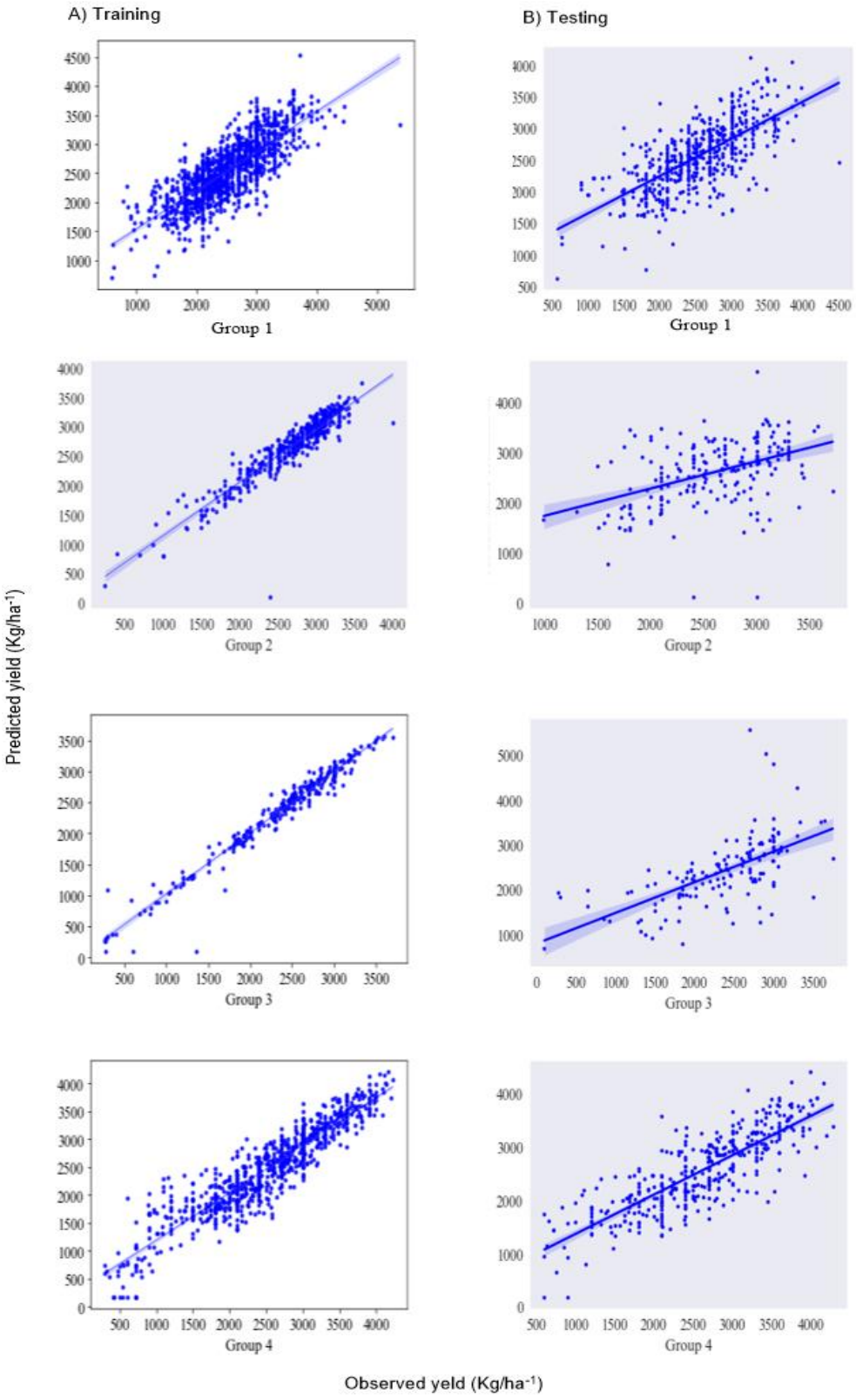

Figura 8. Treinamento e teste do modelo MLP dos grupos de localidades produtoras de soja estudadas.

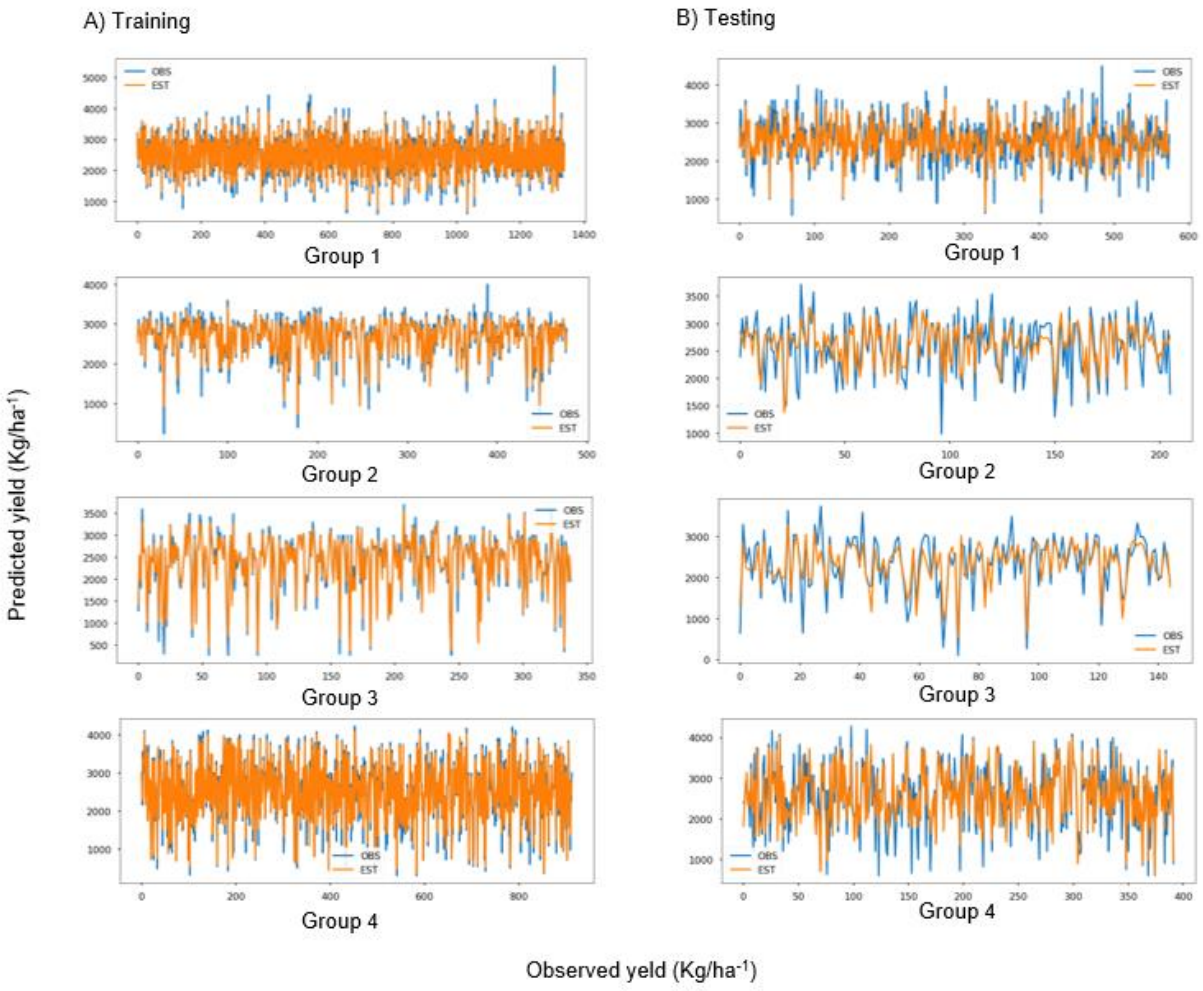

Figura 9. Treinamento e teste do modelo RF dos grupos de localidades produtoras de soja estudadas.

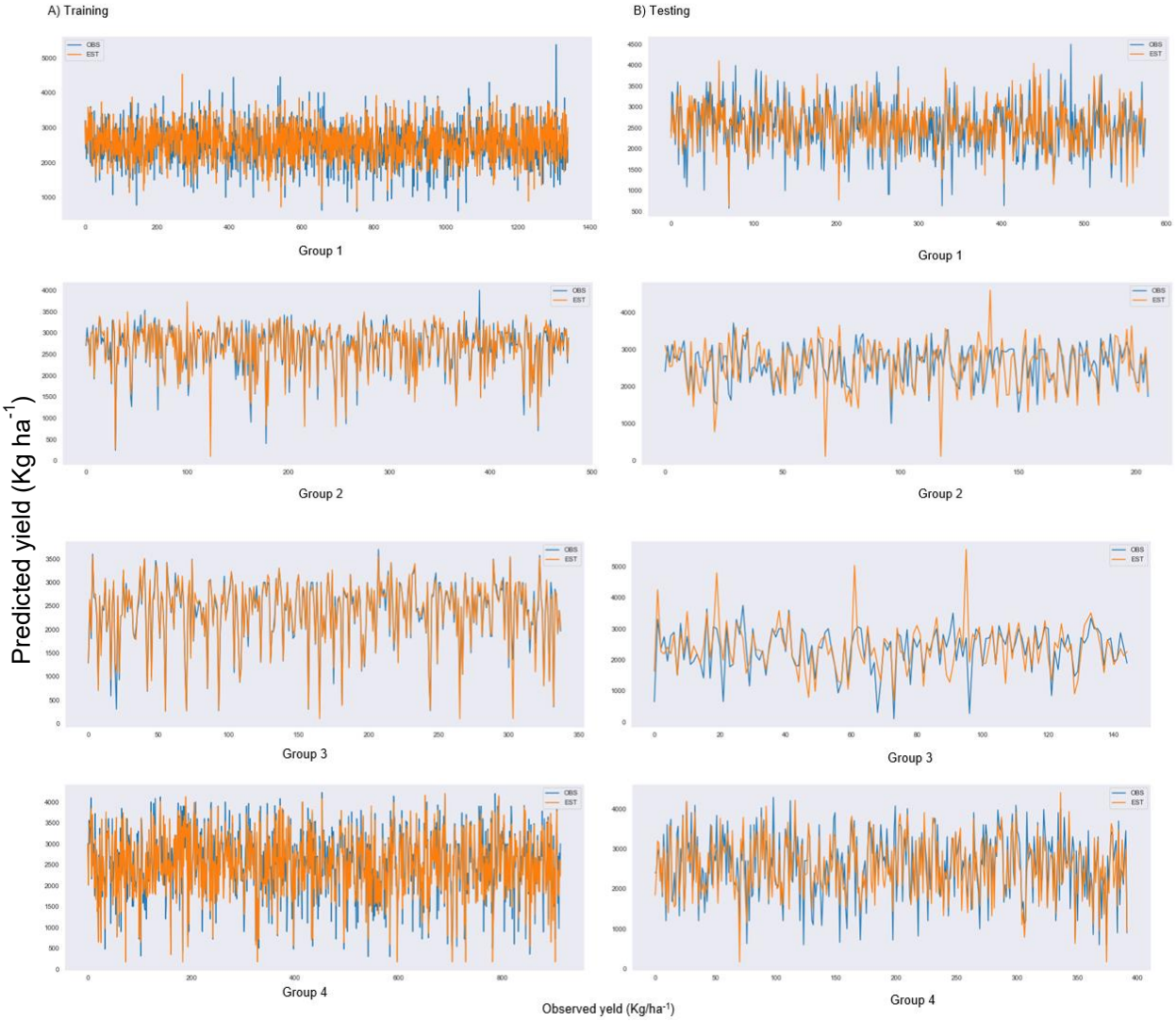

Figura 10. Treinamento e teste do modelo SVM dos grupos de localidades produtoras de soja estudadas.

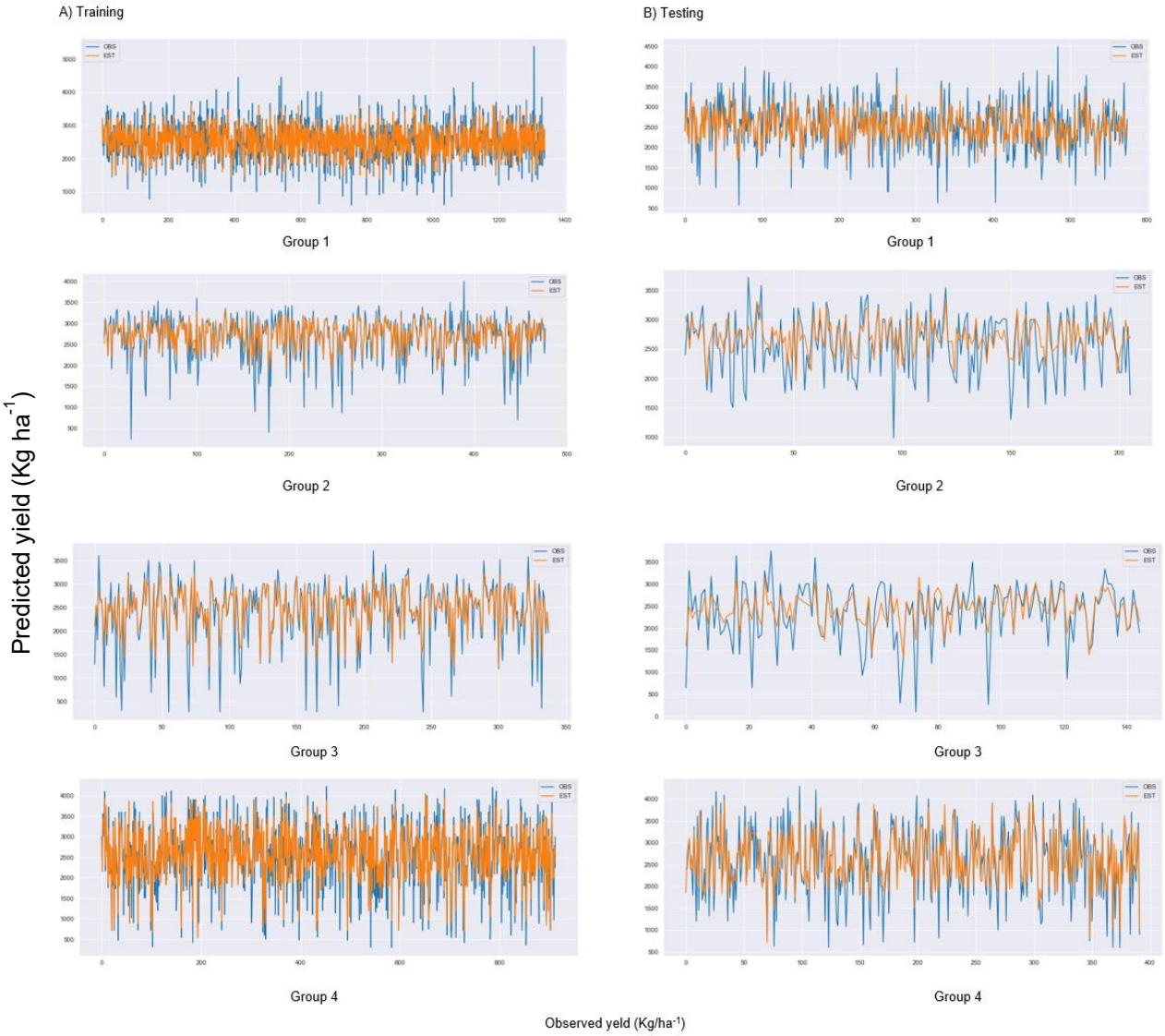

Figura 11. Treinamento e teste do modelo MLP dos grupos de localidades produtoras de soja estudadas.

3.4 Conclusão

É possível utilizar técnicas de aprendizado de máquina para estimação e predição de produtividade de soja no Brasil. O algoritmo Random Forest apresentou o melhor desempenho para estimação e predição de produtividade, com $R^2\ adj$ de 0,97 e $RMSE$ de 128,4 Kg ha^{-1}. O algoritmo mostrou maior

desempenho ao predizer a produtividade dos locais agrupados no grupo 4, esses locais fazem parte da região sul e sudeste.

Referências

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration-Guidelines for computing crop water requirements-FAO. **Irrigation and drainage.** 56. *Fao*, *300* (9), D05109

Aparecido LEO, de Meneses KC, Lorençone PA, Lorençone JA, de Moraes JRSC, de Souza Rolim G (2023) Climate classification by Thornthwaite (1948) humidity index in future scenarios for Maranhão State, Brazil. **Enviroment, Development and Sustaintability**, 25:855-878**.** DOI: https://doi.org/10.1007/s10668-021-02082-9

Aparecido LEO, Lorençone JA, Lorençone PA, Torsonim GB, Lima RF, de Moraes JRSC (2022) Predicting coffee yield based on agroclimatic data and machine learning. **Theoretical and Applied Climatology**, 148:899-914. DOI: https://doi.org/10.1007/s00704-022-03983-z

Aparecido LEO, Meneses KC, de Souza Rolim G, Carvalho MJN, Pereira WBS, da Silva PA, Santos TS, de Moraes JRSC (2020) Algorithms for forecasting cotton yield based on climatic parameters in Brazil. **Archives of Agronomy and Soil Science** 68:984-1001. DOI: https://doi.org/10.1080/03650340.2020.1864821

Auler FR, Minuzzi RB (2022) Estimativa de radiação solar global diária com base na temperatura do ar para diferentes condições de nebulosidade em Florianópolis, Santa Catarina. **Journal of Environmental Analysis and Progress**, 7: 026–032. DOI: https://doi.org/10.24221/jeap.7.01.2022.3996.026-032

Basak D, Pal S, Patranabis DC (2007) Support vector regression. **Neural Information Processing-Letters and Reviews**, 11(10), 203–224.

Beisekenov NA, Anuarbekov TB, Sadenova MA, Varbanov PS, Klemeš JJ, Wang J (2021) Machine Learning Model Identification for Forecasting of Soya Crop Yields in Kazakhstan. In: 2021 6th INTERNATIONAL CONFERENCE ON SMART AND SUSTAINABLE TECHNOLOGIES. **Anais …** Bol and Split, Croácia: SpliTech, p. 1-6. DOI: https://doi.org/10.23919/SpliTech52315.2021.9566376

Borges AL, Souza LS (2004) Exigências Edafoclimáticas. In: Borges AL, Souza LS (Eds.). **O cultivo da bananeira.** Cruz das Almas: Embrapa, p. 15-22.

Branco JA, Pires AM (2017) **Introdução aos Métodos Estatísticos Robustos**. Lisboa: Edições SPE.

Chen M, Lijuan X, (2014) A preliminary study of aflatoxin b1 detection in peanut oil by terahertz time-domain spectroscopy. **Transactions of the ASABE**, 57:1793-1799. DOI:doi: 10.13031/trans.57.10725

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2020) **Acompanhamento da Safra Brasileira de Grãos 2019/2020**: relatório técnico. Brasília, v. 8.

CONAB (Governo Federal). Companhia Nacional de Abastecimento (2022) **Acompanhamento da Safra Brasileira de Grãos 2021/2022**: relatório técnico. Brasília, v. 9.

Cortes C e Vapnik V (1995) Support vector networks. **Machine Learning**, 20(3), 273–297.

Curto L, Covi M, Gassmann MI (2019) Actual evapotranspiration and the pattern of soil water extraction of a soybean (*Glycine max*) crop. **Revista de la Facultad de Ciencias Agrarias. Universidade Nacional de Cuyo** 51:125-141.

Cutler A, Cutler DR, Stevens JR (2012). Random Forests. In: Zhang C, Ma Y (Eds.) **Ensemble Machine Learning: Methods and Applications**, Boston: Springer, p.157–175. DOI:https://doi.org/10.1007/978-1-4419-9326-7

Doorenbos J, Kassam AH (1979) **Yield response to water**. Rome: Food and Agriculture Organization of the United Nations, p.257.

Duarte YCN, Sentelhas PC (2020) Intercomparison and Performance of Maize Crop Models and Their Ensemble for Yield Simulations in Brazil. **International Journal of Plant Production** 14:127–139. DOI: https://doi.org/10.1007/s42106-019-00073-5

Farias JRB, Neumaier N, Nepomuceno AL (2007) **Ecofisiologia da soja.** Londrina: Embrapa Soja, 9 p. (Embrapa Soja, Circular Técnica, 48).

Gill M, Anderson R, et al. (2022) Machine learning models outperform deep learning models, provide interpretation and facilitate feature selection for soybean trait prediction. **BMC Plant Biology** 22:180. DOI: https://doi.org/10.1186/s12870-022-03559-z

Haykin S (1999) Neural networks: a comprehensive foundation. **Printice-hall**. Inc., New Jersey, p. 120-134.

Hoffmann R (2016) **Análise de regressão: uma introdução à econometria [recurso eletrônico]**. Piracicaba: O Autor p.393.

IBGE (Instituto Brasileiro de Geografia e Estatística), (2020) Sistema IBGE de Recuperação Automática - SIDRA: Produção Agrícola Municipal. Disponível em: https://sidra.ibge.gov.br/pesquisa/ppm/quadros/brasil/2020. (Acesso em 10/09/2020).

Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI). **Anais …** São Francisco, CA, USA: IJCA'95, p. 1137-1143.

Monteiro LA, Sentelhas PC, Pedra GU (2018) Assessment of NASA/POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. **Internation Journal of Climatology** 38:1571-1581. DOI: https://doi.org/10.1002/joc.5282

Moraes JRSC, de Souza Rolim G, Martorano LG, de Oliveira Aparecido LE, Padilha de Oliveira MS, de Farias Neto JT (2020) Agrometeorological models to forecast açaí (*Euterpe oleracea* Mart.) yield in the Eastern Amazon. **Journal of the Science of Food and Agriculture**, 100: 1558-1569. DOI: https://doi.org/10.1002/jsfa.10164

Morelli-Ferreira F, Maia NJC, Tedesco D, Kazama EH, Morlin Carneiro F, Santos LB, Seben Junior GF, Rolim GS, Shiratsuchi LS, Silva RP (2021) Comparison of Machine Learning Techniques in Cotton Yield Prediction Using Satellite Remote Sensing. **Preprints** 2021, DOI: https://doi.org/10.20944/preprints202112.0138.v2.

Oliveira GQ, Biscaro GA, Lopes AS, Jung LH, Schwerz F (2014) Comparison between global solar radiation models in aquidauana, "alto pantanal" region. **Comunicata Scientiae**, 5:222-228.

Ometto JC (1981) **Bioclimatologia Vegetal.** São Paulo: Ed. Agronômica Ceres p. 440.

Pereira AR, Angelocci LR, Sentelhas PC (2002) **Agrometeorologia: Fundamentos e Aplicações Práticas.** Porto Alegre: Guaíba Agropecuária p. 478.

Portugal I, Alencar P, Cowan D (2018) The use of machine learning algorithms in recommender systems: A systematic review. **Expert Systems with Applications** 97:205–227. DOI: https://doi.org/10.1016/j.eswa.2017.12.020

Reis IS, Silva TR, Klering E, Lindemann DS, Maier, ELB. Análise da Variabilidade da Temperatura do Ar em Rio Grande – RS (1913-2016). **Revista Brasileira de Geografia Física** 15:583-601.

Rolim GS, Aparecido LEO, Souza PS, Lamparelli RAC, Santos ER (2020) Climate and natural quality of Coffea arábica L. drink. **Theoretical and Applied Climatology** 141:87-98. DOI: https://doi.org/10.1007/s00704-020-03117-3

Roriz M, Carvalho SMP, Vasconcelos MW (2014) High relative air humidity influences mineral accumulation and growth in iron deficient soybean plants. **Frontiers in Plant Science** 5:726. DOI: https://doi.org/10.3389/fpls.2014.00726

Santos BV, Moreno Ferreira dos Santos A, da Silva Cabral de Moraes JR, de Oliveira Vieira IC, de Souza Rolim G (2022) Machine learning algorithms for soybean yield forecasting in the Brazilian Cerrado. **Journal of the Science of Food and Agriculture,** 102: 3665-3672. DOI: https://doi.org/10.1002/jsfa.11713

Segal MR (2004) **Machine Learning Benchmarks and Random Forest Regression**. São Francisco: UCSF Center for Bioinformatics and Molecular Biostatistics p. 15.

Sentelhas PC, Battisti R, Câmara GMS, Farias JRB, Hampf AC, Nendel, C (2015) The soybean yield gap in Brazil – magnitude, causes and possible solutions for sustainable production. **The Journal of Agricultural Science** 153:1394-1411. DOI: https://doi.org/10.1017/S0021859615000313

Sentelhas PC, Santos DL, Machado RE (2008) Water deficit and water surplus maps for Brazil, based on FAO Penman-monteith potential evapotranspiration. **Revista Ambiente e Água** 3:28-42. DOI: doi:10.4136/ambi-agua.59

Silva WL, Júnior ACO, Cavalcanti IFA, Treistman F. (2021) An overview of precipitation climatology in Brazil: space-time variability of frequency and intensity associated with atmospheric systems. **Hydrological Sciences Journal** 66:289-308. DOI: https://doi.org/10.1080/02626667.2020.1863969

Souza A, dos Santos CM, Ihaddadene R, Cavazzana G, Abreu MC, Oliveira-Júnior JF, Pobocikova I, de Gois G, Lins TMP (2022) Analysis of extreme monthly and annual air temperatures variability using regression model in Mato Grosso do Sul, Brazil. **Modeling Earth Systems and Environment**, 8:647-663. DOI: https://doi.org/10.1007/s40808-021-01096-6

Tagliapietra, EL, Zanon, AJ et al. (2021) Biophysical and management factors causing yield gap in soybean in the subtropics of Brazil. **Agronomy Journal** 113:1882-1894. DOI: https://doi.org/10.1002/agj2.20586

Tanveer M, Rajani T, Rastogi R, Shao YH, Ganaie MA (2022) Comprehensive review on twin support vector machines. **Annals Operations Research**. DOI: https://doi.org/10.1007/s10479-022-04575-w

Thornthwaite CW, Mather JR (1955) The water balance. **Drexel Institute of Technology**, Laboratory of Climatology, Centerton, NJ, p 104

Tsukahara RY, Fonseca ICB, Silva MA A, Kochinski EG, Neto JP, Suyama JT (2016) Soybean yield as a consequence of harvest delay and environmental conditions. **Pesquisa Agropecuária Brasileira – Agrometeorologia** 51:905-915.

Werkema MCC, Aguiar S (1996) **Análise de regressão: como entender o relacionamento entre as variáveis de um processo.** Belo Horizonte: Werkema Editora.

Zanon AJ, Streck NA, Grassini P (2016) Climate and Management Factors Influence Soybean Yield Potential in a Subtropical Environment. **Agronomy Journal** 108:1447-1454. DOI: https://doi.org/10.2134/agronj2015.0535.

Zhang L, Zhou W, Jiao L (2004) Wavelet support vector machine. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 34, n. 1, p. 34-39.

Stackhouse PW, Westberg D, Hoell JM, Chandler WS, Zhang T. (2017) Prediction Of Worldwide Energy Resource (POWER)--**Sustainable Buildings Methodology**-(1.0 o Latitude by 1.0 o Longitude Spatial Resolution).

CAPÍTULO 4. CONSIDERAÇÕES FINAIS

1. Atualmente as mudanças climáticas é uma das questões mais preocupantes para a agricultura, visto que o clima é um dos fatores que mais influenciam o desenvolvimento e estabelecimento das plantas. A soja é um dos grãos mais cultivados no mundo, entretanto, é uma planta extremamente sensível a variabilidade climática. Portanto, a identificação da influência dos elementos climáticos na fenologia de plantas de soja e aplicação de técnicas de estimação de produtividade, são ferramentas importantes para tomadas de decisão de manejo da cultura, visando minimizar os impactos sofridos na produtividade pelos elementos climáticos.

2. A identificação da relação dos elementos climáticos com a produtividade da soja em cada estádio fenológico da planta pode ajudar principalmente nas estratégias de manejo da cultura, conforme observado no capitulo 2 deste trabalho. Buscando minimizar as perdas de produtividade diante de um cenário onde cada vez mais há instabilidade climática.

3. A utilização de modelos de inteligência artificial mostrou ser uma técnica capaz de estimar a produtividade da soja com base em dados climáticos (capitulo 3). O modelo Random Forest (RF) apresentou o melhor desempenho entre os modelos estudados, podendo ser utilizado para estimar e predizer a produtividade da soja no Brasil de forma acurada e precisa.

4. Os resultados alcançados neste trabalho poderão ser utilizados como base para novas pesquisas, visando o estudo aprofundado da aplicação de técnicas de inteligência artificial para melhor entender a complexidade de dados climáticos e suas influências no ciclo fenológico de plantas de

soja. Além do estudo da utilização de técnicas economicamente viáveis e ambientalmente mais sustentáveis aplicadas a agricultura e prevenção de perdas no campo.