

ÍCARO ALVES RAMIRO

**RELAÇÃO ENTRE O FLORESCIMENTO DA CANA-DE- AÇÚCAR E A
DISPONIBILIDADE HÍDRICA PARA A REGIÃO DE TABULEIROS COSTEIROS -
ALAGOAS, BRASIL**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Meteorologia Aplicada, para obtenção do título de *Magister Scientiae*.

Orientadora: Hewlley Maria Acioli Imbuzeiro

Coorientadores: Genaina Ap. de Souza
Guilherme Bastos Lyra

**VIÇOSA - MINAS GERAIS
2021**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

R173r
2021 Ramiro, Ícaro Alves, 1993-
Relação entre o florescimento da cana-de-açúcar e a
disponibilidade hídrica para a região de Tabuleiros Costeiros -
Alagoas, Brasil / Ícaro Alves Ramiro. – Viçosa, MG, 2021.
1 dissertação eletrônica (97 f.): il. (algumas color.).

Inclui anexos.

Orientador: Hewlley Maria Acioli Imbuzeiro.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Engenharia Agrícola, 2021.

Referências bibliográficas: f. 81-93.

DOI: <https://doi.org/10.47328/ufvbbt.2021.012>

Modo de acesso: World Wide Web.

1. Cana de açúcar - Cultivo - Modelos estatísticos. 2.
Saccharum spp. 3. Análise de regressão. I. Imbuzeiro, Hewlley
Maria Acioli. II. Universidade Federal de Viçosa. Departamento
de Engenharia Agrícola. Programa de Pós-Graduação em
Meteorologia Aplicada. III. Título.

CDD 22. ed. 633.61

Bibliotecário(a) responsável: Renata de Fátima Alves x

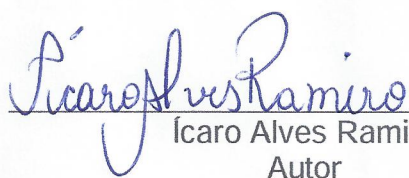
ÍCARO ALVES RAMIRO

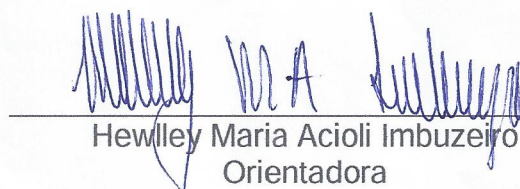
**RELAÇÃO ENTRE O FLORESCIMENTO DA CANA-DE- AÇÚCAR E A
DISPONIBILIDADE HÍDRICA PARA A REGIÃO DE TABULEIROS COSTEIROS -
ALAGOAS, BRASIL**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Meteorologia Aplicada, para obtenção do título de *Magister Scientiae*.

APROVADA: 26 de julho de 2021.

Assentimento:


Icaro Alves Ramiro
Autor


Hewlley Maria Acioli Imbuzeiro
Orientadora

AGRADECIMENTOS

Dedico à minha família, em especial à tia Ana e minha mãe, que sempre acreditaram em mim e não pouparam esforços para realização desta conquista. Ao André, Gisele e Pedro; companheiros nas horas difíceis e alegrias. À minha coorientadora, Genaina, e Professora Fernanda pelos valiosos conselhos. Aos meus colegas e professores do Programa de Meteorologia Aplicada (DEA/UFV), em especial à Graça e ao grupo Micromet.

À Deus.

À Universidade Federal de Viçosa, pela oportunidade de realizar a pós-graduação.

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

“Nunca entendi a luta senão como um meio de acabar com ela. Nunca aceitei o rigor senão como meio de deixar de existir o rigor”.

(Pablo Neruda)

RESUMO

RAMIRO, Ícaro Alves, M.Sc., Universidade Federal de Viçosa, julho de 2021. **Relação entre o florescimento da cana-de-açúcar e a disponibilidade hídrica para a região de Tabuleiros Costeiros - Alagoas, Brasil.** Orientadora: Hewlley Maria Acioli Imbuzeiro. Coorientadores: Genaina Aparecida de Souza e Guilherme Bastos Lyra.

O florescimento da cana-de-açúcar é um fenômeno natural prejudicial, quando o objetivo da produção é a fabricação de biocombustíveis. Isso porque, nesse processo, a cana-de-açúcar desvia as reservas energéticas para a produção das panículas, reduzindo o acúmulo de sacarose, que é convertida em etanol pela indústria. Esse fenômeno é dependente da disponibilidade hídrica no período de indução floral, seja para a absorção de nutrientes ou pela relação água-fixação de carbono. Assim, o objetivo deste trabalho foi estimar a lâmina de água acumulada no solo, chuvas e evapotranspiração, ou seja, a disponibilidade hídrica que induz o florescimento, na região de Tabuleiros Costeiros em Alagoas (AL). Essas lâminas foram obtidas a partir do balanço hídrico sequencial diário e observações meteorológicas em Barra de São Miguel (AL). Depois de obtidas, essas lâminas foram analisadas pelo modelo de regressão logística, juntamente com o número de panículas observadas no mesmo município. Dos modelos gerados, apenas aqueles que tiveram chuva e evapotranspiração como regressores foram significativos ($p < 0.05$). Os regressores desses modelos foram: chuva soma (C_s) e chuva média da soma (C_{sm}), do volume precipitado em 10 dias consecutivos; e evapotranspiração diária (EToHS) e média (EToHS_m), calculadas pelo método de Hargreaves-Samani. Esses modelos também foram os que apresentaram os menores valores de critério de informação de Akaike (AIC): 103.26 para C_s , 102.31 para C_{sm} , 103.85 para EToHS e 103.48 para EToHS_m. Os resultados das razões de chance (OR) mostraram que para cada variação de 1 mm.dia⁻¹ (EToHS e EToHS_m), ou de 1mm.10 dias⁻¹ (C_s e C_{sm}), as chances de ocorrência de, pelo menos, 1% de florescimento diminuem em 0.9981 (C_s) e 0.9979 (C_{sm}) vezes; e aumentam 2.3279 (EToHS) e 2.9066 (EToHS_m) vezes. A partir da matriz de confusão, as acurácias calculadas desses modelos foram 64.94% (C_s), 66.23% (C_{sm}), 71.43% (EToHS) e 72.73% (EToHS_m). Analisando a *Receiver Operating Characteristic Curve* (curva ROC), os valores de área sob a curva (AUC) foram: 0.618 (C_s), 0.640 (C_{sm}), 0.649 (EToHS) e 0.662 (EToHS_m). Os valores de

especificidade dos modelos foram: 0.312 (C_s), 0.656 (C_{sm}), 0.531 (EToHS) e 0.438 (EToHS_m). Já os valores de sensibilidade foram: 1.0 (C_s), 0.644 (C_{sm}), 0.844 (EToHS) e 0.933 (EToHS_m). Os modelos com C_s e C_{sm} apresentaram maiores valores de “pseudo” R^2 com valores de 8,92% (C_s) e 10,47% (C_{sm}) contra 7,95% (EToHS) e 8,56% (EToHS_m). Analisando o desempenho geral, entre os modelos que tiveram chuva como regressor, o modelo com C_{sm} tem desempenho geral levemente maior que o com C_s . Enquanto que, entre os modelos que tiveram evapotranspiração como regressores, o modelo com EToHS_m tem desempenho geral levemente maior que o com EToHS. Os resultados sugerem que a influência da chuva sobre o florescimento da cana se dá de maneira acumulativa, enquanto a evapotranspiração afeta de maneira média.

Palavras-chave: *Saccharum spp.* Validação. Intercomparação de modelos. Regressão Logística. *Proxies*.

ABSTRACT

RAMIRO, Ícaro Alves, M.Sc., Universidade Federal de Viçosa, July, 2021. **Relationship between sugarcane flowering and water availability in the Tabuleiros Costeiros region - Alagoas, Brazil.** Advisor: Hewlley Maria Acioli Imbuzeiro. Co-advisors: Genaina Aparecida de Souza and Guilherme Bastos Lyra.

The flowering of sugarcane is a harmful natural phenomenon when the objective of production is the manufacture of biofuels. This is because, in this process, sugarcane diverts energy reserves for the production of panicles, reducing the accumulation of sucrose, which is converted into ethanol by the industry. This phenomenon depends on water availability during the floral induction period, whether for nutrient absorption or for the water-carbon fixation ratio. Thus, the aim of this work was to estimate the amount of water accumulated in the soil, rainfall and evapotranspiration, that is, the water availability that induces flowering, in the region of Tabuleiros Costeiros in Alagoas (AL). We obtained these slides from daily sequential water balance and meteorological observations in Barra of São Miguel (AL). Once obtained, we analyzed these slides by the logistic regression model, together with the number of panicles observed in the same municipality. Of the generated models, only those that had rainfall and evapotranspiration as regressors were significant ($p < 0.05$). The regressors of these models were rain sum (Cs) and mean rain sum (Csm), of the volume precipitated in 10 consecutive days; and daily (EToHS) and mean (EToHSm) evapotranspiration, calculated by the Hargreaves-Samani method. These models were also the ones that presented the lowest values of the Akaike information criterion (AIC): 103.26 for Cs, 102.31 for Csm, 103.85 for EToHS and 103.48 for EToHSm. The odds ratio (OR) results showed that for each variation of 1 mm.day⁻¹ (EToHS and EToHSm), or 1mm.10 days⁻¹ (Cs and Csm), the chances of occurrence of at least 1% flowering decreases by 0.9981 (Cs) and 0.9979 (Csm) times; and increase 2,3279 (EToHS) and 2,9066 (EToHSm) times. From the confusion matrix, the calculated accuracies of these models were 64.94% (Cs), 66.23% (Csm), 71.43% (EToHS) and 72.73% (EToHSm). Analyzing the Receiver Operating Characteristic Curve (ROC curve), the area under the curve (AUC) values were 0.618 (Cs), 0.640 (Csm), 0.649 (EToHS) and 0.662 (EToHSm). The specificity values of the models were 0.312 (Cs), 0.656 (Csm), 0.531 (EToHS) and 0.438 (EToHSm). The sensitivity values were 1.0 (Cs), 0.644 (Csm),

0.844 (EToHS) and 0.933 (EToHSm). The models with Cs and Csm presented higher Rho^2 with values of 8.92% (Cs) and 10.47% (Csm) against 7.95% (EToHS) and 8.56% (EToHSm). Analyzing the overall performance, among the models that had rain as a regressor, the model with Csm had a slightly higher overall performance than the one with Cs. Whereas, among the models that had evapotranspiration as regressors, the model with EToHSm had a slightly higher overall performance than the one with EToHS. The results suggest that the influence of rain on sugarcane flowering occurs in an accumulative way, while evapotranspiration affects it in an average way.

Keywords: *Saccharum spp.* Validation. Model intercomparison. Logistic Regression. Proxies.

LISTA DE ILUSTRAÇÕES

Figura 1 Processo de preenchimento de falhas das variáveis meteorológicas em estudo, evidenciando as duas validações.	39
Figura 2. Espaçamento inicial do plantio de cana-de-açúcar na área experimental.	42
Figura 3 Fenologia da Cana-de-açúcar.	42
Gráfico 1 A) Climatologia de Barra de São Miguel (AL). B) Balanço hídrico climatológico de Maceió (AL)	33
Gráfico 2 Porcentagem de falhas nos dados meteorológicos da estação de Barra de São Miguel (AL).	35
Gráfico 3 Porcentagem de falhas nos dados meteorológicos da estação de Coruripe (AL)	39
Gráfico 4 Número de inflorescências observadas em campo para cada uma das variedades de cana-de-açúcar do experimento nos anos de: A) 2014, B) 2015, C) 2016 e D) 2017.	43
Gráfico 5 Curva de retenção característica para Latossolo Amarelo.	46
Gráfico 6 Gráfico de calendário evidenciando: A) média do volume de chuvas, em milímetros, de 10 dias consecutivos (Cm) e B) média do volume acumulado chuvas, em milímetros, de 10 dias consecutivos (Csm).	53
Gráfico 7 Comportamento da evapotranspiração de referência (ET _o), calculada através de diferentes métodos, para Barra de São Miguel (AL)	66
Gráfico 8 Comportamento da chuva, ET _o e ARM para Barra de São Miguel (AL)	68
Gráfico 9 Balanço hídrico sequencial diário, dos anos de 2014 a meados de 2017, para Barra de São Miguel (AL)	69

Gráfico 10 Gráfico de probabilidade, evidenciando função de distribuição acumulada estimada (FDA), a partir dos dados das variáveis independentes observadas e estimadas.	71
Gráfico 11 Curva Logit, tendo como regressor: A) Chuva soma; B) Chuva média; C) Evapotranspiração por Hargreaves-Samani e D) Média da evapotranspiração por Hargreaves-Samani.	73
Gráfico 12 Resultados da curva ROC por modelo: A) Cs; B) Csm; C) EToHS e D) EToHSm.	75
Mapa 1 Localização do município de Barra de São Miguel (AL)	29
Mapa 2 A) Geomorfologia de Alagoas. B) Modelo de elevação digital de Alagoas.	30
Mapa 3 A) Mapa de solos de Alagoas. B) Detalhe dos solos presentes em Barra de São Miguel.	31
Mapa 4 Área plantada com cana-de-açúcar em Alagoas (2019)	32
Mapa 5 Mesorregiões de Alagoas.	38

LISTA DE TABELAS

Tabela 1 Siglas e respectivos solos e superfícies presentes no Mapa 3.	32
Tabela 2 Especificação dos bancos de dados usados para preenchimento e validação dos dados do Inmet.	36
Tabela 3 Estatísticas de validação de dados meteorológicos para preenchimento.	37
Tabela 4 Falhas nos dados meteorológicos da estação de Maceió (AL)	38
Tabela 5 Características gerais, ano de plantio e número de talhões das variedades de cana-de-açúcar usadas neste trabalho.	40
Tabela 6 Parâmetros da equação de Van Genuchten para construção da curva de retenção.	45
Tabela 7 Parâmetros das equações de Penman-Monteith, Priestley-Taylor e Hargreaves-Samani para obtenção da evapotranspiração de referência.	48
Tabela 8 Regressores usados na modelagem do florescimento da cana-de-açúcar.	53
Tabela 9 Modelo teórico de uma Matriz de confusão.	56
Tabela 10 Resultados por estatística de validação das variáveis meteorológicas da estação do INMET em Coruripe (AL)	59
Tabela 11 Resultados por estatística de validação das variáveis meteorológicas da estação local de Barra de São Miguel (AL)	61
Tabela 12 Resultados por estatística de validação da evapotranspiração de referência (ET _o)	66

Tabela 13 Resultados dos testes de Shapiro-Wilk e Lilliefors, ao nível de significância de 5%.	70
Tabela 14 Resultado da regressão Logit por regressor.	72
Tabela 15 Resultado da razão de chances (OR) e intervalo de confiança (CI) por modelo.	74
Tabela 16 Resultados da Matriz de confusão por modelo.	74
Tabela 17 Resultado de "pseudo" R^2 , RMSE e MAE por modelo.	76

LISTA DE EQUAÇÕES

Equação 1 Coeficiente de correlação de Pearson.	37
Equação 2 Coeficiente de correlação.	37
Equação 3 Viés.	37
Equação 4 Erro quadrático médio.	37
Equação 5 Erro médio absoluto.	37
Equação 6 Porcentagem de erro relativo médio.	37
Equação 7 Equação de Van Genuchten.	45
Equação 8 Capacidade de água disponível.	47
Equação 9 Evapotranspiração por Penman-Monteith.	48
Equação 10 Evapotranspiração por Hargreaves-Samani.	48
Equação 11 Parâmetro Q_0 da Equação 10.	48
Equação 12 Teste de Shapiro-Wilk.	50
Equação 13 Parâmetro b do teste de Shapiro-Wilk.	50
Equação 14 Teste de Lilliefors.	50
Equação 15 Modelo logit.	54
Equação 16 Critério de Informação de Akaike.	55
Equação 17 Acurácia.	56
Equação 18 Sensibilidade.	57
Equação 19 Especificidade.	57
Equação 20 “Pseudo” R^2 por Nagelkerke.	58
Modelo 1 Modelo logit com C_s como regressor.	73
Modelo 2 Modelo logit com C_{sm} como regressor.	73
Modelo 3 Modelo logit com EToHS como regressor.	73

Modelo 4 Modelo logit com EToHSm como regressor.	73
Modelo 3.1 Modelo 3 em função da temperatura.	77
Modelo 4.1 Modelo 4 em função da temperatura.	77

LISTA DE SIGLAS E ABREVIATURAS

Ageitec	Agência Embrapa de Informação Tecnológica
AIC	Critério de Informação de Akaike
AL	Alagoas
AUC	<i>Area Under the ROC Curve</i>
CAD	Capacidade de água disponível
CC	Capacidade de campo
Ceca	Centro de Ciências Agrárias
CO ₂	Dióxido de Carbono
Conab	Companhia Nacional de Abastecimento
Copersucar	Cooperativa brasileira de açúcar e etanol
ECMWF	<i>European Centre for Medium-Range Weather Forecasts</i>
Embrapa	Empresa Brasileira de Pesquisa Agropecuária
ERA5	<i>Fifth generation ECMWF reanalysis</i>
ET _o	Evapotranspiração de referência
FAO	<i>Food and Agriculture Organization</i>
GD	Graus-dia
GEE	Gases de Efeito Estufa
Inmet	Instituto Nacional de Meteorologia
MAE	Erro médio absoluto
NCAR	<i>National Center for Atmospheric Research</i>
NCEP	<i>National Centers for Environmental Prediction</i>
NOAA	<i>National Oceanic and Atmospheric Administration</i>
OGTR	<i>Office of the Gene Technology Regulator</i>

OMM	Organização Meteorológica Mundial
PMGCA	Programa de melhoramento genético da cana-de-açúcar
PMP	Ponto de murcha permanente
PMRE	Porcentagem de erro relativo médio
RAD	Radiação
Ridesa	Rede interuniversitária para o desenvolvimento do setor sucroalcooleiro
RMSE	Erro médio quadrático
ROC	<i>Receiver Operating Characteristic Curve</i>
Sindaçúcar	Sindicato da Indústria do Açúcar e do Alcool
Socicana	Associação dos Fornecedores de Cana de Guariba
SP	São Paulo
SRTM	<i>Shuttle Radar Topography Mission</i>
Tmax	Temperatura máxima
Tmed	Temperatura média
Tmin	Temperatura mínima
U ₂	Velocidade do vento a 2 metros de altura
UDOP	União Nacional da Bioenergia
UFAL	Universidade Federal de Alagoas

LISTA DE SÍMBOLOS

%	Porcentagem
°C	Graus Celsius
°S	Graus ao Sul
°W	Graus a oeste
arccos	Arco cosseno
cos	Cosseno
e	Média estimada
es – e	Déficit de pressão de vapor
G	Fluxo de calor no solo
γ	Constante psicrométrica
h	Horas
km	Quilômetros
KPa	Quilo Pascal
m	Metros
m	Média mensurada
$m\ s^{-1}$	Metros por segundo
$m^3.m^{-3}$	Metro cúbico por metro cúbico
$MJ\ m^{-2}\ dia^{-1}$	Megajoule por metro quadrado por dia
mm	Milímetros
$mm\ dia^{-1}$	Milímetro por dia
$Mw\ m^{-2}$	Megawatt por metro quadrado
N	Norte
n	Número de observações

N°	Número
p	Parâmetro empírico de forma da curva de retenção de solos
q	Parâmetro empírico de forma da curva de retenção de solos
r	Coefficiente de correlação de Pearson
R^2	Coefficiente de correlação
R_n	Saldo de radiação
sen	Seno
tg	Tangente
α	Inverso do valor de entrada de ar no solo
β	Viés
Δ	Declividade da curva de pressão de vapor de saturação
η	Horas de luz
θ	Conteúdo volumétrico de água
π	Número irracional ($\pi = 3,1415926535897932\dots$)
Ψ	Valor absoluto do potencial mátrico

SUMÁRIO

1. INTRODUÇÃO	20
2. REVISÃO DE LITERATURA	23
3. METODOLOGIA	27
3.1. Área experimental: localização e aspectos edáficos	28
3.2. Aspectos Climáticos e Dados Meteorológicos	32
3.3. Dados da cultura	40
3.4 Análise dos dados	45
3.4.1 Construção da curva de retenção e obtenção da CAD.....	45
3.4.2 Cálculo da evapotranspiração	47
3.4.3 Cálculo do balanço hídrico sequencial diário	49
3.4.4 Teste de normalidade	50
3.4.5. Modelagem	51
3.4.5.1 Regressores	52
3.4.5.2 Modelo qualitativo: Regressão Logística (Logit)	53
3.4.5.3 Avaliações de desempenho do modelo Logit	55
3.4.5.3.1 Matriz de Confusão	55
3.4.5.3.2 Curva ROC	56
3.4.5.3.3 “Pseudo” R ²	57
4. RESULTADOS E DISCUSSÕES	59
4.1 Validações de dados meteorológicos	59
4.2 Cálculos de evapotranspiração e balanço hídrico	65
4.3 Cálculos de normalidade dos dados de florescimento	69
4.4 Modelagens	71
5. CONCLUSÕES	79
REFERÊNCIAS	81
ANEXO 1 – Roteiro de cálculo de saldo de radiação (R _n) diário	94
ANEXO 2 – Roteiro de cálculo de balanço hídrico sequencial diário	96

1. INTRODUÇÃO

A cana-de-açúcar (*Saccharum spp.*) tem papel de destaque no cenário energético mundial, obtendo o *status* de *commodity*. Esta planta é matéria-prima para a produção do etanol e subprodutos associados, usados na fabricação de biocombustíveis (Conab, 2021). Os biocombustíveis são uma alternativa aos combustíveis fósseis e têm sido cada vez mais demandados (TEIXEIRA et al., 2016; HANAKI & PEREIRA-PORTUGAL, 2019). Isto, devido ao esgotamento das jazidas de petróleo, fonte dos combustíveis fósseis, o que tem aumentado seu preço (HANAKI & PEREIRA-PORTUGAL, 2019).

Aliado a isso, há aumento da consciência ambiental da sociedade civil, que rejeita os combustíveis fósseis com o intuito de mitigar as emissões de Gases de Efeito Estufa (GEEs), como o dióxido de carbono (CO₂) (OGTR, 2011; FILOSO et al., 2015; JAISWAL et al., 2017; HANAKI & PEREIRA-PORTUGAL, 2019). Essa tendência também é observada no Brasil (MACEDO et al., 2004; MARIN & NASSIF, 2013; CHERUBIN et al., 2016; OLIVEIRA et al., 2016).

O Brasil se destaca como o país que mais produz cana-de-açúcar no mundo, com 40% da produção mundial (CHAVES & FARIAS, 2008; GREGO et al., 2011; DALCHIAVON, 2013; MORAES et al., 2015; Conab, 2021). Na safra de 2020/2021 foram produzidas 654,80 milhões de toneladas (Conab, 2021). Deste total, a região Nordeste produziu 48,4 milhões de toneladas, com destaque para o estado de Alagoas, que foi responsável pela produção de 17 milhões de toneladas (Conab, 2021).

Especificamente no estado de Alagoas, a agroindústria de produção de etanol e açúcar tem grande impacto social e econômico. Segundo informações do Sindicato da Indústria do Açúcar e do Alcool no Estado de Alagoas (Sindaçúcar-AL), o setor gerou, só no início da safra de 2020/2021, 16,5 mil postos de trabalho. De modo geral, esse número durante toda safra, alcança entre 50 a 55 mil empregos no campo e nas indústrias (Sindaçúcar-AL, 2020). Estima-se, ainda, que 17,5 milhões de toneladas serão processadas nessa safra. Isso, a depender, dentre outras coisas, das condições climáticas (Sindaçúcar-AL, 2021). As baixas precipitações de 2015 e 2016, por exemplo, afetaram substancialmente a safra de 2017/2018. A colheita da safra de 2017/18, em decorrência dessas baixas precipitações, apresentou a menor produção

e produtividade da série histórica de Alagoas (Conab, 2017). Precipitações médias e altas também reduzem a produtividade da cana-de-açúcar, visto que, dentre vários fatores, induzem o florescimento.

Quando o objetivo da produção é o melhoramento genético, o florescimento é encarado como processo benéfico, pois, a partir dele, obtém-se sementes (SILVA & CESAR, 2003; AMARAL et al., 2012; CARABALLOSO et al., 2012; ABU-ELLAIL et al., 2019). O melhoramento genético clássico da cana-de-açúcar é prejudicado pela falta de sincronização do florescimento entre as plantas. Isso limita a escolha de progenitores e, por consequência, restringe as possibilidades de transferência de novas características. Sendo assim, o conhecimento das condições ambientais envolvidas no processo de indução floral e seu controle é de extrema importância para o desenvolvimento de novas cultivares de cana-de-açúcar pelo melhoramento clássico (GLASSOP et al., 2014).

No entanto, o florescimento da cana-de-açúcar é considerado prejudicial quando o objetivo da produção é a fabricação de biocombustíveis. Isto porque, nesse processo, a cana-de-açúcar desvia as reservas energéticas para a produção das panículas, ao invés de utilizar para o acúmulo de sacarose, que é convertida em etanol pela indústria (SILVA & CESAR, 2003; ARALDI et al., 2010; AMARAL et al., 2012; CARABALLOSO et al., 2012; OGTR, 2011). Aliado a isso, o florescimento aumenta o teor de fibras, fenômeno conhecido como isoporização ou chocamento da cana, o que prejudica a colheita mecanizada e o processamento (ARALDI et al., 2010, SILVA & SILVA, 2012; ABU- ELLAIL et al., 2019).

Adicionalmente, a instabilização do BRIX da cana-de-açúcar é um efeito adverso causado pelo florescimento (Almeida et al., 1945). Essa mesma conclusão foi obtida por Silva Neto et al. (2011), em relação aos 'efeitos colaterais' do florescimento. Em termos de quantificação de perdas, o florescimento reduz o rendimento da cana-de-açúcar em aproximadamente 30%. Assim, por exemplo, se há uma estimativa de produção de 100 toneladas por hectare, com plantas florescidas, essa produção cai para 70 toneladas por hectare (ALVES, 2016). Experimentos no Havaí mostraram que, se 35% das plantas em campo florescerem, há uma redução de 6 a 7 toneladas de cana por hectare (HUMBERT, 1974).

O fenômeno do florescimento da cana-de-açúcar é dependente da disponibilidade hídrica no período de indução floral. Dessa forma, chuvas e água armazenada no solo afetam tal fenômeno. A razão é que as chuvas têm correlação

positiva com o volume de água acumulado no solo e é do solo que as raízes da cana-de-açúcar retiram água para a manutenção do ciclo de vida. Além disso, a evapotranspiração aumenta com o volume precipitado de chuvas e aumento de temperatura, reduzindo o volume acumulado de água no solo.

Apesar destas referências sobre a relação entre chuva, umidade no solo e florescimento, não foi identificada uma estimativa da lâmina de água armazenada no solo, chuva ou evapotranspiração, que induz o florescimento. Esse conhecimento auxiliaria, por exemplo, no manejo do uso dos maturadores, como o ethephon. O ethephon atua liberando etileno, que em contato com o tecido vegetal é capaz de suprimir o desenvolvimento da gema floral. Mas a eficiência do seu uso está associada com a dose e a época de aplicação. Quando mais precoce a aplicação em condições de florescimento, mais eficiente (DEUBER & IRVINE, 1987; CAMPOS, 1992; CASTRO et al., 2001; CAPUTO et al., 2007; LEONEZI et al., 2016).

Sabendo da importância econômica e social da cultura da cana-de-açúcar para Alagoas e do fato de que o florescimento é um fenômeno que ocorre naturalmente nesse estado, fica evidente a necessidade de conhecimentos e tecnologias que auxiliem os produtores no manejo da cultura. Isso, principalmente porque muitos produtores têm deixado a atividade para plantar eucalipto, devido à desvalorização do açúcar (LIMA & BARBOSA, 2021). Além disso, parte do sucesso da atividade canavieira em Alagoas é explicado pelo clima adequado para o cultivo. No entanto, esse clima vem mudando. O sexto relatório de avaliação (AR6) do IPCC (Painel Intergovernamental de Mudanças Climáticas) aponta, com média confiança, que no Nordeste brasileiro haverá um aumento na duração do período seco (IPCC, 2021).

Assim, modelos que quantifiquem o efeito da disponibilidade hídrica sobre a produção da cana-de-açúcar é de fundamental importância para o desenvolvimento da cultura em Alagoas. Além disso, pode auxiliar a tomada de decisão dos agricultores e agentes públicos no âmbito da previsão de safra e contratos futuros.

Assim sendo, o objetivo principal deste trabalho foi estimar a lâmina de água acumulada no solo, chuvas ou evapotranspiração, ou seja, a disponibilidade hídrica que induz o florescimento da cana-de-açúcar, na região dos Tabuleiros Costeiros no estado de Alagoas, a partir do balanço hídrico do solo, observações meteorológicas da região e modelagens de regressão estatística.

2. REVISÃO DE LITERATURA

A floração da cana-de-açúcar afeta negativamente o rendimento e a qualidade da cana (BERDING & HURNEY, 2005). Isso, porque, coloca um limite máximo no crescimento da cultura, devido a limitação da produção de fitômeros no ápice do eixo principal (colmo). Dessa forma, causa mau uso da radiação interceptada em termos de acumulação de biomassa (PHILIPPUS et al., 2010).

O florescimento é induzido por vários fatores, incluindo aspectos meteorológicos, tais como o fotoperíodo, temperatura e umidade do ar, radiação solar e chuva, como também a maturidade da planta e a nutrição mineral. (COLEMAN, 1963; CLEMENTS & AWADA, 1967; COLEMAN, 1968; JULIEN et al., 1977; PURSEGLOVE, 1972 apud OGTR, 2011; ARALDI et al., 2010; DURAI et al., 2014; PATIL et al., 2015). Em regiões com climas tropical e equatorial, como é o caso do Nordeste brasileiro, o florescimento precoce é mais frequente do que em regiões de clima subtropical (Sudeste), principalmente devido a maior frequência de altas temperaturas do ar e do fotoperíodo, próximo de doze horas ao longo de todo o ano (SHANMUGAVADIVU & RAO, 2009; ARALDI et al., 2010). O florescimento, nessas regiões, vai depender, quase que exclusivamente, de oscilações nas precipitações (SRIVASTAVA et al., 2006; ARALDI et al., 2010). Isso, porque, o estresse hídrico pode atrasar a floração, se ocorrer antes que surjam as condições indutivas da flor (CARR & KNOX, 2011).

A época de indução do florescimento, constantemente relatada na literatura, concentra-se entre os meses de abril a julho. Essa época corresponde ao período de chuvas em Alagoas, sendo as condições edafoclimáticas mais favoráveis e, portanto, é a época mais provável que ocorra a indução (ARALDI et al., 2010). Amaral e colaboradores (2012), trabalhando com armazenamento de pólen de cana para fins de melhoramento genético no estado de Alagoas, classificaram as variedades de acordo com a época e porcentagem de florescimento. Nessa classificação, inicialmente florescem as variedades de ciclo precoce (16%); depois as intermediárias (43%) e por fim as tardias (34%), sendo que algumas variedades são ainda relutantes ao florescimento (7%).

Existem poucas referências na literatura sobre qual volume de chuva induz o florescimento da cana (ARALDI et al., 2010; TOPPA et al., 2010; CHEAVEGATTI-

GIANOTTO, 2011; SILVA NETO et al., 2011; CARDOZO & SENTELHAS, 2013). A lâmina de água irrigada apresenta correlação positiva direta com o florescimento da cana-de-açúcar (GOSNELL, 1973). Em experimento irrigado na África, em 1969, com diversas variedades, esse mesmo autor obteve 5 flores por hectare, quando a irrigação foi de 457 mm e 57 flores por hectare, quando a irrigação foi de 1219 mm.

Já, em 1972, os valores de florescimento foram muito maiores, segundo o autor supracitado, devido às precipitações, que chegaram à 430 mm no período de janeiro a março. Naquele ano foram obtidas 620 flores por hectare, quando a irrigação foi de 306 mm e 1263 flores por hectare, quando a irrigação foi de 867 mm. Lâminas de irrigação superiores à 867mm tenderam a reduzir o número de flores por hectare.

A relação entre a média de chuva em determinado período com o florescimento foi estabelecida em experimento conduzido em Araras (SP) (PEREIRA et al., 1983). Os autores verificaram que quando o volume precipitado era em média 200 mm em 10 dias, havia florescimento. Por outro lado, quando o volume precipitado era de 65 mm em 6 dias, não havia florescimento. Outros autores fazem referência à umidade do solo e a relação com a indução do florescimento da cana. Em condições de baixa umidade do solo, a porcentagem de brotação dos colmos florescidos foi menor que a dos colmos não florescidos (BARBIERI & SILVA, 2011).

A relação entre temperatura média do ar e o florescimento também foi estudada, o resultado foi um modelo de previsão de florescimento baseado em temperatura (Pereira, 1985). Esse modelo e pesquisas posteriores, como a de Segato & Pereira (2006), evidenciaram que a indução do florescimento ocorre com maior frequência com temperaturas na faixa de 18°C a 31°C, desde que haja umidade suficiente no solo (2/3 da capacidade de campo). Temperaturas nessa faixa devem ocorrer em, no mínimo, 10 dias, não necessariamente consecutivos, para que ocorra a indução. Uma outra medida associada à temperatura, que também foi estudada, foi o Grau-dia. Isler e colaboradores (2014), trabalhando com modelos computacionais, chegaram à conclusão que com 1100 Graus-dia acumulados (GDA) o risco de florescimento é muito alto. Como a disponibilidade hídrica e temperatura são fatores ambientais difíceis de controlar, substâncias químicas têm sido usadas no controle do florescimento, os chamados maturadores.

A aplicação de maturadores é usada com o intuito de mitigar os efeitos da floração. O maturador mais comumente usado em campo e relatado na literatura é o Ethephon, que estimula a biossíntese de etileno na cana. A eficácia e eficiência do

Ethephon sobre a inibição do florescimento depende de fatores, como a época de aplicação e condições ambientais.

Com relação à época de aplicação, Bertollo (2020) constatou que a expressão do genes ligados ao florescimento não foi significativa após a aplicação do ethephon. O gene avaliado foi o ScFT3, que é ortólogo ao *FLOWERING LOCUS T* (FT) em cana-de-açúcar. A pesquisa foi feita em Ribeirão Preto, São Paulo, e avaliada 15 e 30 dias após a aplicação do ethephon, que aconteceu em fevereiro. Corroborando com esse resultado, Cardozo (2017) concluiu que a antecipação da aplicação do ethephon (acima de 25 a 30 dias) mostrou-se ineficiente no controle do florescimento. Por outro lado, o ponto ótimo de controle do florescimento ocorreu quando a aplicação foi realizada entre 9 a 6 dias antes do início do período de indução. Esses resultados evidenciam a necessidade de identificação da época correta de aplicação e a diferença de respostas entre cultivares.

Com relação às condições ambientais, Donaldson & Singels (2004) verificaram que plantios irrigados de cana-de-açúcar têm maior probabilidade de florescer que plantios em sequeiro. Com esse resultado, os autores concluíram que aplicações anuais de ethephon seriam economicamente viáveis para cultivares com floração moderada a abundante, enquanto que, para cultivares de floração rara, as aplicações se justificam apenas nos anos em que a floração é prevista, ou seja, com chuva e temperatura favoráveis. Isso evidencia a complexidade do processo fisiológico de florescimento, uma vez que envolve fatores genéticos e ambientais (ARALDI et al., 2010; GLASSOP et al., 2013; MELLONI et al., 2015).

Esse processo, que resulta na produção de uma inflorescência, envolve uma ou mais vias, incluindo a percepção do fotoperíodo por fitocromos; a maturidade da planta; hormônios vegetais, como a giberelina e vias autônomas. A indução e/ou repressão de genes ao longo dessas vias é mediada por flórigenos. A identificação desses flórigenos é essencial no entendimento das condições indutivas e não indutivas; para seleção do momento de detecção de mudanças durante a indução floral; no entendimento das diferenças entre variedades responsivas e não responsivas à indução e para identificar genes que podem ser manipulados para aumentar ou inibir a floração (GLASSOP & RAE, 2019).

Embora esse flórigeno não tenha sido identificado, várias pesquisas apontam genes envolvidos nas vias de florescimento e possíveis candidatos à flórigeno. Como

Medeiros (2010), que inferiu que o gene CYP51 (Citocromo P450) pode ser utilizado como marcador molecular para o melhoramento clássico, ou seja, usando sementes, que vieram de flores. Coelho et al. (2013), trabalhando com bancos de dados genéticos, identificaram supostos genes envolvidos na via de floração da cana-de-açúcar controlada por fotoperíodo. O uso de DNA's complementares (cDNAs) também foi estudado.

Souza (2014), por meio de análises *in silico*, prospectou cDNAs que apresentaram similaridade para calmodulina, fator de transcrição NAC e fosfatidilinositol. Essas substâncias já haviam sido descritas na literatura como tendo um papel no processo de desenvolvimento floral de outras plantas, mas não para a cana-de-açúcar. Sousa (2015), trabalhando com análises genômicas do florescimento da cana, constatou que existem pelo menos dois genes distintos homólogos à proteína HINT em cana-de-açúcar. Este trabalho também constatou que a expressão diferencial de HINT está relacionada com fatores climáticos presentes na região Nordeste do Brasil, estresses bióticos e hormônios vegetais.

Além destes trabalhos, Glassop e Rae (2019) verificaram mudanças significativas no nível de expressão de 16 genes durante um ciclo de 24 horas de claro-escuro. Santos (2020) validou 8 diferentes candidatos a genes de referência envolvidos na indução floral fotoperiódica (*UBQ1*, *UBQ2*, *RPL*, *GAPDH*, *25SrRNA1*, *TUB*, *EF1- α* e *TIPS-41*). Os resultados destes trabalhos podem ajudar na identificação de genes associados na percepção do fotoperíodo, que é um elemento importante para regiões que estejam localizadas em latitudes distantes do Equador.

3. METODOLOGIA

Este estudo foi realizado em 5 etapas: (i) definição dos parâmetros do solo para obtenção da capacidade de água disponível no solo (CAD); (ii) cálculo da evapotranspiração; (iii) cálculo do balanço hídrico sequencial diário para obtenção da água armazenada no solo (ARM); (iv) teste de normalidade das variáveis independentes (regressores) e (v) estimativa do modelo de regressão logística (logit). Inicialmente, tem-se a definição dos parâmetros do solo da estação experimental localizada em Barra de São Miguel (AL). Esses parâmetros foram utilizados no cálculo do ponto de murcha permanente (PMP) e capacidade de campo (CC), ambos em $\text{m}^3.\text{m}^{-3}.\text{dia}$.

O PMP e a CC foram utilizados no cálculo da capacidade de água disponível no solo (CAD) em $\text{m}^3.\text{m}^{-3}.\text{dia}^{-1}$. Além da CAD, a evapotranspiração também foi calculada. Os métodos usados para o cálculo da evapotranspiração foram Penman-Monteith (ET_o) e Hargreaves-Samani (ET_oHS) em $\text{mm}.\text{dia}^{-1}$. As chuvas observadas (C) em $\text{mm}.\text{dia}^{-1}$, a CAD, a ET_o e a ET_oHS foram usadas nas estimativas do balanço hídrico sequencial diário. A partir dessas estimativas, a quantidade de água armazenada no solo (ARM), em $\text{mm}.\text{dia}^{-1}$, pôde ser calculada. A partir dos cálculos supracitados, ARM, C, ET_o e ET_oHS foram usadas como *proxies* do fenômeno de florescimento.

Os dados de florescimento foram cedidos pelo programa de melhoramento genético da cana-de-açúcar (PMGCA) do Centro de Ciências Agrárias (Ceca) da Universidade Federal de Alagoas (UFAL). Esses dados foram coletados na estação de cruzamento de cana-de-açúcar, localizada em Barra de São Miguel (AL), no período de 2014 a 2017. Além dos dados de florescimento, variáveis meteorológicas como chuva ($\text{mm}.\text{dia}^{-1}$), temperatura máxima ($^{\circ}\text{C}$), temperatura média ($^{\circ}\text{C}$), temperatura mínima ($^{\circ}\text{C}$), radiação ($\text{MW}.\text{m}^{-2}$), velocidade do vento ($\text{m}.\text{s}^{-1}$) e umidade relativa (%) também foram cedidas para o mesmo período. Essas variáveis foram usadas no cálculos de evapotranspiração e balanço hídrico.

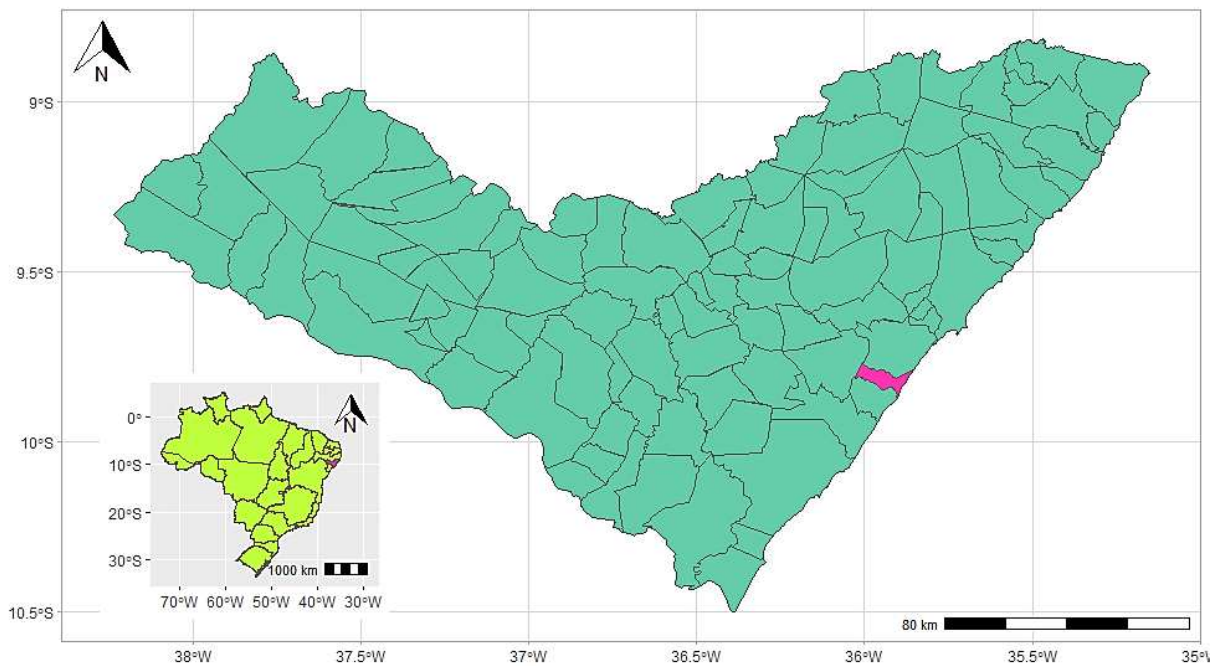
Depois de calculados, os volumes de ARM, C, ET_o e ET_oHS foram tratados. Esses tratamentos consistiram em calcular a média, o somatório e a média do somatório dessas variáveis, o que resultou em 4 situações por variável, totalizando 16 regressores. Cada um desses regressores foram analisados pelo modelo de

regressão logística, juntamente com os dados de florescimento transformados em variável binária. Isso, para verificar o quanto influenciam na floração da cana.

Essas análises indicaram quais regressores eram significativos ($P < 0,005$). Dessa forma, os modelos gerados a partir dos regressores significativos tiveram o desempenho testado. Esses testes indicaram qual o modelo que melhor explica a relação existente entre os regressores e o fenômeno do florescimento para a região de Tabuleiros Costeiros no estado de Alagoas. A seguir, cada uma dessas partes da metodologia é explicada com mais detalhes.

3.1. Área experimental: localização e aspectos edáficos

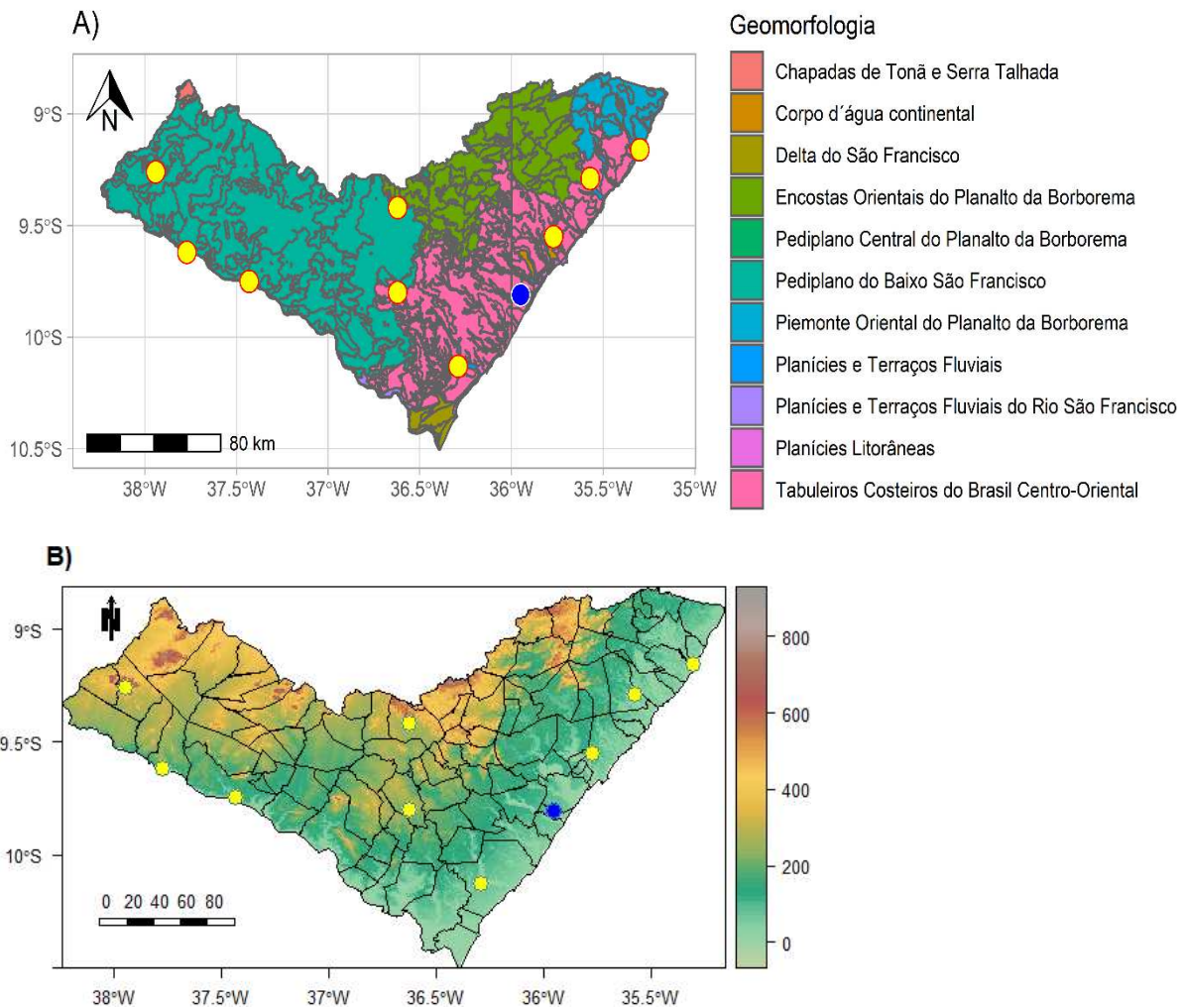
Os dados usados neste trabalho foram fornecidos pelo programa de melhoramento genético da cana-de-açúcar (PMGCA) do Centro de Ciências Agrárias (Ceca) da Universidade Federal de Alagoas (UFAL). Esses dados foram coletados na estação de cruzamento de cana-de-açúcar, localizada em Barra de São Miguel, município situado no litoral sudeste do estado de Alagoas, região Nordeste do Brasil ($9,81^{\circ}\text{S}$ de latitude, $35,95^{\circ}\text{W}$ de longitude) (Mapa 1). Essa localização, em termos geomorfológicos, está compreendida nos chamados Tabuleiros Costeiros (Mapa 2.A), que se caracterizam por apresentarem desde áreas de relevo plano a ondulados. Essa variação é oriunda do processo de formação deste ambiente, que foi formado por sedimentos do Grupo Barreiras (FERREIRA et al., 2015).



Mapa 1 - Localização do município de Barra de São Miguel (AL). Fonte: Autor.

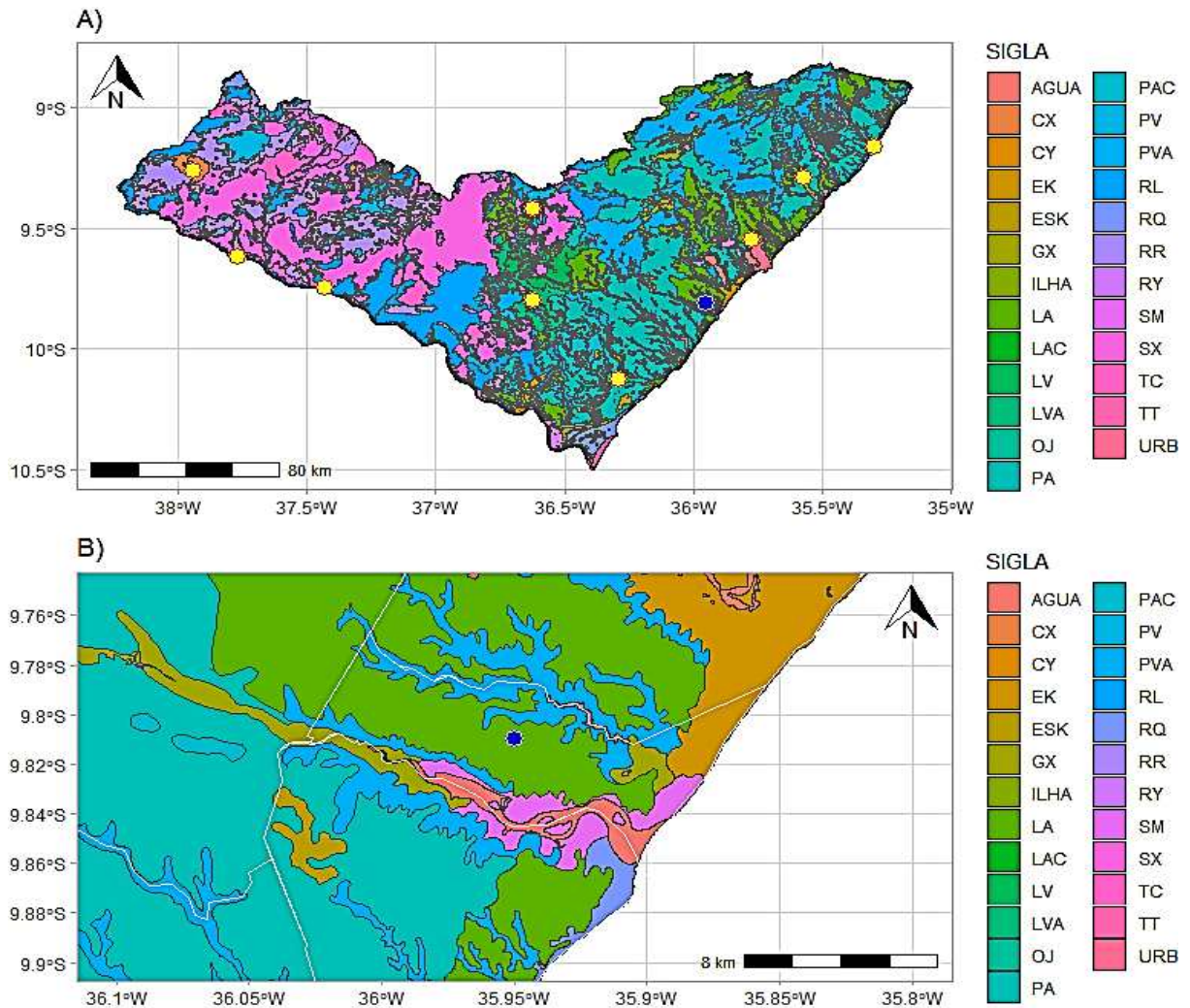
Esses sedimentos conferem grau de entalhamento diferencial à área, o que deu origem aos desníveis altimétricos, com altitudes que variam de 50-100 metros (Mapa 2.B). Nestes locais podemos encontrar morros testemunhos, planaltos de topo convexo, platôs de origem sedimentar, vales estreitos, encostas abruptas, encostas suaves e amplas várzeas. O depósito desses sedimentos, de textura média a argilosa, muito intemperizados, dão origem a solos coesos (LIMA NETO et al., 2009; EMBRAPA, 2017).

Esses solos são classificados como Argissolos e Latossolos Amarelos. Em Barra de São Miguel, nas coordenadas da estação de origem dos dados, os solos têm, em sua maioria, textura média/argilosa, classificado como Latossolo Amarelo (Mapa 3.A, 3.B e Tabela 1) (BARROS et al., 2012; DE ABREU et al., 2013). As características mais importantes do solo, para este trabalho, estão relacionadas com a água armazenada, que será efetivamente usada pela planta.



Mapa 2- A) Geomorfologia de Alagoas. B) Modelo de elevação digital de Alagoas. O ponto azul representa a localização da estação de Barra de São Miguel e os amarelos representam as estações do Inmet. Fonte: Autor com base em IBGE (2021) e SRTM (2021), respectivamente.

Essa quantidade de água armazenada, utilizada pela planta, está compreendida entre a capacidade de campo (CC) e o ponto de murcha permanente (PMP). É denominada de capacidade de água disponível (CAD), cuja disponibilidade depende do tipo de solo e das suas características de retenção de água. A partir das características do solo e das variáveis meteorológicas da região foi obtida uma curva de retenção característica de Latossolo Amarelo, o que é essencial no cálculo do balanço hídrico climatológico sequencial. Ainda, na região desses solos, que grande parte da atividade agrícola e urbana se desenvolvem.



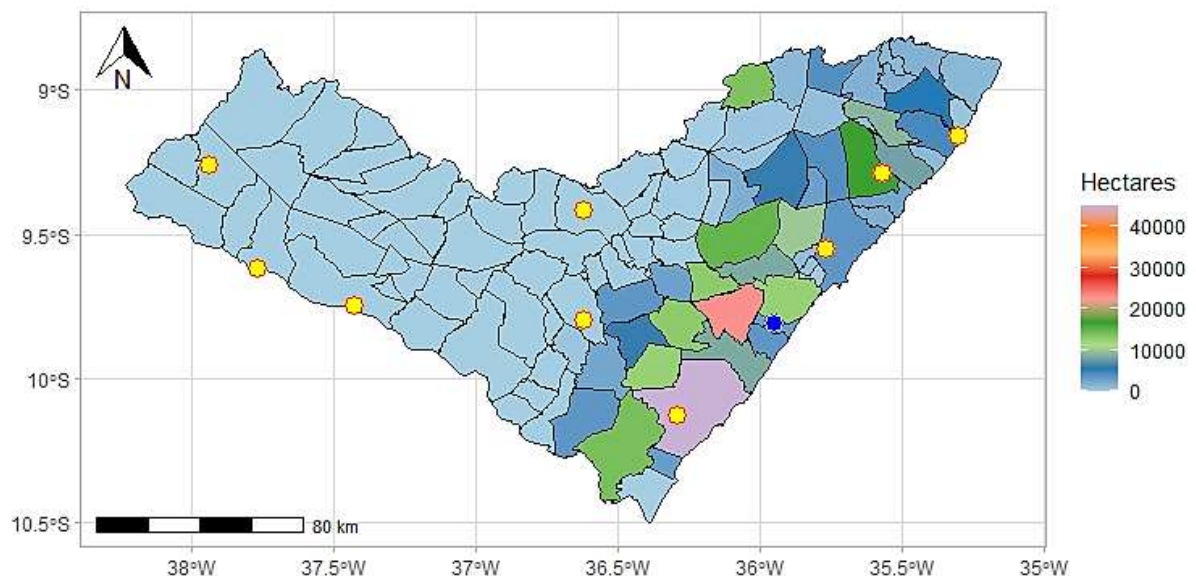
Mapa 3 - A) Mapa de solos de Alagoas. B) Detalhe dos solos presentes em Barra de São Miguel. O ponto azul representa a localização da estação de Barra de São Miguel e os amarelos representam as estações do INMET. Fonte: Autor com base em Embrapa (2021).

SIGLA	TIPO DE SOLO/ SUPERFÍCIE
ÁGUA	Água
CX	Cambissolos Háplicos
CY	Cambissolos Flúvicos
EK	Espodossolos Humilúvicos
ESK	Espodossolos Ferrihumiluvicos
GX	Gleissolos Háplicos
ILHA	Ilha
LA	Latossolos Amarelos
LAC	Latossolos Acinzentados
LV	Latossolos Vermelhos
LVA	Latossolos Vermelhos-Amarelos
OJ	Organossolos Tiomorficos
PA	Argissolos Amarelos
PAC	Argissolos Acinzentados
PV	Argissolos Vermelhos

PVA	Argissolos Vermelhos-Amarelos
RL	Neossolos Litólicos
RQ	Neossolos Quartzorênicos
RR	Neossolos Regolíticos
RY	Neossolos Flúvicos
SM	Solos de Mangue
SX	Planossolos Hápicos
TC	Luvissolos Crômicos
TT	Tipo De Terreno (Dunas Móveis)
URB	Área Urbana

Tabela 1 - Siglas e respectivos solos e superfícies presentes no Mapa 3. Fonte: Autor.

Devido à essa concentração de atividade agrícola e urbana, a região dos Tabuleiros Costeiros, área de origem dos dados, é representativa no que tange a importância, tanto do ponto de vista socioeconômico, como da produção agrícola, com destaque para a produção de cana-de-açúcar (CRUZ et al., 2017). É possível observar que grande parte da produção canavieira de Alagoas se desenvolve nessa área (Mapa 4). A produção de cana-de-açúcar nessa área enfrenta riscos. Os riscos mais frequentes estão associados às condições ambientais. Dentro destas condições, o tempo e o clima são considerados os mais expressivos. Esses podem alterar a disponibilidade de recursos naturais, como a água, essencial à produção da cana-de-açúcar.

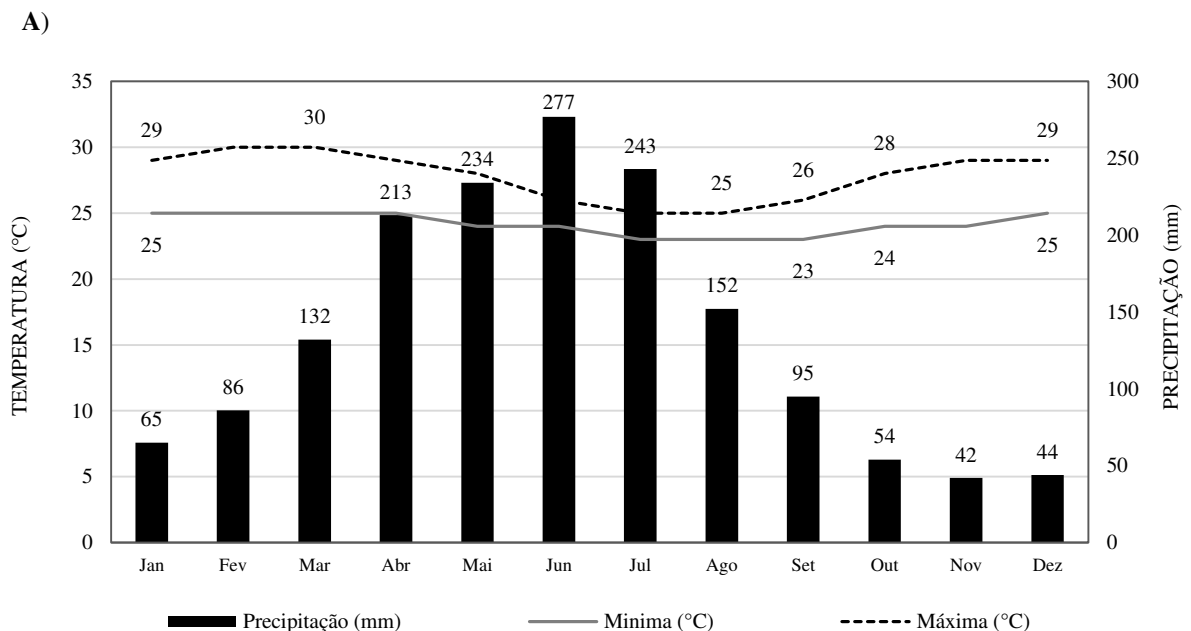


Mapa 4- Área plantada com cana-de-açúcar em Alagoas (2019). O ponto azul representa a localização da estação de Barra de São Miguel e os amarelos representam as estações do INMET. Fonte: Autor com base em IBGE (2020b).

3.2. Aspectos Climáticos e Dados Meteorológicos

O município de Barra de São Miguel, segundo a classificação de Thorthwaite, apresenta um clima subúmido (C2sA'a'), no qual ocorre maior umidade do solo durante todo o ano, quando comparado a outros tipos climáticos de Alagoas (BARROS et al., 2012). A chuva média anual da região é da ordem de 1.000 mm, mal distribuída, sendo o período de fevereiro a agosto correspondente a estação chuvosa (BARROS et al., 2012; DE ABREU et al., 2013; Climatempo, 2021).

A chuva total média na estação seca é da ordem de 300 mm, com valor mínimo no mês de novembro, cerca de 42 mm, e máximo no mês de janeiro, 65 mm. Na estação chuvosa a chuva total média é da ordem de 1400 mm, com valor máximo no mês de junho com 277 mm (Gráfico 1.A). A menor temperatura mínima média diária em Barra de São Miguel é em torno de 23°C e a máxima média é de 30°C, nos meses de julho e fevereiro, respectivamente (Gráfico 1.A) (Climatempo, 2021).



B)

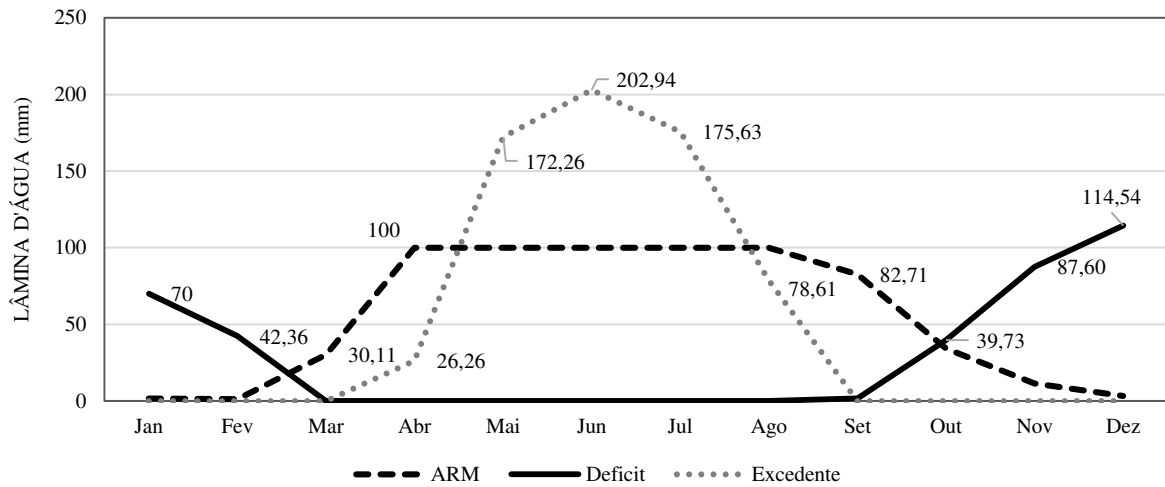


Gráfico 1 - A) Climatologia de Barra de São Miguel (AL). B) Balanço hídrico climatológico de Maceió (AL). Fonte: Autor com base em Climatempo (2020) e Inmet (2020), respectivamente.

O Balanço Hídrico Climatológico (1961 a 1990) para a capital de Alagoas, Maceió, situada à 30 km de distância de Barra de São Miguel, demonstra que o período de retirada de água do solo ocorre no início da estação seca, nos meses de outubro a janeiro. O déficit hídrico na região compreende os meses de outubro a fevereiro, ultrapassando 50 mm. O período de reposição hídrica inicia nos dois primeiros meses da estação chuvosa, março e abril, e os excedentes nos meses de abril a setembro, com valores totais em torno de 30 mm a 200 mm (BARROS et al., 2012) (Gráfico 1.B).

Dados meteorológicos também foram fornecidos pelo PMGCA da UFAL. Esses dados foram coletados diariamente no período de agosto de 2014 a dezembro de 2017. Os dados utilizados neste trabalho foram: temperatura máxima em °C (Tmax), temperatura mínima em °C (Tmin), temperatura média em °C (Tmed), radiação em Mw/m² (RAD), velocidade do vento a 2 metros de altura em ms⁻¹ (U₂), umidade relativa do ar em % (UR) e chuva total em mm (C). Esses dados foram coletados e armazenados por uma estação automática, instalada no mesmo local do plantio de cana-de-açúcar.

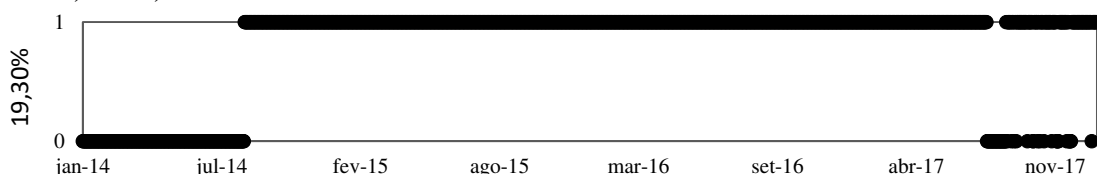
Os dados dessa estação local continham, de forma geral, falhas, lacunas e descontinuidade temporal em todas as variáveis meteorológicas, exceto a chuva (Gráfico 2). Essas falhas se concentram no início de 2014 e final de 2017. Assim, houve necessidade de validação com outras bases de dados e posterior preenchimento de falhas.

O processo de validação consistiu em comparar estatisticamente os dados meteorológicos da estação de Barra de São Miguel com dados mensurados. Os dados mensurados foram baixados dos *sites* de três bancos gradeados: ERA5, NCEP/NCAR e Xavier; além de duas estações meteorológicas do Instituto Nacional de Meteorologia (INMET), próximas a Barra de São Miguel: Coruripe e Maceió (Tabela 2). Esses dados, inicialmente, continham o período de 2014 a 2017, mas foram cortados, de forma que sobrasse o mesmo período de dados da estação de Barra de São Miguel (Gráfico 2). Esses cortes foram fundamentais para a aplicação das estatísticas de comparação, pois as séries em comparação precisam ter o mesmo tamanho.

As estatísticas utilizadas no processo de validação foram: coeficiente de correlação de Pearson (r), coeficiente de correlação (R^2), viés (β), erro quadrático médio (RMSE), erro médio absoluto (MAE) e porcentagem de erro relativo médio (PMRE) (Tabela 3). Essas estatísticas, de modo geral, comparam duas séries de dados, uma série de dados observados e uma série de dados mensurados e foram escolhidas por serem corriqueiramente usadas em trabalhos de validação de dados meteorológicos (RAMAPAZO et al., 2018; BABAR et al., 2019; ZULUAGA et al., 2021). A série de dados observados correspondeu aos dados da estação meteorológica de Barra de São Miguel, enquanto a série de dados mensurados correspondeu aos dados do ERA5, NCEP/NCAR, Xavier, INMET Coruripe e INMET Maceió. Dessa forma, uma tabela com os resultados das estatísticas, por variável, foi montada.

De acordo com os resultados dessas estatísticas, o banco de dados que apresentou maiores valores de r e R^2 e menores valores de β , RMSE, MAE e PMRE; foi o que melhor representou a variável observada, ou seja, apresentou melhor desempenho. Assim sendo, o banco de dados que apresentou melhor desempenho, por variável, foi usado no preenchimento das falhas.

A) Tmax, Tmed, Tmin e UR



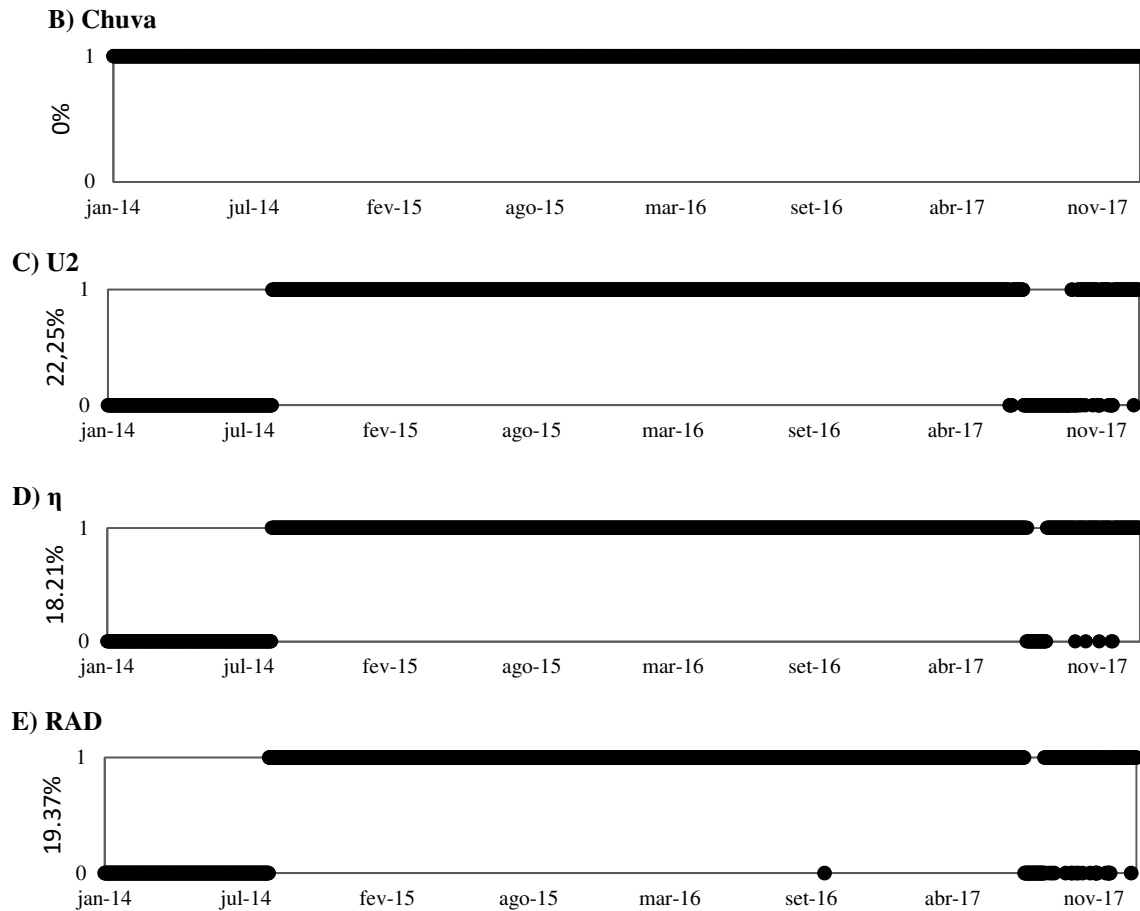


Gráfico 2 - Porcentagem de falhas nos dados meteorológicos da estação de Barra de São Miguel (AL). 1: representa presença de dados e 0: ausência de dados. Fonte: Autor.

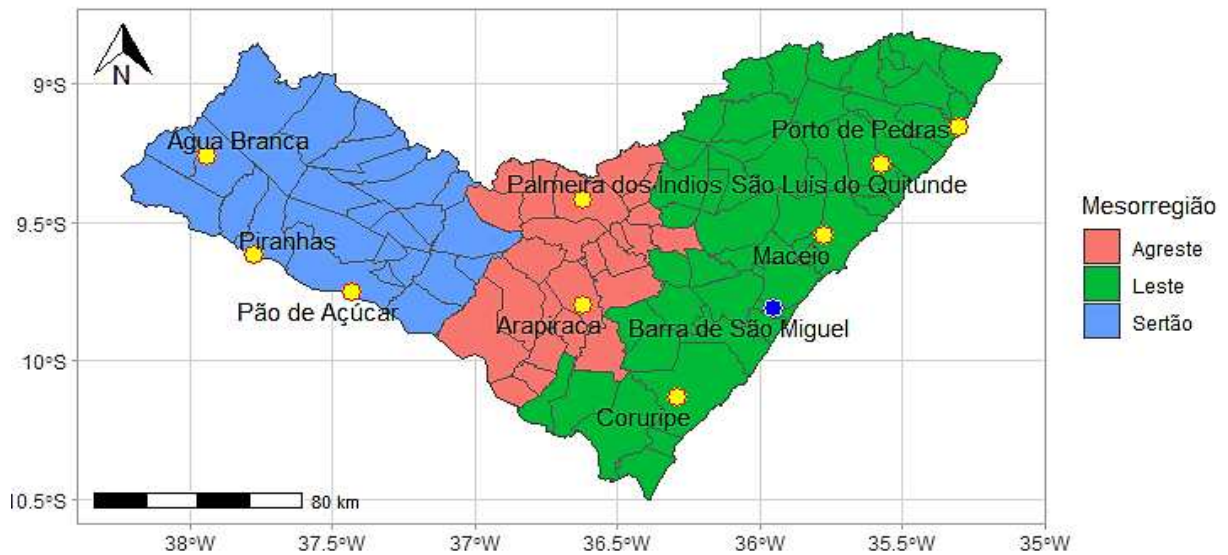
BANCO DE DADOS	INSTITUIÇÃO DE ORIGEM	RESOLUÇÃO
ERA5	ECMWF (<i>European Centre for Medium-Range Weather Forecasts</i>)	0,25° X 0,25°
NCEP/NCAR	NCEP (<i>National Centers for Environmental Prediction</i>) em parceria com o NCAR (<i>National Center for Atmospheric Research</i>)	2,5° x 2,5°
XAVIER	Xavier et al.	0,1° X 0,1°
Estação Coruripe	Instituto Nacional de Meteorologia (INMET)	Pontual
Estação Maceió	Instituto Nacional de Meteorologia (INMET)	Pontual

Tabela 2 - Especificação dos bancos de dados usados para preenchimento e validação dos dados do Inmet. Fonte: Autor com base em ECMWF (2020), NOAA (2020) e Xavier et al. (2016).

ESTATÍSTICA	FÓRMULA	Equação
r	$\frac{\sum_{i=1}^n (e_i - \bar{e}) \cdot (m_i - \bar{m})}{\sqrt{(\sum_{i=1}^n ((e_i - \bar{e}))^2) \cdot (\sum_{i=1}^n ((m_i - \bar{m}))^2)}}$	(1)
R²	$r \cdot (r)$	(2)
β	$\frac{1}{n} \sum_{i=1}^n (e_i - m_i)$	(3)
RMSE	$\sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - m_i)^2}$	(4)
MAE	$\frac{1}{n} \sum_{i=1}^n e_i - m_i $	(5)
PMRE	$\frac{100}{n} \sum_{i=1}^n \left \frac{e_i - m_i}{m_i} \right $	(6)

Tabela 3 - Estatísticas de validação de dados meteorológicos para preenchimento. As variáveis: e , m e n ; representam média estimada (calculada a partir dos bancos de dados), média observada (calculada a partir dos dados de Barra de São Miguel) e número de observações, respectivamente. Fonte: Autor.

As estações do INMET, localizadas em Maceió e Coruripe, foram escolhidas pela proximidade com Barra de São Miguel. A estação de Maceió fica a 30,8 km e a de Coruripe a 62,6 km de Barra de São Miguel (Google Maps, 2021). Além disso, pertencem à mesma mesorregião (Mapa 5). Esse critério pôde ser usado, pois a própria criação dessa mesorregião levou em consideração peculiaridades geográficas dos municípios que a compõe. Isso, como a caracterização baseada na Geografia Física da região (geologia, relevo, clima e vegetação natural) e secundariamente nos fatos da Geografia Humana (produção, comunicações, densidade demográfica, formação histórica, etc.) (BRASIL, 1941b; DINIZ, BATELLA, 2005). Nesse contexto, o uso dessas estações pôde ser feito sem muitas extrapolações.



Mapa 5 - Mesorregiões de Alagoas. O ponto azul representa a localização da estação de Barra de São Miguel e os amarelos representam as estações do INMET. Fonte: Autor com base em IBGE (2020a).

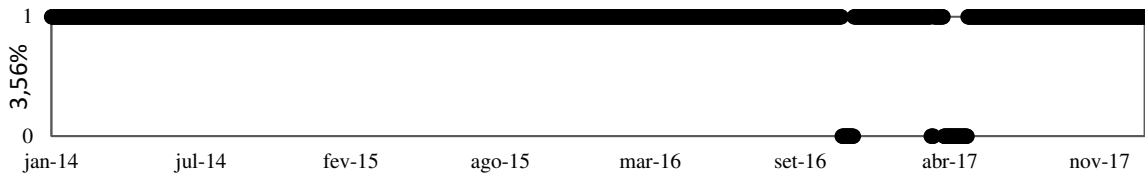
Os outros bancos de dados foram escolhidos pela resolução espacial (alta, média e baixa), disponibilidade de período de dados e variáveis (Tabela 2). Nesse processo, é importante salientar que os dados do INMET também apresentam falhas. Na estação de Maceió, essas falhas (Tabela 4) puderam ser corrigidas de forma simples, apenas retirando as horas faltantes e fazendo a média com os dados que restaram. Já na estação de Coruripe, esse processo foi mais complexo, pois haviam falhas de até um mês (Gráfico 3), o que exigiu uma validação, que foi feita de forma semelhante ao que foi feito para Barra de São Miguel, salvo algumas diferenças.

Essas diferenças foram: os (i) bancos de dados utilizados foram ERA5, NCEP/NCAR, e Xavier (ii) a série de dados observados, utilizada nas estatísticas de validação, foi a de Coruripe e as séries de dados mensurados foram obtidas dos bancos de dados (ERA5, NCEP/NCAR, e Xavier); (iii) As séries de dados observados foram cortadas de forma que tivessem o mesmo tamanho da série de dados da estação de Coruripe (Gráfico 3). Na Figura 1, observa-se o esquema dos dois processos de validação e preenchimento.

ANO	DIA/MÊS	FALHA
2014	07/04	Falha em todas variáveis entre 12:00 h e 13:00hrs.
2015	01/01	Falha em todas variáveis entre 10:00 h e 11:00 h.
	13/01	Falha em todas variáveis entre 17:00 h e 18:00 h.
	21/01	Falha em todas variáveis entre 21:00 h e 22h.
	09/08	Falha em todas variáveis entre 17:00 h e 18:00 h.
2017	30/10	Falha em todas variáveis entre 2:00 h e 3:00 h.

Tabela 4 - Falhas nos dados meteorológicos da estação de Maceió (AL). Fonte: Autor.

(A) Tmax, Tmin, UR, U₂ e η



(B) Tmed

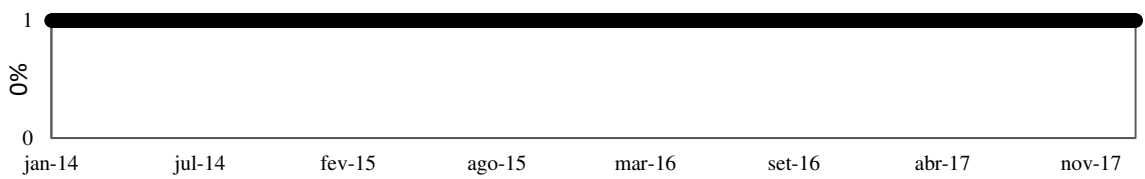


Gráfico 3 – Porcentagem de falhas nos dados meteorológicos da estação de Coruripe (AL). 1: representa a presença de dados e 0: ausência de dados. Fonte: Autor.

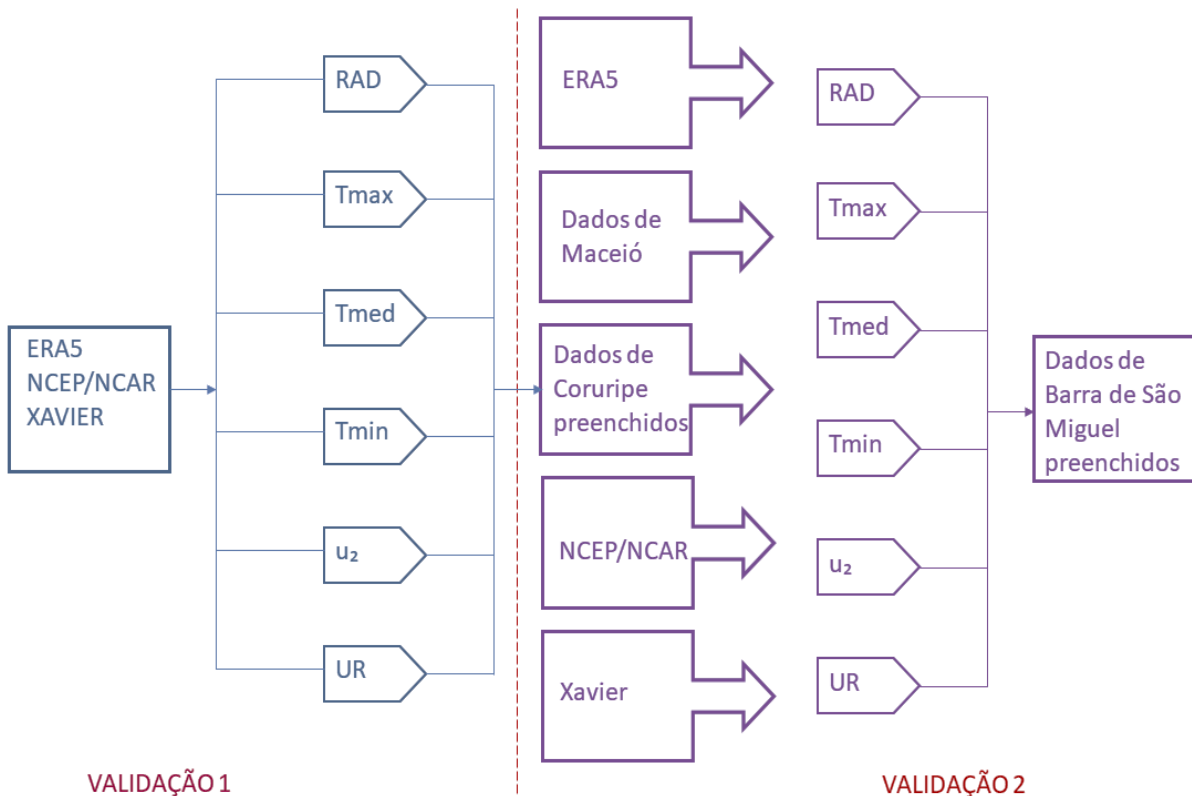


Figura 1 - Processo de preenchimento de falhas das variáveis meteorológicas em estudo, evidenciando as duas validações. Fonte: Autor.

3.3. Dados da cultura

As 24 variedades de cana-de-açúcar, das quais obteve-se os dados de florescimento, estão descritas na Tabela 5, bem como: característica de maturação, hábito de florescimento e ano de plantio. Entre essas variedades estão as do grupo RB e SP, que correspondem a 45,4% e 45,6% da área produzida de cana-de-açúcar em Alagoas, respectivamente. Todas as variedades identificadas com as iniciais RB são desenvolvidas pela Ridesa (Rede interuniversitária para o desenvolvimento do setor sucroalcooleiro). Enquanto as SP foram desenvolvidas pela Copersucar, a maior cooperativa brasileira de exportação de açúcar e etanol. Destacam-se as variedades RB92579 e SP79-1011, que juntas somam mais de 45% da área produzida nesse estado (Sindaçúcar, 2009). Dessa forma, as variedades presentes no banco de dados analisado são representativas da produção alagoana.

VARIEDADE DE CANA-DE-AÇÚCAR	MATURAÇÃO	FLORESCIMENTO	ANO DO PLANTIO	NÚMERO DE PARCELAS			
				2014	2015	2016	2017
1) IAC862210	Precoce	Eventual	2013	1	1	1	1
2) IAC873396	Precoce	Eventual	2014	0	1	1	1
3) IAC911099	Média tardia	Frequente	2014	0	1	1	2
4) IACSP942094	Média tardia	Raro	2014	0	1	1	1
5) IACSP955000	Média	Raro	2013	2	5	5	5
6) IACSP955094	Média	Raro	2014	0	2	2	2
7) IACSP974039	Precoce	Eventual	2014	0	3	3	3
8) RB72454	Média tardia	Eventual	2013	1	2	2	2
9) RB835486	Precoce média	Eventual	2013	1	1	1	1
10) RB855156	Precoce	Frequente	2014	0	1	1	1
11) RB855453	Precoce	Frequente	2014	0	2	2	2
12) RB855511	Precoce média	Eventual	2013	1	1	1	1
13) RB867515	Média tardia	Eventual	2014	0	2	2	2
14) RB92579	Média tardia	Eventual	2013	2	2	2	2
15) RB931011	Tardia	Frequente	2013	1	1	1	2
16) RB93509	Tardia	Frequente	2013	1	1	1	1
17) RB951541	Precoce	Eventual	2013	1	1	1	1
18) RB99395	Precoce	Eventual	2013	1	1	1	1
19) SP79-1011	Média tardia	Raro	2013	2	2	2	2
20) SP80-1842	Precoce	Eventual	2014	0	1	1	2
21) SP80-185	Média	Médio	2013	2	2	2	3
22) SP80-3280	Média	Eventual	2014	0	3	3	3
23) SP81-3250	Média	Eventual	2013	1	2	2	1
24) SP83-2847	Média	Frequente	2014	0	1	1	1

Tabela 5 - Características gerais, ano de plantio e número de talhões das variedades de cana-de-açúcar usadas neste trabalho. Fonte: Autor com base em RIDESA (2010), Ageitec (2020), Socicana (2020) e UDOP (2020).

O plantio da cana foi feito em parcelas com cana inteira. Cada parcela continha 5 metros e foi constituída de uma quantidade de linhas variáveis, a depender do ano e da variedade (Tabela 5). O espaçamento entre linhas foi de 1 metro e o espaçamento entre plantas variável, a depender do estágio fenológico (Figura 2). A fenologia da cana foi considerada em quatro fases: emergência/brotação, perfilhamento, desenvolvimento/crescimento da parte aérea e maturação (MARIN et al., 2009). No início do plantio e emergência, o espaçamento das socas foi de 0,50 m e, ao final do perfilhamento, houve o fechamento da linha (Figura 3).

Cada uma dessas fases pode sofrer atrasos ou adiantamentos a depender do material de origem: soca, ressoça, planta. A mudança de material de origem ocorre, pois, a cana é uma planta semiperene, que passa por vários cortes ao longo da vida útil produtiva do canavial. Para manter uma boa densidade de plantas na lavoura, no plantio da cana inteira, foi respeitado a proporção de 12 gemas viáveis por metro.

Normalmente, o plantio da cana-de-açúcar é realizado, em Barra de São Miguel, em outubro e a primeira colheita é feita em outubro do ano seguinte. Ao todo, a série de dados contempla os anos de 2013 a 2017. A data de florescimento e o número de panículas nas parcelas foram coletadas durante todos os anos do experimento, diariamente no período de florescimento (Gráfico 4).

Como essas parcelas tinham 5 metros e considerando a densidade de 12 gemas por metro, o *stand* seria de 60 colmos por parcela. Como o número de parcelas eram diferentes entre variedades diferentes no mesmo ano e entre as mesmas variedades em anos diferentes (Tabela 5); o florescimento foi somado, para as mesmas variedades no mesmo ano, e divididas pelo número de parcelas da referida variedade.

Dessa forma, o florescimento foi obtido por variedade por parcela por ano. Em seguida, foram feitas médias desses valores entre variedades diferentes, quando observados na mesma data, obtendo-se o número de médio de inflorescências. Esses valores foram divididos por 5, a fim de se obter o número médio de inflorescências por metro. Por fim, esses valores foram transformados em porcentagem e depois transformados em variável binária, em que:

$$Y = \begin{cases} 0 & \text{– houve florescimento de menos de 1\% das plantas em campo} \\ 1 & \text{– houve florescimento de 1\% ou mais das plantas em campo} \end{cases}$$

Estes valores binários foram usados como regressando (variável dependente) nas modelagens, usando a regressão logística. A porcentagem de 1% foi usada, pois, a aplicação de maturadores pode ser feita mesmo após ocorrido o florescimento de algumas plantas, desde que aplicados o mais cedo possível (COLETI et al., 1986). Os maturadores são substâncias que inibem o desenvolvimento das gemas florais (LEONEZI et al., 2016). Além disso, não foi identificado na literatura um nível de dano econômico exato. Esta informação identificaria o número de inflorescências limite, em que a tomada de decisão, ou seja, a aplicação de maturadores seria economicamente viável.

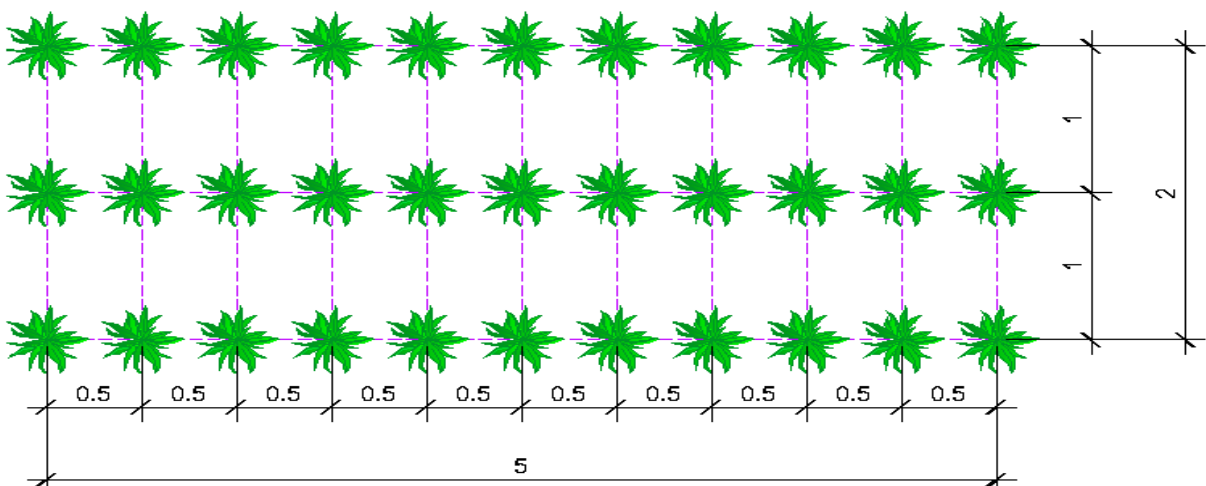


Figura 2 - Espaçamento inicial do plantio de cana-de-açúcar na área experimental. Fonte: Autor.

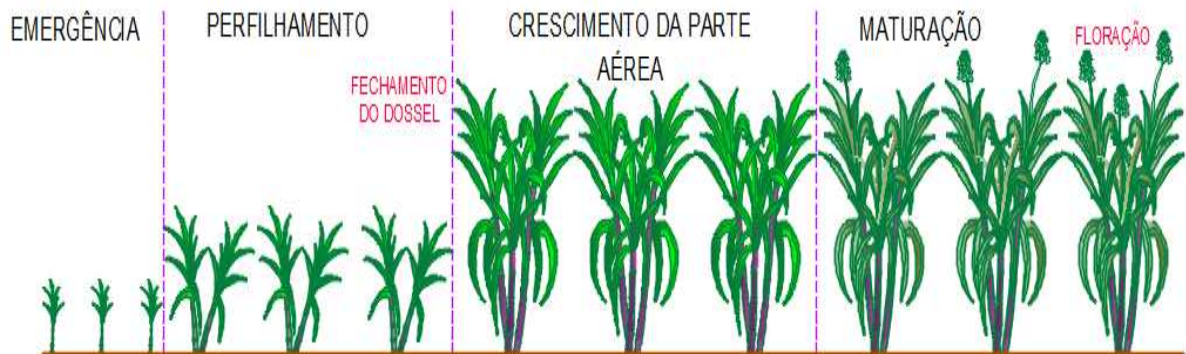
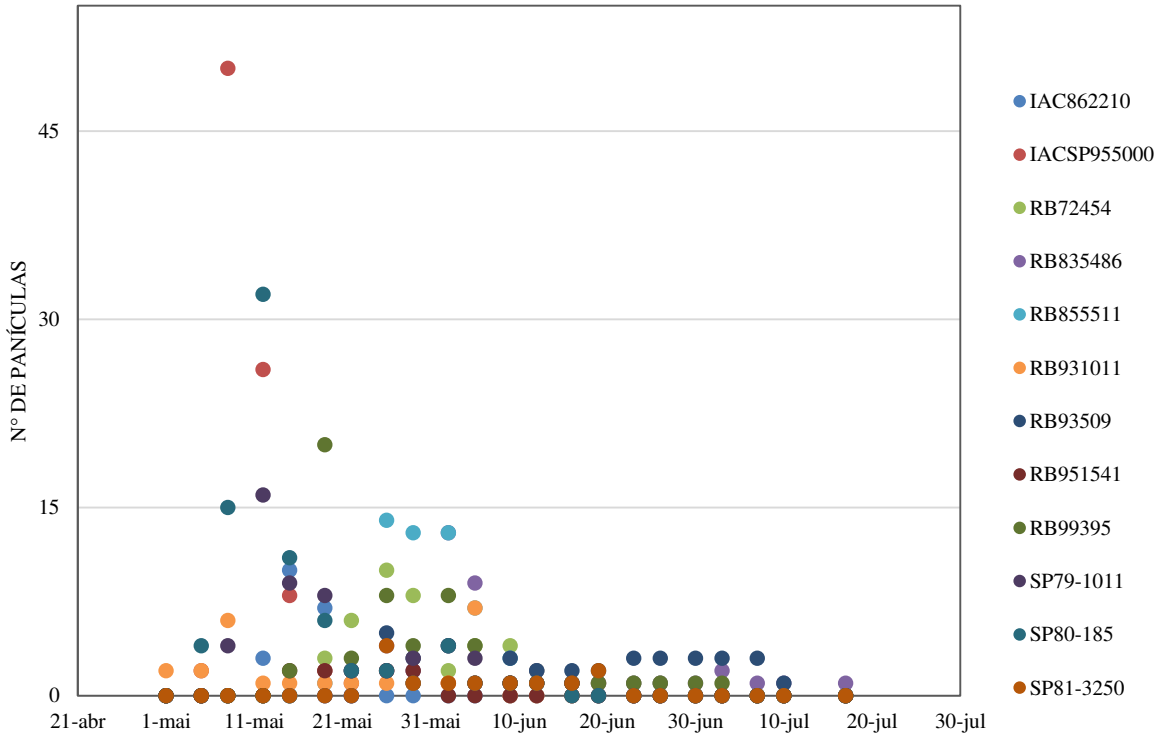
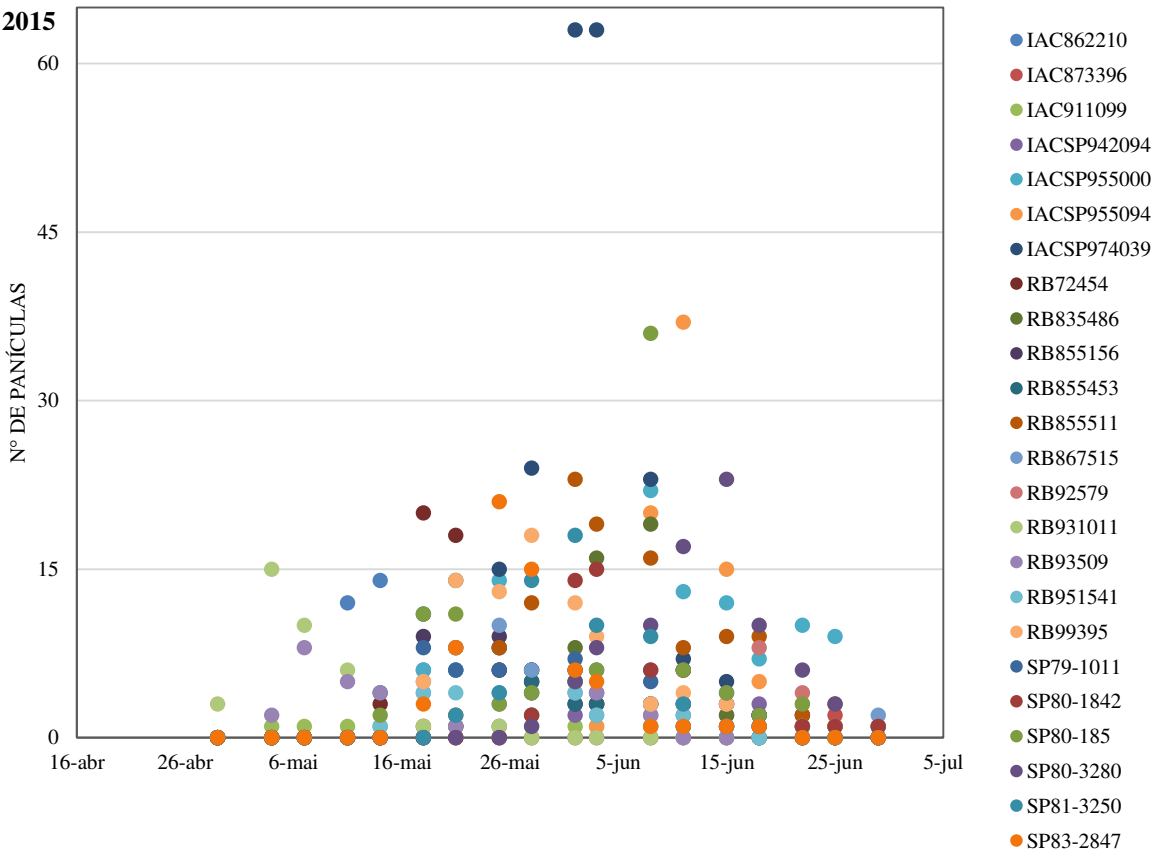


Figura 3 - Fenologia da Cana-de-açúcar. Fonte: Autor com base em Marin et al., 2009.

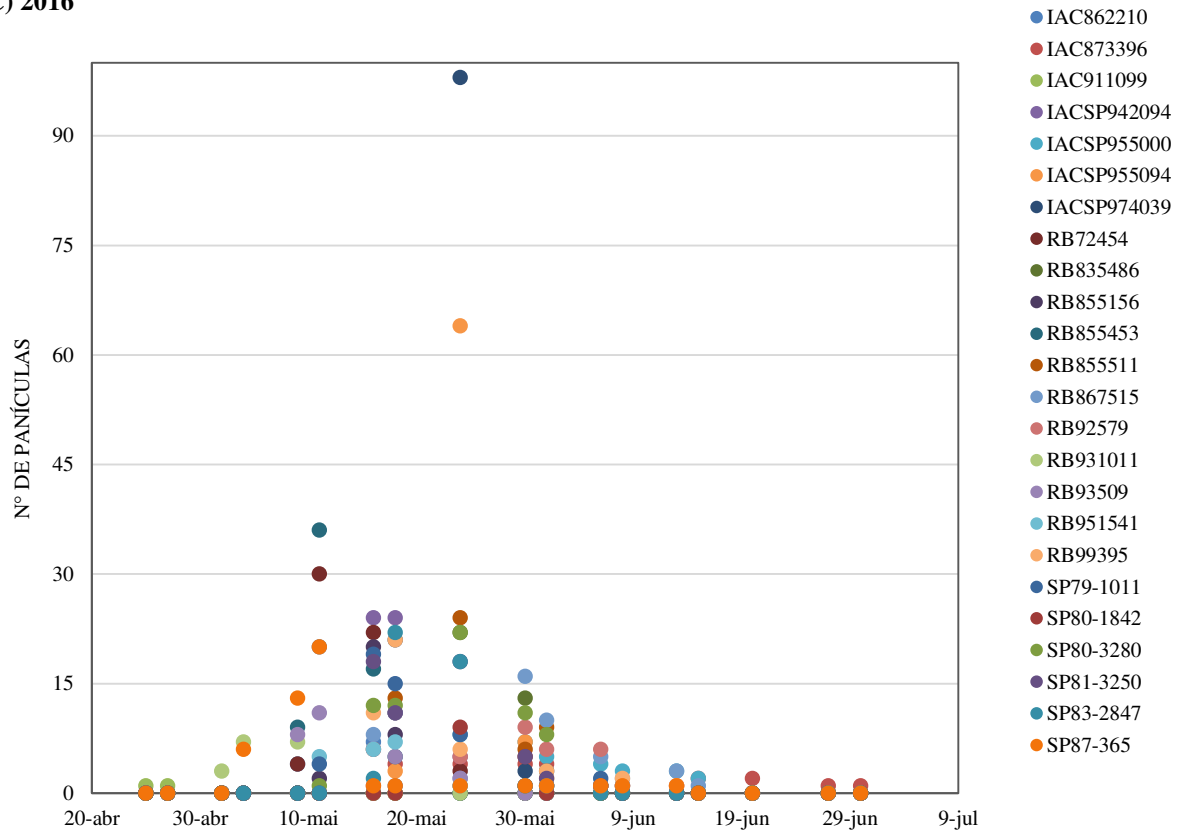
A) 2014



B) 2015



C) 2016



D) 2017

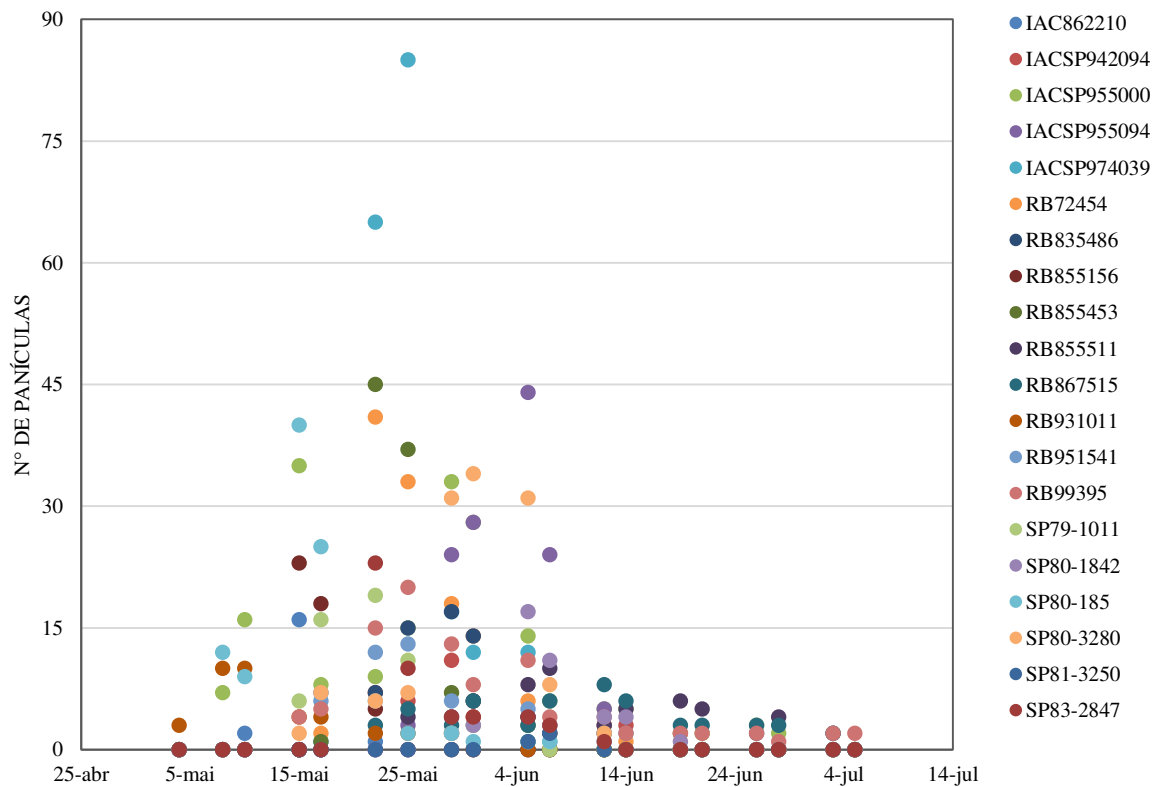


Gráfico 4 - Número de inflorescências observadas em campo para cada uma das variedades de cana-de-açúcar do experimento nos anos de: A) 2014, B) 2015, C) 2016 e D) 2017. Fonte: Autor.

3.4 Análise dos dados

3.4.1 Construção da curva de retenção e obtenção da CAD

De modo geral, a construção de uma curva característica de retenção de água envolve o conhecimento empírico da relação, numa amostra de solo, entre a tensão da água (potencial mátrico) e o conteúdo de água (NEBEL, 2009). Entretanto, nesse trabalho, a curva de retenção não foi obtida por métodos tradicionais como: a câmara de pressão de Richards, método da centrífuga, tensiometria e funil de placa porosa.

Foi utilizado o método proposto por Maia (2016), que reuniu, para Latossolos, dados da literatura com critério de seleção estatístico, a fim de obter os parâmetros da equação de Van Genuchten, que é um modelo empírico ou semi-empírico de curva de retenção de água no solo (Equação 7). Os valores ideais dos parâmetros, unidades de medida e significado podem ser observados na Tabela 6, e a curva característica de retenção para Latossolo Amarelo está representada no Gráfico 5. A partir desses parâmetros e da Equação 7, é possível calcular a CAD.

$$\Theta = \Theta_r + \frac{(\Theta_r - \Theta_s)}{[1 + (\alpha\psi)^p]^{1/q}} \quad (7)$$

PARÂMETRO	SIGNIFICADO	UNIDADE DE MEDIDA	VALOR IDEAL
θ	Conteúdo volumétrico de água	$m^3.m^{-3}$	
θ_s	Conteúdo de saturação de água no solo	$m^3.m^{-3}$	0,3722
θ_r	Conteúdo residual de água no solo	$m^3.m^{-3}$	0.1214

ψ	Valor absoluto do potencial mátrico	KPa	Variável na obtenção de diferentes pontos da curva de retenção
α	Inverso do valor de entrada de ar	KPa ⁻¹	0,0176
p	Parâmetro empírico de forma da curva.	-	1,6425
q	Parâmetro empírico de forma da curva.	-	0,3885

Tabela 6 - Parâmetros da equação de Van Genuchten para construção da curva de retenção. Fonte: Autor com base em Maia (2016).

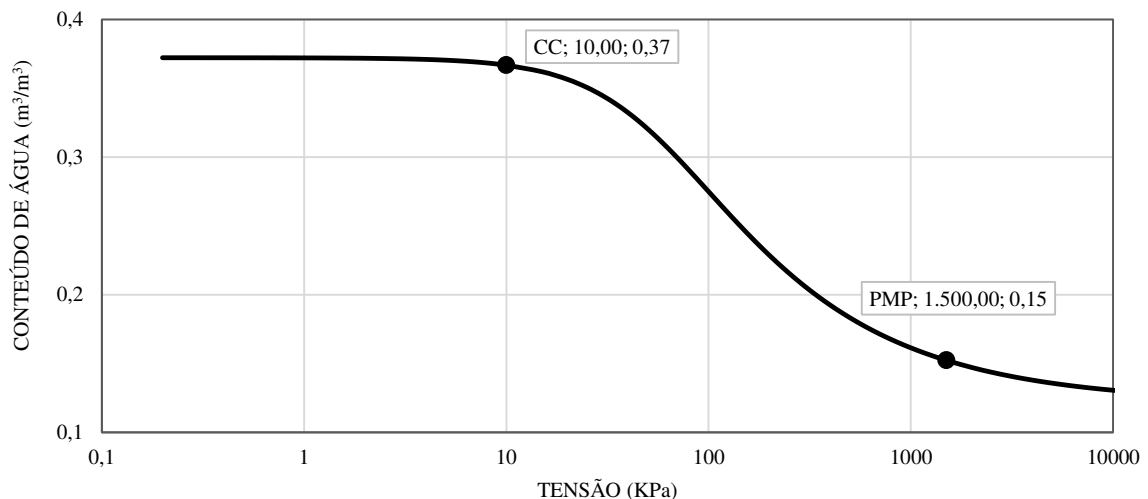


Gráfico 5 - Curva de retenção característica para Latossolo Amarelo, os pontos vermelhos evidenciam a capacidade de campo (CC) e ponto de murcha permanente (PMP). Fonte: Autor com base em Maia (2016).

A capacidade de água disponível (CAD), pode ser definida como a quantidade de água armazenada no solo, que está compreendida entre a capacidade de campo (CC) e o ponto de murcha permanente (PMP) (Equação 8). A CC, por sua vez, corresponde às condições após uma chuva ou irrigação. Nessas condições, a drenagem é considerada desprezível. Além disso, o teor de água é considerado constante para uma

dada profundidade. Já o PMP, se refere ao conteúdo de água crítico que as raízes não são capazes de absorver, uma vez que, essa água se encontra fortemente ligada à matriz do solo. Para quebrar essa ligação, seria necessário grande despendimento de energia pela planta, o que torna inviável (MAIA, 2016).

Para o cálculo da CAD, assumimos os conteúdos de água nas tensões de 10 e 1.500 kPa. A partir, dessas tensões e da equação de Van Genuchten, foi possível encontrar os valores de conteúdo de água na CC ($0.367 \text{ m}^3/\text{m}^3$), correspondente a 10 KPa, e no PMP ($0.152 \text{ m}^3/\text{m}^3$), correspondente a 1500 KPa (Gráfico 5). Com esses valores e a Equação 8, chegou-se a uma CAD de $0,215 \text{ m}^3/\text{m}^3$.

$$\text{CAD} = \text{CC} - \text{PMP} \quad (8)$$

3.4.2 Cálculo da evapotranspiração

Neste trabalho, a evapotranspiração foi calculada inicialmente pelo método de Penman-Monteith, que é o método parametrizado pelo guia 56 (*Irrigation and drainage*) da FAO – 56 (Food and Agriculture Organization). Esse método se baseia em termos energéticos e aerodinâmico, que, por sua vez, são controlados pelas resistências ao transporte de vapor da superfície para a atmosfera (Equação 9). Por essa razão, esse método apresenta bom desempenho em diferentes regiões, sob condições variadas (ALLEN et al., 1998).

Apesar disso, essa metodologia de cálculo requer muitos dados de entrada. No presente trabalho, variáveis como: temperatura, radiação, velocidade do vento e umidade relativa; puderam ser preenchidas com alguns bancos de dados (Tabela 2), após validação. Outras variáveis, como a irradiância solar diária no topo da atmosfera (R_0), pôde ser calculada através de relações astronômicas Terra-sol (itens 1 e 2 do Anexo 1). Entretanto, o preenchimento ou cálculo do período de brilho solar efetivo (η) não foi possível. Isso, devido à falta de bancos de dados que contenham esses valores. A estação meteorológica do INMET em Maceió (AL) contém algumas

observações dessa variável, porém com excessivas falhas, que não puderam ser preenchidas.

Dessa forma, a falta desses valores impediu os cálculos de irradiância global (R_g), que impediram os cálculos de balanço de ondas curtas e longas (BOC e BOL). Com isso, não foi feito o cálculo do saldo de radiação (R_n) (itens 3,4 e 5 do Anexo 1).

Entretanto, estes impedimentos foram contornados pelo uso do método de Hargreaves-Samani (Equações 10 e 11), como forma alternativa de cálculo de E_{To} , já que esse método é baseado em temperatura e relações astronômicas Terra-sol. Os resultados desse método mais o banco de dados de Xavier et al. (2016) foram validados por meio das estatísticas da Tabela 4 para posterior preenchimento da série de dados, calculada pelo método de Penman-Monteith.

a) Penman-Monteith - Padrão - FAO 1998

$$E_{To} = \frac{0,408 \Delta (R_n - G) + \gamma \frac{900}{T_{med} + 273} U_2 (e_s - e)}{\Delta + \gamma (1 + 0,34 U_2)} \quad (9)$$

b) Hargreaves & Samani - FAO 1998

$$E_{To} = 0,0023 \cdot Q_0 \cdot (T_{med} + 17,8) \cdot \sqrt{T_{max} - T_{min}} \quad (10)$$

$$Q_0 = R_0 / 2,45 \quad (11)$$

PARÂMETRO	SIGNIFICADO	UNIDADE DE MEDIDA
<i>E_{To}</i>	Evapotranspiração de referência	mm dia ⁻¹
<i>R_n</i>	Saldo de radiação à superfície	MJ m ⁻² dia ⁻¹
<i>G</i>	Fluxo de calor no solo	MJ m ⁻² dia ⁻¹
<i>T_{med}</i>	Temperatura média diária do ar	°C
<i>U₂</i>	Velocidade do vento a 2 m de altura, período de 24 horas	m s ⁻¹
<i>(e_s - e)</i>	Déficit de pressão de vapor	kPa
Δ	Declividade da curva de pressão de vapor de saturação	kPa °C ⁻¹

γ	Constante psicrométrica	kPa °C ⁻¹
<i>R_o</i>	Irradiância solar no topo da atmosfera	MJ/m ² .dia

Tabela 7 - Parâmetros das equações de Penman-Monteith e Hargreaves-Samani para obtenção da evapotranspiração de referência. Fonte: Autor.

3.4.3 Cálculo do balanço hídrico sequencial diário

Uma vez calculada a evapotranspiração potencial (ET_o), o balanço hídrico sequencial diário também pôde ser estimado. Isso porque, a ET_o é um importante componente do balanço hídrico. O balanço hídrico sequencial, no contexto deste trabalho, foi a contabilização diária de entradas e saídas de água. Essas contabilizações foram feitas pelo método de Thornthwaite-Mather, para as coordenadas da estação de cruzamento de cana-de-açúcar, em Barra de São Miguel (9,81°S de latitude, 35,95°W de longitude).

A sequência usada no cálculo desse balanço hídrico foram os dias dos anos em análise, ou seja, de 2014 a 2017. Uma grande parte desses dias não estavam compreendidos no período de indução floral da cana-de-açúcar, ou seja, entre os meses de abril a julho. Isso ocorre porque o balanço hídrico sequencial diário não é calculado a partir de um ano cíclico (normal). Dessa forma, o início dos cálculos de balanço deve acontecer, preferencialmente, quando houver uma sequência de períodos com chuva menos evapotranspiração maiores que zero [(P-ET_o) > 0]. Essa sequência deve ser grande o suficiente para garantir que armazenamento de água no solo se aproxime da capacidade de armazenamento (ARM = CAD). Assim, como desejou-se determinar o balanço hídrico sequencial diário a partir de março de 2014, foi conveniente iniciar os cálculos em 21 de fevereiro de 2014, já que nessa data houve uma sequência de [(P-ET_o) > 0] (SENTELHAS & ANGELOCCI, 2012).

Já a data final foi 1 de agosto de 2017. Essa data foi escolhida, pois, foi o período compreendido no banco de dados gradeados de Xavier e colaboradores (2016). Esperou-se com esse balanço hídrico sequencial conhecer os perfis diários de disponibilidade hídrica do solo, deficiência e excedente hídricos e, com isso, identificar a variabilidade dessas variáveis. No Anexo 2, é apresentado o roteiro de cálculo do balanço hídrico sequencial diário (SENTELHAS & ANGELOCCI, 2012).

3.4.4 Teste de normalidade

Antes da escolha do modelo de regressão, a normalidade dos regressores (variáveis independentes) foi verificada. A justificativa disso é que o atendimento à pressuposição de normalidade dos dados é condição necessária para alguns modelos de regressão, como o probit ou normit. Além disso, a normalidade é condição necessária para análise dos resíduos de equações ajustadas. Para tal verificação, existem vários testes estatísticos, nesse trabalho foram usados os testes de Shapiro-Wilk e Lilliefors (SHAPIRO & WILK, 1965; DALLAL & WILKINSON, 1986).

O teste de Shapiro-Wilk (1965), baseia-se na estatística W dada por:

$$W = \frac{b^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (12)$$

Em que:

x_i são os valores da amostra ordenados de forma crescente, ou seja, $x_{(1)}$ é o menor.

b é determinado da seguinte forma:

$$b = \begin{cases} \sum_{i=1}^{n/2} a_{n-i+1} \cdot (x_{n-i+1} - x_{(i)}) & \text{se } n \text{ é par} \\ \sum_{i=1}^{(n+1)/2} a_{n-i+1} \cdot (x_{n-i+1} - x_{(i)}) & \text{se } n \text{ é ímpar} \end{cases} \quad (13)$$

a_{n-i+1} são constantes tabeladas, seus valores são gerados a partir de médias, variâncias e covariâncias de uma amostra de tamanho n .

Já o teste de Lilliefors é uma modificação do teste de Kolmogorov-Smirnov. Essa modificação consiste em estimativas de média e variância dos dados, já que o teste de Kolmogorov-Smirnov é realizado para uma função de distribuição específica, com média e variância conhecidas. O teste de Lilliefors baseia-se na estatística D dada por:

$$D_n = \text{Max} | F(x_i) - F_n(x_i) | \quad (14)$$

Em que:

$[F(x_i)]$ é a distribuição de probabilidades acumuladas observadas.

$[F_n(x_i)]$ é a distribuição acumulada teórica (distribuição de Z).

Esse teste consiste em verificar se a distância máxima entre a distribuição de probabilidades acumuladas observadas padronizadas $[F(x_i)]$ e a distribuição acumulada teórica (distribuição de Z) $[F_n(x_i)]$; é estatisticamente significativa ou não, a um dado nível de probabilidade. Esse nível de probabilidade foi adotado em 5%, neste trabalho.

A interpretação dos testes é semelhante, já que ambos testam as seguintes hipóteses:

$$\begin{cases} H_0: \text{o regressor segue uma distribuição normal} \\ H_1: \text{o regressor não segue uma distribuição normal} \end{cases}$$

Para determinar se o regressor em análise segue ou não uma distribuição normal, o p-valor foi comparado com o nível de significância de 5% ($\alpha=0,05$). Esse nível de significância indica um risco de 5% de concluir que o regressor segue uma distribuição normal, quando, na verdade, realmente a segue. Resumindo:

- Se p-valor $\leq \alpha$: o regressor segue uma distribuição normal (Rejeita-se a H_0).
- Se p-valor $> \alpha$: não é possível concluir que o regressor não segue uma distribuição normal (não deve rejeitar H_0).

3.4.5. Modelagem

Neste trabalho, o florescimento, obtido conforme a metodologia do item 3.3 da Metodologia, foi a variável dependente ou regressando. Essa variável foi predita por um modelo qualitativo, a chamada Regressão Logística (Logit). Nesse modelo, as variáveis regressoras ou independentes foram: a água armazenada no solo (ARM), obtida através do balanço hídrico sequencial diário; volume de chuvas, obtido por observações na estação experimental de Barra de São Miguel; e evapotranspiração, calculada pelo método de Penman-Monteith e Hargreaves-Samani.

3.4.5.1 Regressores

Cada um desses regressores foram tratados para 4 situações, totalizando 16 modelos. A primeira situação corresponde ao valor instantâneo da variável 30 dias antes do evento de florescimento observado, por exemplo, para um evento de florescimento no dia 8 de maio, a ARM, chuva ou evapotranspiração correspondente seria a registrada no dia 8 de abril.

A segunda seria a soma das variáveis durante o período de indução floral, contados a partir de 15 de março até 30 dias antes do evento de florescimento observado. Por exemplo, para um evento de florescimento no dia 8 de maio, a ARM, chuva ou evapotranspiração correspondente seria a soma dos volumes do dia 15 de março até o dia 8 de abril.

As duas últimas situações se referem à média da primeira e segunda situação, por 10 dias consecutivos, agrupados a partir do dia 15 de março. No Gráfico 6 exemplificam-se essas situações, mostrando que, a partir do dia 15 de março, 10 dias consecutivos têm mesmo valor de volume de chuvas, que correspondem à: A) média do volume de chuvas nesses 10 dias (Chuva média – C_m) e B) média do volume de chuvas acumulado por 10 dias (Chuva soma média – C_{sm}). Na Tabela 8, estão compilados todos os regressores usados na obtenção dos modelos.

Essas formas de tratamento dos regressores foram escolhidas a fim de entender se as variáveis afetam o florescimento de modo instantâneo ou acumulativo. Além de elucidar se a data de 15 de março é compatível com o início da indução de florescimento relatado na literatura (ARALDI et al., 2010).

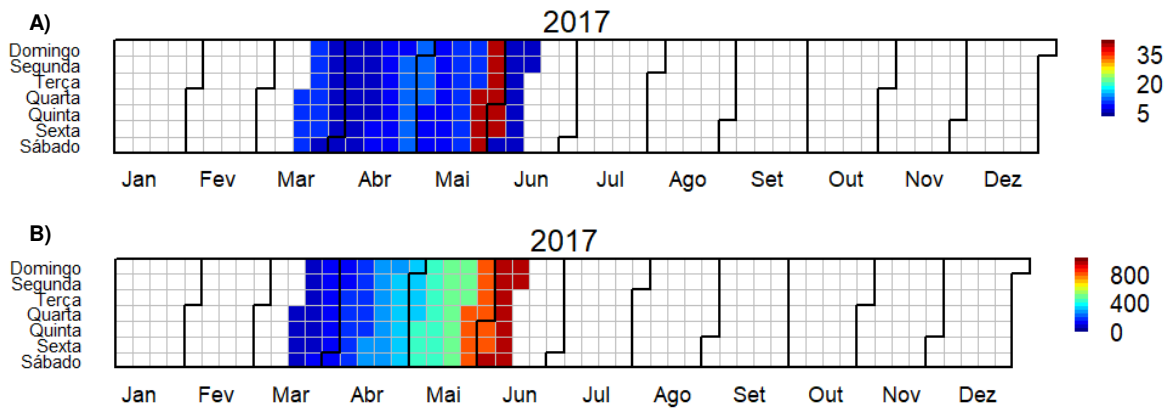


Gráfico 6 - Gráfico de calendário evidenciando: A) média do volume de chuvas, em milímetros, de 10 dias consecutivos (Cm) e B) média do volume acumulado chuvas, em milímetros, de 10 dias consecutivos (Csm). Fonte: Autor.

SIGLA	REGRESSOR	UNIDADE
ARM	Quantidade de água armazenada no solo	mm.dia ⁻¹
ARM_s	Quantidade somada de água armazenada no solo em 10 dias consecutivos	mm.10 dias ⁻¹
ARM_m	Média da quantidade de água armazenada no solo em 10 dias consecutivos	mm.dia ⁻¹
ARM_{sm}	Média da quantidade somada de água armazenada no solo em 10 dias consecutivos	mm.10 dias ⁻¹
C	Chuva diária	mm.dia ⁻¹
C_s	Soma do volume de chuva em 10 dias consecutivos	mm.10 dias ⁻¹
C_m	Média da chuva diária	mm.dia ⁻¹
C_{sm}	Média da soma do volume chuva em 10 dias consecutivos	mm.10 dias ⁻¹
ET_o	Evapotranspiração diária por Penman-Monteith	mm.dia ⁻¹
ET_{os}	Soma do volume de evapotranspiração por Penman-Monteith em 10 dias consecutivos	mm.10 dias ⁻¹
ET_{om}	Média da evapotranspiração diária por Penman-Monteith	mm.dia ⁻¹
ET_{osm}	Média da soma do volume evapotranspiração, calculada pelo método de Penman-Monteith, em 10 dias consecutivos	mm.10 dias ⁻¹
ET_{oHS}	Evapotranspiração diária por Hargreaves-Samani	mm.dia ⁻¹
ET_{oHSs}	Soma do volume de evapotranspiração por Hargreaves-Samani em 10 dias consecutivos	mm.10 dias ⁻¹
ET_{oHSm}	Média da evapotranspiração diária por Hargreaves-Samani	mm.dia ⁻¹
ET_{oHSsm}	Média da soma do volume evapotranspiração, calculada pelo método de Hargreaves-Samani, em 10 dias consecutivos	mm.10 dias ⁻¹

Tabela 8 – Regressores usados na modelagem do florescimento da cana-de-açúcar. Fonte: Autor.

3.4.5.2 Modelo de Regressão Logística (Logit)

Nesse modelo, o florescimento foi simulado por meio de uma variável binária.

Em que:

$$Y = \begin{cases} 0 & \text{– não houve florescimento de 1% das plantas em campo} \\ 1 & \text{– houve florescimento de de 1% das plantas em campo} \end{cases}$$

Uma vez que a lâmina de água no solo, chuva ou evapotranspiração podem ou não levar 1% das plantas em campo florescerem, o florescimento é do tipo sim ou não. Em outras palavras, o regressando é uma variável binária, ou dicotômica, ou *dummy*. Como esse modelo não é afetado pela normalidade dos dados de florescimento, já que o termo de erro segue uma distribuição binomial, o modelo qualitativo do tipo logit é o mais adequado mesmo que seja verificada a não normalidade dos regressores pelos métodos descritos no item 3.4.4 da Metodologia (GUJARATI; PORTER, 2011).

O modelo logit pode ser escrito como:

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = \beta_1 + \beta_2 X_i + u_i \quad (15)$$

Em que:

L: é o logaritmo da razão de chances, chamado de logit.

$P_i/(1 - P_i)$: é a razão entre as chances de florescimento ou não de 1% das plantas em campo.

L positivo: significa que, quando a lâmina de água no solo aumenta (X_i), as chances de 1% das plantas florescerem ser igual a 1, aumentam.

L for negativo: as chances de 1% das plantas florescerem ser igual a 1, diminuem; à medida que o valor de X aumenta.

β_1 : intercepto, representa o valor do logaritmo das chances quando x é igual a zero.

β_2 : coeficiente angular, mede a variação em L para uma unidade de variação em X.

u_i : representa o erro estocástico.

A fim de facilitar a interpretação dos coeficientes, dos modelos ajustados, fora calculada a razão das chances (OR - Odds Ratio) para a variável regressora. Esse cálculo consiste na exponenciação da variável de regressão. Isso porque no modelo de regressão logística os estimadores resultantes ficam na forma logarítmica, ou seja, o log das chances de florescer 1% das plantas. Outra questão, ligada a esses estimadores, é que são obtidos pelo método de máxima verossimilhança. Esta informação é útil na seleção dos melhores modelos.

Dentre os critérios para seleção dos modelos, os critérios de p-valor significativo ($<0,05$) e menor AIC foram usados. O Critério de Informação de Akaike (AIC) é baseado no máximo da função de verossimilhança (MFV) e busca selecionar, dentre um grupo de modelos avaliados, o modelo que mais se aproxima do modelo real, que é aquele que descreve os dados e é desconhecido. O modelo com menor valor de AIC é considerado o modelo de melhor ajuste (BOZDANGAN, 1987). O cálculo de AIC é dado por:

$$\text{AIC} = -2L + 2K \text{ (16)}$$

Em que:

L: representa o MFV do modelo com os parâmetros “ θ ”.

K: o número de parâmetros

Após selecionados os modelos, em que os regressores são lâmina de água no solo, chuva ou evapotranspiração; esses modelos terão seus desempenhos avaliados quanto a capacidade de predição.

3.4.5.3 Avaliações de desempenho do modelo Logit

A qualidade do ajuste dos modelos, com p-valor significativo e menores AIC, foram avaliadas por diferentes métodos: matriz de regressão, curva ROC, “pseudo” R^2 , erro médio quadrático (RMSE) e erro médio absoluto (MAE). Estes métodos foram calculados utilizando valores preditos pelos modelos ajustados. A seguir, cada um desses métodos é descrito, com exceção de RMSE e MAE, que já foram descritos no item 3.2 da Metodologia.

3.4.5.3.1 Matriz de Confusão

Uma das formas de verificação da qualidade do ajuste dos modelos, com p-valor significativo, foi a chamada matriz de confusão. Essa matriz consiste na classificação das probabilidades de florescimento, de acordo com os “erros” e os “acertos” de predição do modelo ajustado (HOSMER & LEMESCHOW, 2000).

Para verificar se houve acerto ou erro na predição, primeiramente é preciso estipular uma regra de predição ou ponto de corte. Para este trabalho, o ponto de corte adotado foi o de 0,5. Probabilidades estimadas em 0,5 ou mais são classificadas como 1 ($\hat{Y}=1$); e abaixo de 0,5 são classificadas como 0 ($\hat{Y}=0$) (HOSMER & LEMESCHOW, 2000). Estipulado o ponto de corte, avaliou-se o poder de discriminação do modelo, pelo seu desempenho em classificar os “houve 1% de florescimento” dos “não houve 1% de florescimento”. A partir disso, comparando com os valores reais de florescimento ($Y=1$ e $Y=0$), os acertos e erros são compilados numa matriz, que contém valores das observações de Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN) (Tabela 9).

	Valor Observado	
	Y=1	Y=0
Valor Estimado	$\hat{Y}=1$	$\hat{Y}=0$
	VP	FP
	FN	VN

Tabela 9 - Modelo teórico de uma Matriz de confusão. Fonte: Adaptado de Fawcett (2006).

3.4.5.3.2 Curva ROC

Os valores de VP, FP, FN E VN, usados para construir a matriz de confusão, também foram usados na confecção da curva ROC (*Receiver Operating Characteristic Curve*) e obtenção da acurácia - AC (Equação 17). Essa curva foi usada para visualizar, organizar e classificar os modelos propostos com base na *performance* preditiva. Isso porque a curva ROC mede a capacidade de predição do modelo proposto, através das medidas de sensibilidade e especificidade (Equações 18 e 19).

$$AC = \frac{VP+VN}{VP+VN+FP+FN} \quad (17)$$

$$\text{SENS} = \frac{VP}{VP+FN} \quad (18)$$

$$\text{ESP} = \frac{VN}{VN+FP} \quad (19)$$

A sensibilidade é a probabilidade de o teste apontar que houve 1% de florescimento, dado que, de fato, houve. Já a especificidade é a probabilidade de um teste apontar que não houve 1% de florescimento, dado que houve.

Outra medida importante, encontrada na curva ROC, é a AUC (*Area Under the ROC Curve*). Essa medida representa a probabilidade de que o classificador efetue previsões aleatórias positivas melhores que as negativas. Os valores de AUC variam entre 0 e 1. Quanto mais próximo de 1, melhor. Além disso, valores de AUC devem estar acima de 0,5 para serem considerados realísticos (FAWCETT, 2006).

Essas previsões de verdadeiro positivo e falso positivo são plotadas em uma linha de limites crescentes para criar a curva. Como a curva é constituída de um conjunto de limites, então, a fim de comparar modelos diferentes, um limite ideal é necessário. Esse limite, foi plotado na curva ROC no ponto que está mais próximo do canto superior esquerdo do gráfico (FAWCETT, 2006).

3.4.5.3.3 “Pseudo” R²

Como os estimadores dos modelos são obtidos pelo método de máxima verossimilhança, significa que o R², comumente usado para medir a qualidade do ajuste de modelos, deve ser substituído pelo “pseudo” R² (NAGELKERKE, 1991; HAIR, 2009).

O “pseudo” R² é utilizado para comparar modelos. Existem várias metodologias de cálculo do “pseudo” R², como a Cox e Snell's e Nagelkerke. A metodologia de Cox e Snell's é baseada no log do valor de verossimilhança de um modelo (modelo completo) em comparação com o log do valor de verossimilhança para um modelo base (modelo nulo). Já a metodologia de Nagelkerke é uma versão ajustada de Cox e Snell's; e foi a metodologia adotada neste trabalho (NAGELKERKE, 1991; HAIR, 2009). Segue abaixo o cálculo do “pseudo” R² para essa metodologia:

$$R^2_{\text{Nagelkerke}} = \frac{1 - \left(\frac{L_0}{L_1}\right)^{2/n}}{1 - L_0^{2/n}} \quad (20)$$

Em que:

L_0 : representa o logaritmo do valor de verossimilhança do modelo nulo.

L_1 : representa o logaritmo do valor de verossimilhança do modelo completo.

n : representa o tamanho da amostra

4. RESULTADOS E DISCUSSÕES

4.1 Validações de dados meteorológicos

Pelos resultados da validação dos dados de Coruripe, verificamos que, das variáveis analisadas, as temperaturas mínima (Tmin), máxima (Tmax) e média (Tmed) foram as que tiveram as melhores *performances*, com erros médios (RMSE) que variam de: 2,64°C a 4,92°C (para o ERA5), 2,88°C a 3,49°C (para o NCEP/NCAR) e 1,95°C a 4,77°C (para o Xavier); ou em porcentagem (PMRE): 6,14% a 15,33% (para o ERA5), 6,95% a 9,09% (para o NCEP/NCAR) e 4,91% a 15,36% (para o Xavier). De modo geral, os bancos de dados subestimaram essas temperaturas. Essa subestimação variou de -0,65°C, para temperatura média com o Xavier, a -4,31°C para temperatura mínima, também com Xavier. Verificou-se superestimação apenas na temperatura máxima, usando o ERA5 (+1,96°C) e Xavier (+3,02°C). Superestimações foram verificadas em todos bancos de dados para radiação (Rad) e, na maioria, para velocidade do vento (u_2) (Tabela 10).

VARIÁVEL	BANCO DE DADOS	r	R ²	RMSE	MAE	VIÉS	PMRE (%)
Rad (MW.m ⁻²)	ERA5	0.80	0.64	48.15	38.06	38.06	24.40
	NCEP/NCAR	0.62	0.38	77.36	66.25	66.25	41.13
	Xavier	0.81	0.66	27.04	5.42	5.42	13.17
Tmax (°C)	ERA5	0.49	0.24	3.22	1.96	1.96	10.59
	NCEP/NCAR	0.57	0.33	2.88	1.63	-1.63	6.95
	Xavier	0.72	0.51	3.60	3.02	3.02	11.91
Tmed (°C)	ERA5	0.57	0.33	2.64	1.29	-1.29	6.14
	NCEP/NCAR	0.54	0.29	3.49	2.53	-2.53	9.09
	Xavier	0.78	0.61	1.95	0.65	-0.65	4.91
Tmin (°C)	ERA5	0.48	0.24	4.92	4.28	-4.28	15.33
	NCEP/NCAR	0.47	0.22	3.38	2.32	-2.32	8.94
	Xavier	0.73	0.54	4.77	4.31	-4.31	15.36
u_2 (m.s ⁻¹)	ERA5	0.78	0.60	1.35	1.26	1.26	53.74
	NCEP/NCAR	0.25	0.06	3.42	3.12	3.12	136.03
	Xavier	0.55	0.30	0.78	0.62	-0.62	25.10
UR (%)	ERA5	0.70	0.48	4.85	0.35	-0.35	5.00
	NCEP/NCAR	0.49	0.24	10.67	8.94	8.94	12.37
	Xavier	0.85	0.72	3.53	1.16	-1.16	3.63

Tabela 10 - Resultados por estatística de validação das variáveis meteorológicas da estação do INMET em Coruripe (AL). As unidades de medida de RMSE, MAE e VIÉS são as mesmas das variáveis meteorológicas em teste. Fonte: Autor.

Essas superestimções variaram de +5,42 (Xavier) a +66,25 MW.m⁻² (NCEP/NCAR), para radiação, e de +1,26 (ERA5) a +3,12 m.s⁻¹ (NCEP/NCAR) para velocidade do vento. Apenas o Xavier subestimou a velocidade do vento em -0,62 m.s⁻¹. Por outro lado, a umidade relativa foi subestimada pelo ERA5 (-0,35%) e Xavier (-1,16%) e superestimada pelo NCEP/NCAR (8,94%). Dessas variáveis, a umidade relativa apresentou os melhores resultados, seguida da radiação e, por último, a velocidade do vento. Os erros (RMSE) associados à mensuração da umidade relativa variou de 3,53 (Xavier) a 10,67% (NCEP/NCAR) ou em porcentagem (PMRE) de 3,63% a 12,37%. Para radiação, esses erros variaram de 27,04 (Xavier) a 77,36 MW.m⁻² (NCEP/NCAR) ou de 13,17 % (Xavier) a 41,13% (NCEP/NCAR). Já, para velocidade do vento, esses erros foram bem maiores, variando de 0,78 (Xavier) a 3,42 m.s⁻¹ (NCEP/NCAR) ou de 25,1% (Xavier) a 136,03% (NCEP/NCAR).

Nessa tabela também foi possível observar os valores de coeficientes de correlação (r), obtidos entre os dados observados (INMET - Coruripe) e os bancos de dados: ERA5 (0,48 a 0,80), NCEP/NCAR (0,25 a 0,62) e Xavier (0,55 a 0,85); além dos valores de R^2 , entre os dados observados e os bancos de dados: ERA5 (0,24 a 0,64), NCEP/NCAR (0,06 a 0,38) e Xavier (0,30 a 0,72). Nessa tabela, podemos observar que, para a maioria das variáveis, ERA5 e Xavier apresentaram os melhores valores das estatísticas r e R^2 , ou seja, os resultados se aproximaram de 1 (uma correlação perfeita).

Dessa forma, foi possível verificar que o ERA5 e Xavier possuem potenciais parecidos de preenchimento de falhas, das variáveis meteorológicas em estudo, da estação do INMET em Coruripe (AL). Para este trabalho, ERA5 foi usado para preencher as falhas dos dados de velocidade do vento e o Xavier foi usado para preencher radiação, temperatura (mínima, máxima e média) e umidade relativa.

Já, pelos resultados da validação dos dados de Barra de São Miguel, verificamos que, das variáveis analisadas, as temperaturas (máxima, média e mínima) e umidade relativa foram as que tiveram as melhores *performances*, com erros médios (RMSE) que variam de 0,60 (INMET Maceió) a 6,25°C (INMET Coruripe) e de 6 (NCEP/NCAR) a 8,65% (INMET Coruripe), respectivamente. De modo geral, os bancos de dados subestimaram as temperaturas máxima e média (-0,04 a -3,80°C) e

superestimaram a temperatura mínima (+0,03 a +5,51°C). Já a umidade relativa foi subestimada pela maioria dos bancos de dados (-5,64% a -7,15%), exceto pelo NCEP/NCAR, que superestimou em +1,91%. Superestimações foram verificadas na maioria dos bancos de dados para velocidade do vento (u_2) (Tabela 11).

VARIÁVEL	BANCO DE DADOS	r	R ²	RMSE	MAE	VIÉS	PMRE (%)
Rad (MW.m ⁻²)	INMET Coruripe	0.59	0.34	4.06	0.55	-0.55	18.15
	ERA5	0.65	0.42	4.69	2.65	2.65	22.43
	INMET Maceio	0.59	0.35	4.20	0.32	-0.32	18.59
	NCEP/NCAR	0.45	0.20	13.29	2.56	-2.56	59.40
	Xavier	0.82	0.68	2.84	1.11	-1.11	12.11
Tmax (°C)	INMET Coruripe	0.59	0.35	2.46	0.94	-0.94	7.23
	ERA5	0.71	0.51	1.53	0.55	-0.55	4.01
	INMET Maceio	0.80	0.63	3.92	3.80	-3.80	12.66
	NCEP/NCAR	0.70	0.48	3.60	3.42	-3.42	11.39
	Xavier	0.83	0.69	1.93	1.51	1.51	5.53
Tmed (°C)	INMET Coruripe	0.47	0.23	3.60	2.72	2.72	11.01
	ERA5	0.85	0.72	0.73	0.04	-0.04	2.33
	INMET Maceio	0.90	0.82	0.60	0.08	0.08	1.80
	NCEP/NCAR	0.70	0.49	1.08	0.50	-0.50	3.47
	Xavier	0.81	0.66	1.95	1.72	1.72	6.81
Tmin (°C)	INMET Coruripe	0.20	0.04	6.25	5.51	5.51	25.65
	ERA5	0.73	0.54	1.13	0.03	0.03	4.16
	INMET Maceio	0.68	0.46	3.34	3.10	3.10	14.46
	NCEP/NCAR	0.52	0.27	3.03	2.67	2.67	12.72
	Xavier	0.68	0.46	1.63	1.06	1.06	6.20
u_2 (m.s ⁻¹)	INMET Coruripe	0.55	0.30	0.81	0.16	0.16	42.45
	ERA5	0.64	0.42	1.28	1.05	1.05	73.41
	INMET Maceio	0.73	0.53	0.80	0.47	0.47	44.14
	NCEP/NCAR	0.41	0.17	3.73	3.48	3.48	201.58
	Xavier	0.67	0.45	0.74	0.21	-0.21	29.93
UR (%)	INMET Coruripe	0.65	0.42	8.65	7.15	-7.15	8.96
	ERA5	0.67	0.45	7.83	6.30	-6.30	7.92
	INMET Maceio	0.75	0.57	6.77	5.64	-5.64	7.12
	NCEP/NCAR	0.42	0.17	6.00	1.91	1.91	5.87
	Xavier	0.82	0.68	6.68	5.86	-5.86	7.22

Tabela 11 - Resultados por estatística de validação das variáveis meteorológicas da estação local de Barra de São Miguel (AL). As unidades de medida de RMSE, MAE e VIÉS são as mesmas das variáveis meteorológicas em teste. Fonte: Autor.

Essas superestimacões variaram de +0,16 (INMET Coruripe) a +3,48 m.s⁻¹ (NCEP/NCAR), com exceção do Xavier, que subestimou u_2 em -0,21 m.s⁻¹. Por outro lado, a radiação foi subestimada pela maioria dos bancos de dados, variando de -0,32 (INMET Maceió) a -2,56 MW.m⁻² (NCEP/NCAR). A exceção foi o ERA5, que superestimou Rad em 2,65 MW.m⁻². Os erros (RMSE) associados à mensuração da radiação variaram de 2,84 (Xavier) a 13,29 MW.m⁻² (NCEP/NCAR) ou de 12,11% (Xavier) a 59,40% (NCEP/NCAR). Já, para velocidade do vento, esses erros foram bem maiores, variando de 0,74 (Xavier) a 3,73 m.s⁻¹ (NCEP/NCAR) ou de 29,93% (Xavier) a 201,58% (NCEP/NCAR).

Observamos também os valores de coeficientes de correlação (r), obtidos entre os dados observados (estação de Barra de São Miguel) e os bancos de dados: ERA5 (0,64 a 0,85), INMET Coruripe (0,20 a 0,65), INMET Maceió (0,59 a 0,90), NCEP/NCAR (0,41 a 0,70) e Xavier (0,67 a 0,83); e os valores de R^2 , entre os dados observados e os bancos de dados. Sendo, ERA5 (0,42 a 0,72), INMET Coruripe (0,04 a 0,42), INMET Maceió (0,35 a 0,82), NCEP/NCAR (0,17 a 0,48) e Xavier (0,45 a 0,69). Nessas tabelas, podemos observar que INMET Maceió, ERA5 e Xavier apresentaram os melhores valores das estatísticas r e R^2 , ou seja, os resultados se aproximaram de 1 (uma correlação perfeita).

Dessa forma, foi possível verificar que INMET Maceió, ERA5 e Xavier possuem potenciais parecidos de preenchimento de falhas, das variáveis meteorológicas em estudo, da estação em Barra de São Miguel (AL). Para esse trabalho, INMET Maceió foi usado para preencher as falhas dos dados de velocidade do vento e temperatura média; o ERA5 foi usado para preencher as falhas dos dados de temperatura mínima e o Xavier foi usado para preencher radiação, temperatura máxima e umidade relativa.

Dos bancos de dados em grade (Xavier, NCEP/NCAR e ERA5), o Xavier obteve melhor desempenho para Coruripe (Tabela 10), e Barra de São Miguel (Tabela 11). Nesse banco, em sua versão 2, dados de radiação, velocidade do vento e umidade relativa foram interpolados a partir de várias fontes *in situ*, usando vários modelos como: IDW (*inverse distance weighting*), OPK (*ordinary point kriging*) e ADW (*angular distance weighting*). IDW e ADW mostraram os melhores resultados nessa versão. Já as temperaturas (Tmax e Tmin), na versão 2.1, foram interpoladas usando os modelos IDW e ADW, por apresentarem melhores resultados na versão 2, mas com a adição da informação de relevo. As grades do banco de dados de Xavier e

colaboradores foram geradas por meio desses dados, com resolução de 0,1 grau por 0,1 grau, ou seja, a qualidade desse banco pode ser explicada pela própria forma como esses dados foram mensurados (XAVIER et al., 2016). A superioridade da performance de dados observados pode ser usada para explicar o desempenho das estações do INMET em Coruripe e Maceió.

Essas estações foram escolhidas por estarem próximas de Barra de São Miguel. Pela proximidade, já era esperado que a estação de Maceió, que fica a 30,8 km de Barra de São Miguel, correlacionasse mais e melhor que a estação de Coruripe, que fica localizada a 62,6 km de Barra de São Miguel. Entretanto, os resultados de coeficiente de correlação (r) da estação de Coruripe ficaram muito abaixo dos outros bancos de dados. Isso porque, apesar da proximidade, raio inferior 100 km, que é considerado a distância ideal pela Organização Meteorológica Mundial (OMM) para representar bem uma região, outros fatores atuaram na redução desse desempenho. (BLAINSKI et al., 2012; MELLO et al., 2017; BIELENKI JUNIOR et al., 2018).

Um desses fatores pode ser as várias feições dos Tabuleiros Costeiros, discutidas no item 3.1 da metodologia, além da proximidade do oceano. Esses fatores fazem com que haja diferenças climáticas relevantes em áreas inferiores a um raio de 100 km (BLAINSKI et al., 2012; MELLO et al., 2017; BIELENKI JUNIOR et al., 2018). Outro fator, a quantidade excessiva de falhas, que necessitaram ser preenchidas para a estação de Coruripe, com banco de dados em grade (Gráfico 3), dando origem à uma série sintética. Essa série possui capacidade reduzida de correlação com os dados de Barra de São Miguel (Tabela 11). Embora, medidas de acurácia, como RMSE e viés, sejam menores que os do ERA5 e NCEP/NCAR (Tabela 9).

ERA5 e NCEP/NCAR, assim como Xavier, são dados de grade, porém são reanálises. Esse tipo de dado é baseado na associação de um modelo de previsão e um sistema de assimilação de dados. Esses dados são provenientes de diferentes fontes, tais como: radiossonda; observação de superfície por meio de equipamentos e sensores; micro-ondas; imagens de satélite; entre outros (CHEN et al. 2008; SASHA et al., 2010). O desempenho do ERA5 foi melhor que o do NCEP/NCAR, tanto em Coruripe, como em Barra de São Miguel. A qualidade do ERA5 em representar elementos meteorológicos em várias partes do globo foi discutida por vários autores.

Zuluaga e colaboradores (2021), trabalhando com radiação, chegaram à conclusão que o ERA5 é uma excelente opção para investigar DSWR (radiação de onda curta descendente) em todo o Brasil. Esses autores identificaram que a região

Nordeste, onde fica Barra de São Miguel, foi uma das que mostrou melhores resultados, principalmente nos meses de março-abril-maio. Resultado semelhante foi encontrado por Porfirio e colaboradores (2020), usando modelos físicos GL (radiação global). Esses autores também verificaram superestimções da radiação pelo ERA5 e GL. Essas superestimções variaram de +2 a +40 MW/m², semelhante aos resultados das Tabelas 10 e 11 deste trabalho. Segundo Porfirio e colaboradores (2020), esses erros são provenientes de suposições do algoritmo do GL, principalmente para refletância de céu claro, aerossóis e vapor de água.

Rampazo e colaboradores (2018) concluíram o mesmo. Nesse trabalho, os autores analisaram radiação solar, temperatura do ar, umidade relativa, velocidade do vento e chuva, com dados do modelo ECMWF (mesmo Centro que desenvolve o ERA5) e dados de satélite, para o estado de São Paulo. Constatou-se, nesse trabalho, que houve superestimção da radiação solar em 100% dos pontos analisados, o que está de acordo com nossos resultados (Tabela 10), mas não com os dados na Tabela 11. Para a velocidade do vento, 80% dos pontos foram superestimados, o que está de acordo com nossos resultados (Tabelas 10 e 11). Enquanto subestimaram a umidade relativa em 100% dos pontos, resultado coerente com os dados das Tabelas 10 e 11. Já a temperatura foi subestimada em 50% e superestimada em 50%, esse resultado está de acordo com os dados das Tabelas 10 e 11. Outra conclusão desse trabalho foi que a umidade relativa não apresentou boa exatidão aos dados observados (Tabelas 10 e 11). Esse resultado indica que, principalmente o NCEP/NCAR, é um produto que deve ser melhorado.

Babar e colaboradores (2019), trabalhando com radiação na Noruega, também verificaram superestimção do ERA5. Esse trabalho mostra, que a superestimção dessa reanálise também ocorre em locais de alta latitude. Nesse estudo, os autores verificaram erros de estimativa (MAE), para o ERA5, de 6,8 W.m⁻², o que é consistente com os resultados da Tabela 11. Uma possível explicação, apontada por esses autores, foi a de que o ERA5 representa de forma imprecisa as nuvens. Na China, Zhou e colaboradores (2019), também verificaram superestimção da radiação. Entre outros modelos, esse trabalho foi feito usando o NCEP/NCAR e Modern-ERA (desenvolvido pelo mesmo Centro do ERA5). O NCEP-NCAR, nesse trabalho, apresentou um RMSE de 76.3 W. m⁻² e um MAE de 68.8 W.m⁻², resultado muito similar pode ser observado na Tabela 10.

Também na China, trabalhando com modelos ERA e outros, Zhou e colaboradores (2017) verificaram que as reanálises são hábeis em simular a variância interanual da temperatura de superfície, com valores de coeficiente de correlação que variaram de 50 a 95%, exceto para uma região de grande altitude: o planalto tibetano. Resultados semelhantes de coeficiente de correlação podem ser vistos na Tabela 10. Lihare e colaboradores (2019), analisando também temperatura de superfície, mas no Canadá, encontraram que, para a maioria das estações do ano, o ERA-Interim, subestimava a temperatura (-0,2°C). Esse resultado é similar aos encontrados para Barra de São Miguel (Tabela 11).

De modo geral, os resultados do presente trabalho concordam com a literatura científica mais atual sobre o tema. Vale destacar que, as reanálises estão sendo constantemente melhoradas: sistemas de assimilação, dados de satélite e diversas parametrizações físicas; vêm sendo incorporadas aos modelos, o que vem conferindo grau a mais de refinamento (WANG et al., 2020).

4.2 Cálculos de evapotranspiração e balanço hídrico

Após cálculos, a série de evapotranspiração pelo método de Penman-Monteith precisou ser preenchida. Pelos resultados da validação para posterior preenchimento, pode-se perceber que, em comparação com o método de Penman-Monteith (P&M), os dados gradeados de Xavier e colaboradores (2016) e o método de Hargreaves-Samani (HS) tenderam a subestimar a ETo em -2,01 e -2,65 mm.dia⁻¹, respectivamente (Tabela 12 e Gráfico 7). Esses valores são elevados para estimativas diárias. Os erros associados à essas estimativas (RMSE) foram de 2,11 (Xavier) e 2,74 mm (HS) ou em porcentagem 29,96 (Xavier) e 38,25% (HS). Na Tabela 12, também é possível observar os valores de coeficientes de correlação (r), obtidos entre os dados observados (estação de Barra de São Miguel) e o banco de dados do Xavier (0,84) e o método de HS (0,70).

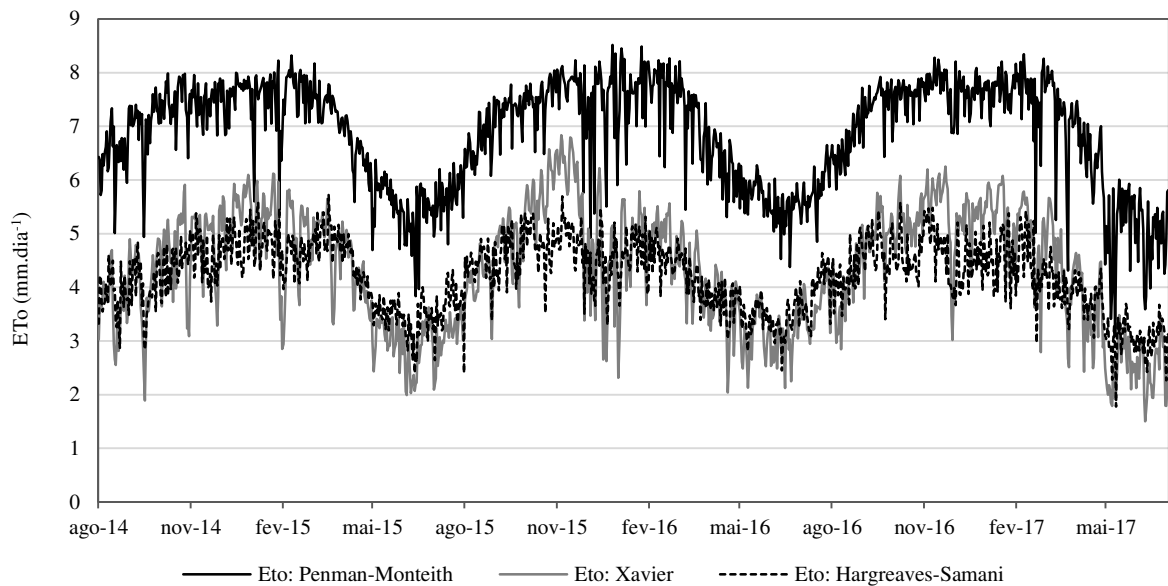


Gráfico 7 - Comportamento da evapotranspiração de referência (ETo), calculada através de diferentes métodos, para Barra de São Miguel (AL). Fonte: Autor.

METODOLOGIAS/DADOS EM COMPARAÇÃO	r	R ²	RMSE	MAE	BIAS	PMRE
XAVIER: PENMAN-MONTEITH	0.84	0.70	2.11	2.01	-2.01	29.96
HARGREAVES-SAMANI: PENMAN-MONTEITH	0.70	0.49	2.74	2.65	-2.65	38.25

Tabela 12 - Resultados por estatística de validação da evapotranspiração de referência (ETo). As unidades de medida de RMSE, MAE e VIÉS são as mesmas de ETo, ou seja, milímetro (mm). Fonte: Autor.

Desses valores de coeficiente, é possível concluir que o banco de dados de Xavier e colaboradores (2016), apresentou melhor valor de r , ou seja, os resultados se aproximaram mais de 1 (uma correlação perfeita). Alguns fatores podem explicar a leve superioridade do banco de dados Xavier em relação ao método de HS.

Um desses fatores é que os valores de Eto, calculados por Xavier e colaboradores (2016), também foram obtidos pelo método de Penman-Monteith. Apesar disso, esses valores foram interpolados, ou seja, dependiam de dados observados, que possivelmente estavam indisponíveis para Barra de São Miguel na obtenção das grades.

Já o método de HS, calculado conforme a Equação 10, tem como dado de entrada apenas a temperatura, ou seja, espera-se que seu desempenho seja inferior ao método P&M, que, por sua vez, se baseia em termos energéticos e aerodinâmico (SENTELHAS et al., 2010). Isso explica o desempenho do banco de dados de Xavier em comparação com o método de HS.

De modo geral, foi possível verificar que o Xavier e HS possuem potenciais parecidos de preenchimento de falhas de ETo para Barra de São Miguel. Para esse trabalho, Xavier foi usado para preencher as falhas dos dados de ETo, inicialmente calculada pelo método P&M. Vale ressaltar que, até a data de execução deste trabalho, o banco de dados de Xavier e colaboradores (2016), contava com valores de Eto de 1980 a meados de 2017. Assim sendo, o método de HS pode ser uma alternativa viável para períodos que não estão compreendidos nesse banco de dados. Estes resultados estão de acordo com algumas pesquisas feitas no Brasil, usando diferentes métodos.

Uma dessas pesquisas é a de Alencar e colaboradores (2015) que, comparando a ETo, estimada na ausência e presença de dados meteorológicos, para várias localidades de Minas Gerais, chegaram às seguintes conclusões: o método de HS, de modo geral, superestimou a ETo em $0,80 \text{ mm dia}^{-1}$; os erros associados à essas estimativas, pelo método de HS, variaram de $0,82$ a $1,44 \text{ mm dia}^{-1}$, com um valor médio de $1,12 \text{ mm dia}^{-1}$. Esses valores de erro médio são inferiores aos observados na Tabela 12, além disso, a ETo foi superestimada para Minas e subestimada para Barra de São Miguel. Isto ocorre, provavelmente, porque o método de HS foi desenvolvido para regiões de clima semiárido, com base em dados temperatura. Dessa forma, espera-se que esse método superestime a ETo em regiões com maior umidade e menores temperaturas.

Resultados semelhantes foram encontrados por Althoff e colaboradores (2020), nesse trabalho, os autores trabalharam na confecção de um banco nacional de dados gradeados de Eto, usando técnicas de *machine learning*. O banco de dados resultante apresentou erros médios de $0,242$ a $1,553 \text{ mm dia}^{-1}$ e vieses de -1.06 a $0.944 \text{ mm dia}^{-1}$. Sendo que, esses valores de erro médio foram maiores no Nordeste, provavelmente pela maior quantidade de registros na região, ou seja, os erros são proporcionais à quantidade de observações.

Costa e colaboradores (2017), trabalhando com dados de Eto também no Nordeste, mais precisamente em Alagoas, constataram que os erros médios associados às estimativas de Eto, pelo método de HS foram de $0,81 \text{ mm.dia}^{-1}$ (Água Branca); $0,73 \text{ mm.dia}^{-1}$ (Maceió); $0,89 \text{ mm.dia}^{-1}$ (Palmeira dos Índios); $0,94 \text{ mm.dia}^{-1}$ (Pão de Açúcar) e $1,28 \text{ mm.dia}^{-1}$ (Porto de Pedras). Esses resultados são inferiores aos 2 mm , encontrados para Barra de São Miguel (Tabela 12). Comparando com estes trabalhos, nossos resultados (Tabela 12) apresentaram desempenho abaixo do

esperado, com maiores erros associados às estimativas de Eto. A série de ETo preenchida, para o período analisado, é apresentada no gráfico a seguir (Gráfico 8).

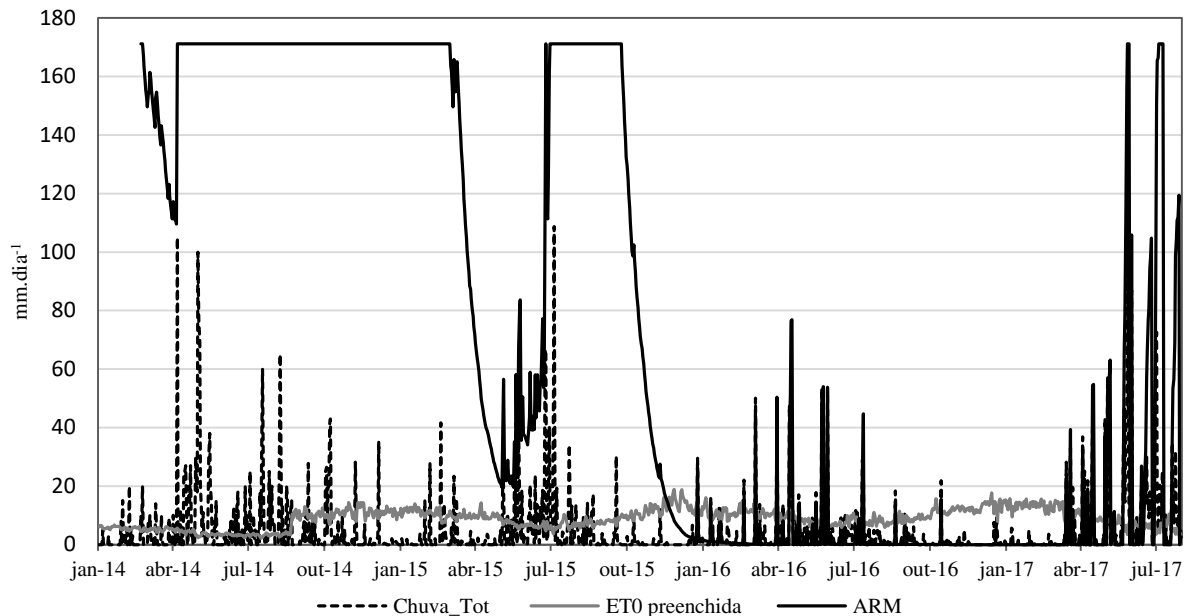


Gráfico 8 - Comportamento da chuva, armazenamento de água no solo (ARM) e evapotranspiração de referência (ETo) para Barra de São Miguel (AL). Fonte: Autor.

Foi possível observar o comportamento da ETo e chuva durante os anos de 2014 a meados de 2017 (Gráfico 8). Com relação a ETo, nos meses mais quentes: novembro (média de 12,30 mm.dia⁻¹), dezembro (média de 12,51 mm.dia⁻¹) e janeiro (média de 10,37 mm.dia⁻¹); a ETo diária é maior. Por outro lado, nos meses mais frios: junho (média de 5,87 mm.dia⁻¹), julho (média de 5,81 mm.dia⁻¹) e agosto (média de 6,93 mm.dia⁻¹); a ETo diária é menor. O período com maiores precipitações são os meses de abril (média de 26,43 mm.mês⁻¹), maio (média de 46,41 mm.mês⁻¹), junho (média de 34,24 mm.mês⁻¹) e julho (média de 36,51 mm.mês⁻¹); e o período com as menores precipitações são os meses de outubro (média de 8,80 mm.mês⁻¹), novembro (média de 2,80 mm.mês⁻¹), dezembro (média de 4,13 mm.mês⁻¹) e janeiro (média de 5,12 mm.mês⁻¹).

Comparando os valores de chuva, de 2014 a 2015, com o Gráfico 1.A (comportamento normal), percebemos que o comportamento sazonal de chuva é normal, porém a quantidade de chuvas nesses anos está muito abaixo do comportamento normal. Segundo Marengo e colaboradores (2015), a causa dessa falta de chuva foi o chamado bloqueio atmosférico. Este sistema de alta pressão,

extremamente raro, prejudicou o transporte de umidade da Amazônia; o desenvolvimento dos principais sistemas causadores de chuva, como as ZCAS (Zona de Convergência do Atlântico Sul) e as frentes frias. Esse sistema anômalo também explica os baixos valores de água armazenada (ARM) e excedente hídrico no solo (EXC) (Gráfico 9), em comparação com o comportamento normal (Gráfico 1.B).

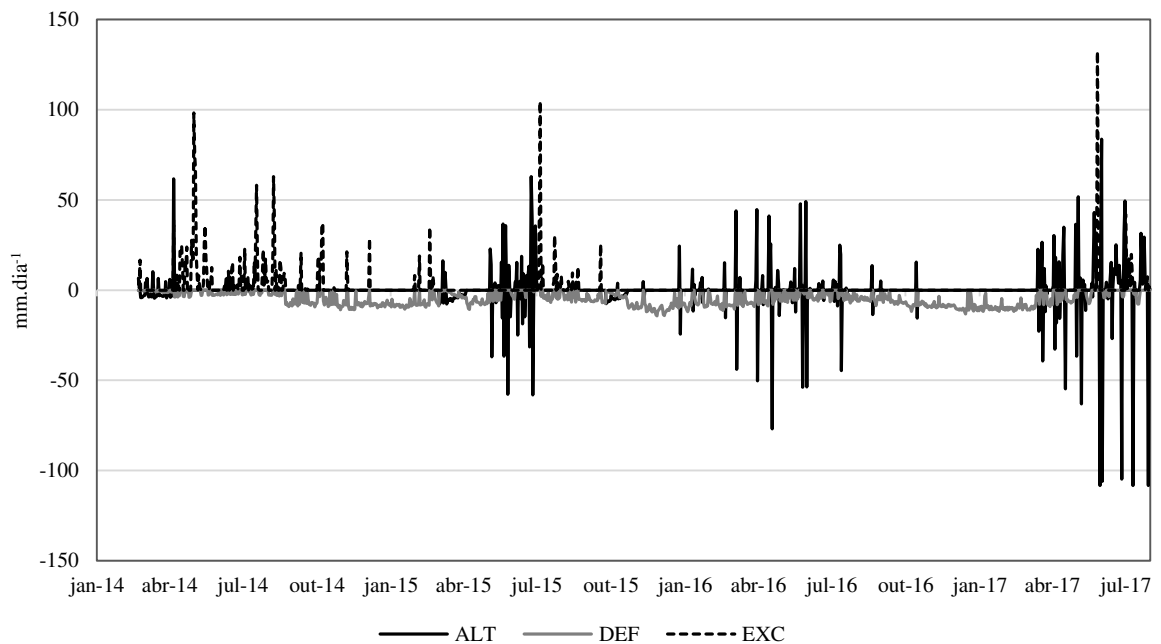


Gráfico 9 – Balanço hídrico sequencial diário, dos anos de 2014 a meados de 2017, para Barra de São Miguel (AL). Fonte: Autor.

A seguir, no Gráfico 10 e Tabela 13, são apresentados os resultados dos testes de normalidade: Shapiro-Wilk e Lilliefors; apresentados na seção 3.3.4 da metodologia.

4.3 Cálculos de normalidade dos regressores

Os resultados dos testes de Shapiro-Wilk e Lilliefors ($P < 0,05$) mostraram-se não normais para a maioria dos regressores. Isso porque os p-valores encontrados são menores que o nível de significância ($p\text{-valor}_w < 0,05$ e $p\text{-valor}_D < 0,05$). As

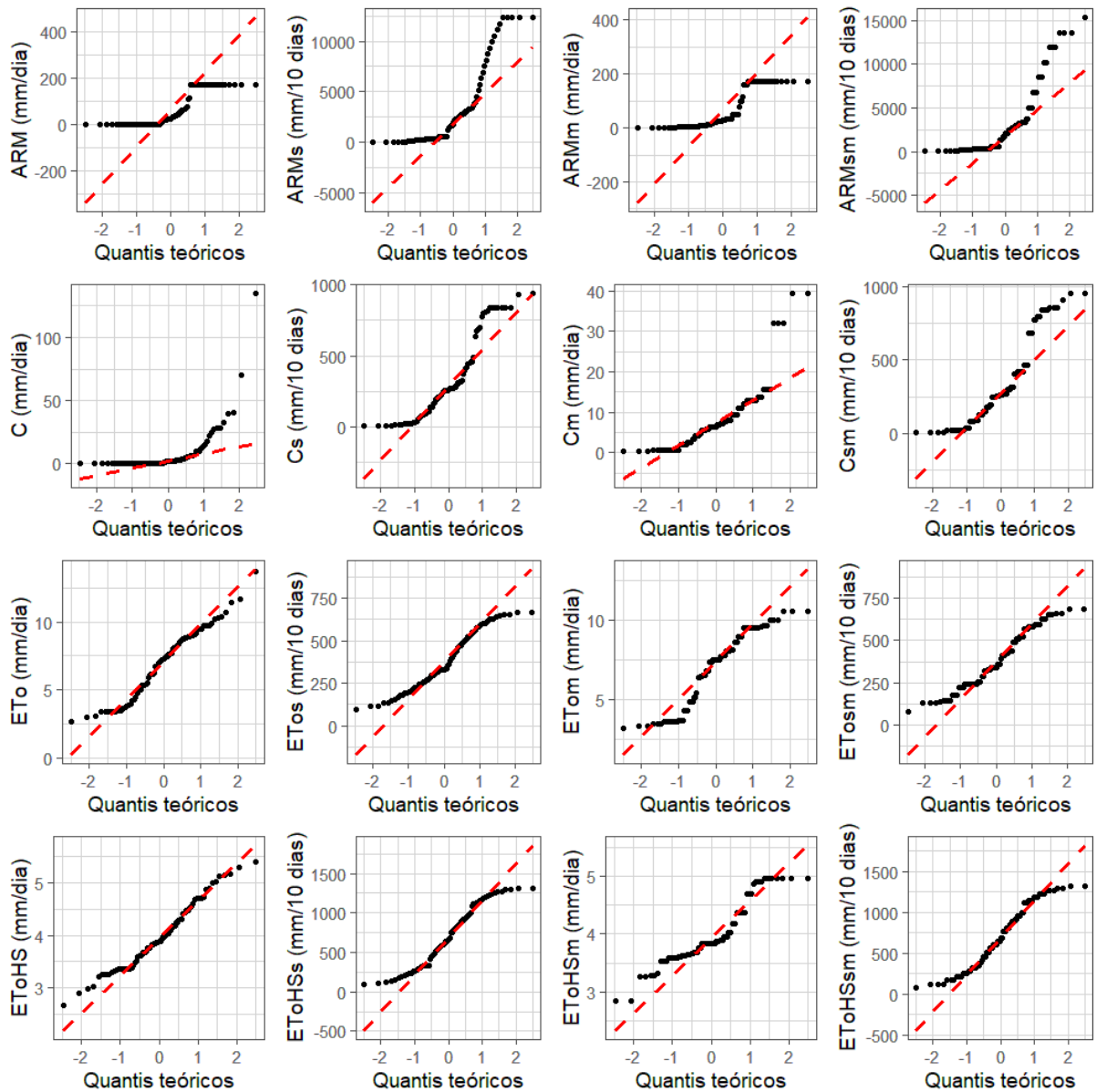
exceções foram: ETo ($p\text{-valor}_w = 0,2072 > 0,05$) e EToHS ($p\text{-valor}_D = 0,2929 > 0,05$ e $p\text{-valor}_w = 0,1081 > 0,05$) (Tabela 13).

Regressor	Nº de observações analisadas	Lilliefors (Kolmogorov - Smirnov)		Shapiro - Wilk	
		D	p-valor	W	p-valor
ARM	77	0,216	$1,50 \cdot 10^{-9}$	0,740	$2,25 \cdot 10^{-10}$
ARM_s	77	0,205	$1,39 \cdot 10^{-8}$	0,770	$1,24 \cdot 10^{-9}$
ARM_m	77	0,271	$2,23 \cdot 10^{-15}$	0,738	$2,03 \cdot 10^{-10}$
ARM_{sm}	77	0,231	$5,37 \cdot 10^{-11}$	0,764	$8,61 \cdot 10^{-10}$
C	77	0,336	$2,20 \cdot 10^{-16}$	0,459	$2,36 \cdot 10^{-15}$
C_s	77	0,175	$3,64 \cdot 10^{-6}$	0,875	$1,79 \cdot 10^{-6}$
C_m	77	0,177	$2,77 \cdot 10^{-6}$	0,756	$5,54 \cdot 10^{-10}$
C_{sm}	77	0,176	$3,28 \cdot 10^{-6}$	0,876	$1,96 \cdot 10^{-6}$
ETo	77	0,083	0,2072*	0,964	$2,72 \cdot 10^{-2}$
ETo_s	77	0,126	$4,17 \cdot 10^{-3}$	0,948	$3,24 \cdot 10^{-3}$
ETo_m	77	0,139	$8,76 \cdot 10^{-4}$	0,915	$7,98 \cdot 10^{-5}$
ETo_{sm}	77	0,115	$1,40 \cdot 10^{-2}$	0,953	$6,34 \cdot 10^{-3}$
EToHS	77	0,078	0,2929*	0,974	0,1081*
EToHS_s	77	0,121	$6,97 \cdot 10^{-3}$	0,938	$9,36 \cdot 10^{-4}$
EToHS_m	77	0,185	$6,59 \cdot 10^{-7}$	0,910	$4,52 \cdot 10^{-5}$
EToHS_{sm}	77	0,105	$3,64 \cdot 10^{-2}$	0,940	$1,25 \cdot 10^{-3}$

Tabela 13 - Resultados dos testes de Shapiro-Wilk e Lilliefors. Os regressores com "*" têm distribuição normal, ou seja, $p\text{-valor} > \alpha = 0,05$. Fonte: Autor.

As funções de distribuição acumulada estimada (FDA), que representa o valor de cada observação contra a probabilidade acumulada estimada da observação, mostra que o ajuste com a reta de distribuição normal é pobre. As FDA's dos regressores tenderam a ficar muito afastadas da linha reta de distribuição ajustada (Gráfico 10).

Esses dados assimétricos formaram uma linha curva, evidenciando a não normalidade dos dados. De qualquer forma, rejeitou-se a H_0 , em que, os regressores (variáveis independentes) seguem uma distribuição normal. Como a regressão Logit não é afetada pela forma funcional das variáveis independentes, já que o termo de erro segue uma distribuição binomial, confirmou-se que esse modelo foi o mais adequado (GUJARATI; PORTER, 2011).



4.4 Modelagens

Os resultados das modelagens mostraram que para nenhum modelo em que ARM e Eto, por Penman-Monteith, foram regressores, os coeficientes foram significativos ($P < 0,05$). Provavelmente, esses resultados estão associados com a grande quantidade de parâmetros para o cálculo dessas variáveis, o que significa que uma grande quantidade de incertezas foi incorporada aos modelos. Por outro lado, os

modelos que tiveram C_s , C_{sm} , $EToHS$ e $EToHS_m$ como regressores foram significativos a 5% (Tabela 14).

REGRESSOR (X_i)	INTERCEPTO (β_1)	COEFICIENTE (β_2)	P-VALOR DE β_2	AIC
ARM	0.3180063	0.0003602	0.911	108.53
ARM_s	4.460e-01	-3.278e-05	0.590	108.25
ARM média	0.3222852	0.0002964	0.930	108.53
ARM_{sm}	4.932e-01	-4.614e-05	0.422	107.89
C	0.307079	0.004292	0.740	108.42
C_s	0.9897604	-0.0019124	0.02572*	103.26
C_m	0.07579	0.03277	0.295	107.33
C_{sm}	1.0565259	-0.0021038	0.01629*	102.31
ETo	-0.004037	0.049407	0.598	108.26
ETo_s	0.856517	-0.001363	0.335	107.6
ETo_m	-0.24043	0.08376	0.411	107.86
ETo_{sm}	1.010168	-0.001745	0.2221	107.02
EToHS	-2.9947	0.8450	0.0381*	103.85
EToHS_s	0.8212294	-0.0006764	0.273	107.32
EToHS_m	-3.8653	1.0670	0.0339*	103.48
EToHS_{sm}	0.8608890	-0.0007290	0.2394	107.13

Tabela 14 - Resultado da regressão Logit por regressor. Os p-valores com "*" são significativos à 5%.
Fonte: Autor.

Esses resultados, provavelmente, estão associados ao fato de que esses regressores foram estimados com observações *in situ*. Isso porque os dados de chuva no período de 2014 a 2017 foram coletados sem falhas (Gráfico 2.B). Da mesma forma, os dados de temperatura, para cálculo da evapotranspiração por Hargreaves-Samani, estavam disponíveis e sem falhas para o período de dados de treino dos modelos, ou seja, de março a junho de 2015 a 2017 (Gráfico 2.A). O nível de significância desses modelos, entre 1 e 5%, também foram encontrados por outros autores, trabalhando em outros contextos de modelagem de florescimento por meio da regressão logística.

Como Parra-Tabla e colaboradores (2000), trabalhando com regressão logística no período floral de orquídeas na Península de Yucatán, México, chegaram a um p-valor também significativo (p-valor<0,05). Resultados semelhantes foram obtidos por Goodwin e colaboradores (1998). Estes últimos autores, modelaram através da regressão logística, a capacidade de invasão de plantas no Canadá, a partir de caracteres biológicos, como a floração.

Esses quatro modelos, com p-valor significativo, também foram os que apresentaram os menores valores de AIC, indicando que são os que mais se

aproximam do modelo real desconhecido, que descreve o florescimento. Os sinais dos coeficientes indicam que a chuva acumulada afeta de maneira negativa as chances de 1% de florescimento (Modelo 1 e 2), enquanto a evapotranspiração afeta de maneira positiva (Modelo 3 e 4). Esses resultados também são mostrados de maneira visual no Gráfico 11.

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = 0.9897604 - 0.0019124 \cdot C_s \quad \text{(Modelo 1)}$$

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = 1.0565259 - 0.0021038 \cdot C_{sm} \quad \text{(Modelo 2)}$$

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = -2.9947 + 0.8450 \cdot \text{EToHS} \quad \text{(Modelo 3)}$$

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = -3.8653 + 1.0670 \cdot \text{EToHS}_m \quad \text{(Modelo 4)}$$

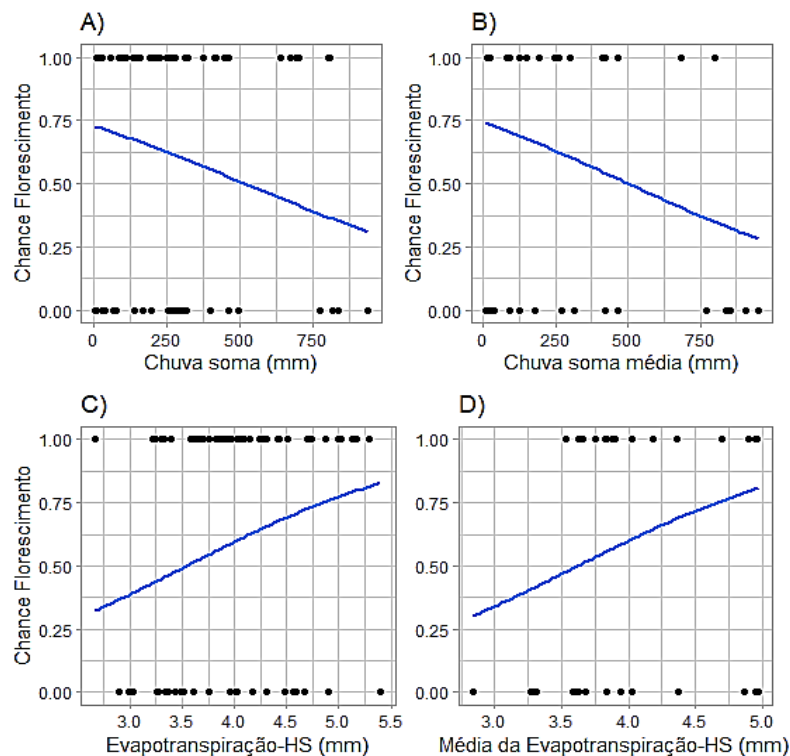


Gráfico 11 - Curva Logit, tendo como regressor: A) Chuva soma; B) Chuva média; C) Evapotranspiração por Hargreaves-Samani e D) Média da evapotranspiração por Hargreaves-Samani. Fonte: Autor.

Os dados de CI's e as OR's por modelo, se mostraram satisfatórias (Tabela 15). Os resultados dos CI's indicam que, para todos os modelos, o coeficiente da variável independente está dentro do intervalo de 95% (Tabela 15). Já os resultados das OR's podem ser interpretados da seguinte maneira: para cada variação de 1 mm.dia⁻¹ (EToHS e EToHS_m) ou de 1mm.10 dias⁻¹ (C_s e C_{sm}), as chances de ocorrência de, pelo menos, 1% de florescimento diminuem em 0.9981 (C_s) e 0.9979 (C_{sm}) vezes; e aumentam 2.3279 (EToHS) e 2.9066 (EToHS_m) vezes (Tabela 15).

	OR (β_2)	CI para β_2 (2.5%_97.5%)
MODELO 1	0.9981	0.9963_0.9997
MODELO 2	0.9979	0.9961_0.9996
MODELO 3	2.3279	1.0807_5.4272
MODELO 4	2.9066	1.1408_8.4383

Tabela 15 – Resultado da razão de chances (OR) e intervalo de confiança (CI) por modelo. Fonte: Autor.

Avaliando o poder de discriminação desses modelos, pela Matriz de Confusão (Tabela 16), pudemos observar que os 4 modelos têm poder de discriminação de Verdadeiro Positivo (VP) semelhantes, que correspondem à 14,29% (11 em 77) para os modelos 1 e 2; 20,78% (16 em 77) para o modelo 3 e 18,18% (14 em 77) para o modelo 4. Em termos Verdadeiro Negativo (VN) e Falso Positivo, verificou-se que o modelo 4 é melhor, com valores de 54,55% (42 em 77) e 3,90% (3 em 77), respectivamente. Com esses valores foi possível calcular a acurácia dos modelos.

MODELO 1			MODELO 2		
Valor Estimado	Valor Observado		Valor Estimado	Valor Observado	
	Y=1	Y=0		Y=1	Y=0
	$\hat{Y}=1$	11		6	$\hat{Y}=1$
$\hat{Y}=0$	21	39	$\hat{Y}=0$	21	40
MODELO 3			MODELO 4		
Valor Estimado	Valor Observado		Valor Estimado	Valor Observado	
	Y=1	Y=0		Y=1	Y=0
	$\hat{Y}=1$	16		6	$\hat{Y}=1$
$\hat{Y}=0$	16	39	$\hat{Y}=0$	18	42

Tabela 16 - Resultados da Matriz de confusão por modelo. Fonte: Autor.

As acurácias calculadas foram: 64,94% (Modelo 1); 66,23% (Modelo 2); 71,43% (Modelo 3) e 72,73% (Modelo 4). Gutierrez e colaboradores (2008),

encontraram valores de VP e FN de 47,32% (530 em 1120) e 2,77% (31 em 1120), respectivamente, trabalhando com florescimento de plantas invasoras na cultura do girassol na cidade de Matabueyes na Espanha. Embora a capacidade de discriminação seja, em porcentagem, parecida, a acurácia do modelo ajustado por esses autores é bem diferente, chegando a 94,6%. Esse valor mais acurado, provavelmente, é resultante do maior número de observações.

Ainda, mensurando a capacidade de predição dos modelos propostos, as curvas ROC (Gráfico 12) mostram que o modelo 1, que teve Cs como regressor, foi melhor em classificar os dados como verdadeiros positivos. Isso porque, o ponto de comparação dessa curva (0,312; 1,00) está mais a noroeste do que os pontos de comparação das curvas que tiveram Csm (0,656; 0,644), EToHS (0,531; 0,844) e EToHSm (0,438; 0,933) como regressores. Em outras palavras possui ponto mais próximo do ponto (0,1), onde a taxa de falsos positivos seria 0 e a taxa de verdadeiros positivos seria 1 (ou 100%).

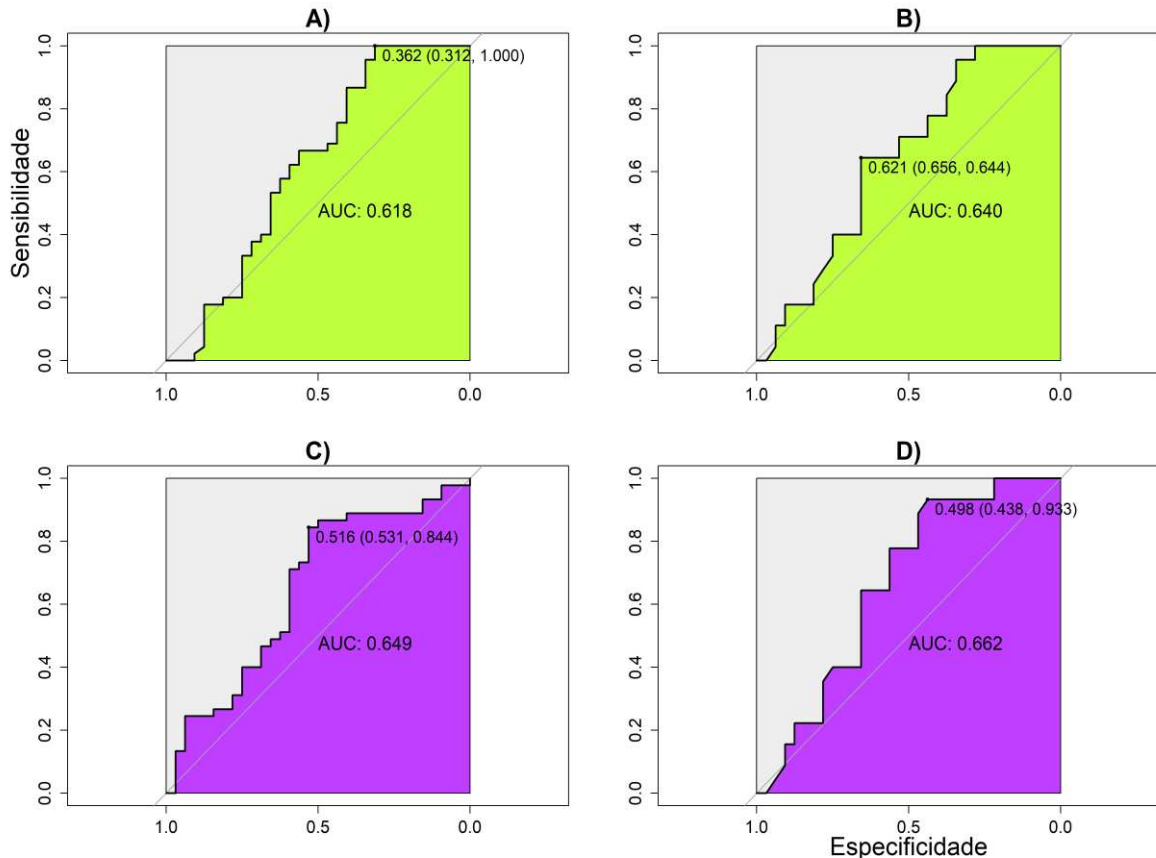


Gráfico 12 - Resultados da curva ROC por modelo: A) Cs; B) Csm; C) EToHS e D) EToHSm. Fonte: Autor.

Analisando os valores das curvas ROC, embora o valor de especificidade do Modelos 1 (0.312) seja menor que as dos outros modelos, os valores de sensibilidade desse modelo é o maior, chegando a 1.0, enquanto para o Modelos 2 é 0,644; para o modelo 3 é 0,844 e para o Modelo 4 é 0.933. Os valores de plotados à esquerda das curvas: 0.362 (Graf. 12.A), 0.621 (Graf. 12.B), 0.516 (Graf. 12.C) e 0.498 (Graf. 12.D), correspondem aos pontos de corte que otimizam a sensibilidade e especificidade, visto que, são duas medidas inversamente proporcionais. Além disso, os valores de AUC mostram que a curva ROC D (0.662) tem uma melhor capacidade preditiva que a curva C (0.649), que tem capacidade preditiva que a curva B (0,640) e A (0,618).

Os Modelos 1 e 2 também apresentam maiores valores de “pseudo” R^2 com valores de 8,92% (Modelo 1) e 10,47% (Modelo 2) contra 7,95% (Modelo 3) e 8,56% (Modelo 4) (Tabela 17). Isso evidencia que os modelos com chuva como regressor são mais bem ajustados e explicativos do florescimento. Embora esses valores pareçam pequenos, vale lembrar que estas são medidas de “pseudo” R^2 e não R^2 , para o “pseudo” R^2 , valores de ajuste na ordem de 20% são considerados ótimos (STEINBERG & COLLA, 2000).

	Pseudo R^2 - Nagelkerke	RMSE	MAE
MODELO 1	0.08920221	1.292	1.082
MODELO 2	0.10467	1.345	1.118
MODELO 3	0.0794787	1.465	1.120
MODELO 4	0.08556351	1.589	1.136

Tabela 17 - Resultado de "pseudo" R^2 , RMSE e MAE por modelo. Fonte: Autor.

Por outro lado, os Modelos 1 e 2 apresentaram os menores valores de RMSE e MAE, mas as unidades de medida desses regressores são diferentes (Tabela 8). Enquanto dos Modelos 1 e 2 estão em $\text{mm} \cdot 10 \text{ dias}^{-1}$, os regressores dos modelos 3 e 4 estão em $\text{mm} \cdot \text{dia}^{-1}$. Isso significa que, os erros associados à esses regressores também serão diferentes.

Analisando o desempenho geral, entre os modelos que tiveram chuva como regressores (Modelos 1 e 2), o Modelo 2 tem desempenho geral levemente maior que o Modelo 1. Enquanto que, entre os modelos que tiveram evapotranspiração por Hargreaves-Samani como regressores (Modelos 3 e 4), o Modelo 4 tem desempenho geral levemente maior que o Modelo 3. Como a evapotranspiração pelo método de

Hargreaves-Samani depende apenas de valores de temperatura máxima, média e mínima (Equações 10 e 11); os Modelos 3 e 4 podem ser escritos da seguinte forma:

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = -2.9947 + [(0.00079 \cdot R_o) \cdot (T_{\text{med}} + 17.8) \cdot (\sqrt{T_{\text{max}} - T_{\text{min}}})] \quad \text{(Modelo 3.1)}$$

$$\text{Ln} \left(\frac{P_{\text{florescer}}}{1-P_{\text{florescer}}} \right) = -3.8653 + [(0.00101 \cdot R_o) \cdot (T_{\text{med}} + 17.8) \cdot (\sqrt{T_{\text{max}} - T_{\text{min}}})] \quad \text{(Modelo 4.1)}$$

Os resultados sugerem que a influência da chuva sobre o florescimento da cana se dá de maneira acumulativa, enquanto a evapotranspiração afeta de maneira média. Já que os valores diários de evapotranspiração sofrem pouca variação para uma mesma época do ano. Os resultados também sugerem que pequenas alterações de volume de chuva acumulado, associados à aumentos de temperatura (e por consequência de evapotranspiração) são capazes de provocar 1% de florescimento na cana-de-açúcar.

Variáveis ambientais como temperaturas e estresse hídrico estão associados à indução floral de várias culturas. Existem relatos que o estresse hídrico proporciona a dormência da gema vegetativa, que, ao receber água, sai do estado de dormência. Isso acontece pois há a reconexão dos vasos condutores que chegam até as gemas e, assim, estimulam a passagem do estágio vegetativo para o reprodutivo.

Esse comportamento foi observado em árvores frutíferas subtropicais, como a jabuticaba (OLIVEIRA et al., 2019; DAVENPORT, 2011). Os botões florais de avelãs precisam de reidratação para saírem do estado de endodormência para alcançar a floração (AYYANATH et al., 2021). A hidratação pós período de *déficit* hídrico, evidenciando a importância da lâmina de água na indução do florescimento, fica evidente em plantas de café (CRISOSTO et al., 1992). Nas quais a irrigação é utilizada não só para induzir o florescimento, mas também para sua uniformização. Outros trabalhos na Literatura também associam condições ambientais e florescimento de plantas.

Como Wells e colaboradores (1998), usando regressão logística, constataram que para orquídeas na Inglaterra, o florescimento é crescente com o aumento de chuvas acumuladas e redução de temperatura. Wallach e colaboradores (2018) trabalhando com modelagem de QTL na previsão do período de floração no feijoeiro (*Phaseolus vulgaris*), constataram que o florescimento depende da temperatura diária e do comprimento do dia. Bogawski e colaboradores (2019) também evidenciaram a

influência da temperatura para modelar o padrão espacial do início da floração de Betulas.

Ohigashi e colaboradores (2019), trabalhando com desenvolvimento de modelos para prever o início do florescimento da soja silvestre, evidenciaram que, para florescerem, as sojas selvagens exigiram temperaturas mais altas mesmo em condições de dias curtos para alguns locais do Japão. Os resultados desses trabalhos corroboram com o presente estudo, onde foram verificadas associações entre ambiente, clima e florescimento.

5. CONCLUSÕES

Os resultados das validações de dados meteorológicos mostram que, ERA5 e Xavier possuem potenciais parecidos de preenchimento de falhas, das variáveis meteorológicas em estudo (temperatura máxima, temperatura média, temperatura mínima, radiação, umidade relativa e velocidade do vento), da estação do INMET em Coruripe (AL). Já, para Barra de São Miguel, os bancos de dados do INMET Maceió, ERA5 e Xavier possuem potenciais parecidos de preenchimento de falhas.

Esses dados validados podem ser utilizados nos cálculos balanço hídrico sequencial diário, pelo método de Thornthwaite-Mather, e evapotranspiração, pelos métodos de Penman-Monteith e Hargreaves-Saman (HS). A série de evapotranspiração pelo método de Penman-Monteith precisa ser preenchida. De modo geral, verificamos que o Xavier e HS possuem potenciais parecidos de preenchimento de falhas de ETo pelo método de Penman-Monteith para Barra de São Miguel. Embora, estes dois conjuntos de dados tendem a subestimar a ETo.

Comparando os valores de chuva observados em Barra de São Miguel, de 2014 a 2015, com dados de chuva normal, percebemos que o comportamento sazonal das chuvas é normal. Porém a quantidade de chuvas nesses anos estava muito abaixo do comportamento normal. Essa situação anômala explica os baixos valores de água armazenada (ARM) e excedente hídrico no solo (EXC), encontrados no balanço hídrico sequencial diário.

Os resultados dos testes de Shapiro-Wilk e Lilliefors à um nível de 5% de significância, mostram que a maioria das variáveis independentes não seguem uma distribuição normal, o que corroborou com o uso da regressão logística.

Dos modelos gerados, apenas aqueles que apresentam chuva (C_s e C_{sm}) e evapotranspiração (EToHS e EToHS_m), pelo método de Hargreaves-Samani, como regressores são significativos no caso em estudo ($P < 0,05$). Esses quatro modelos são os que apresentaram os menores valores de AIC, indicando que são os que mais se aproximam do modelo real desconhecido, que descreve o florescimento. Os sinais dos coeficientes indicam que a chuva acumulada afeta de maneira negativa as chances de 1% de florescimento, enquanto a evapotranspiração afeta de maneira positiva.

Analisando o desempenho geral, entre os modelos nos quais chuva foi utilizada como regressor, o modelo com C_{sm} tem desempenho geral levemente maior

que o modelo com Cs. Enquanto que, entre os modelos que tiveram evapotranspiração por Hargreaves-Samani como regressores, o modelo com EToHSm tem desempenho geral levemente maior que o modelo com EToHS. Os resultados sugerem que a influência da chuva sobre o florescimento da cana se dá de maneira acumulativa, enquanto a evapotranspiração afeta de maneira média.

REFERÊNCIAS

- ABU-ELLAIL, F.F.B., MCCORD, P.H. Temperature and Relative Humidity Effects on Sugarcane Flowering Ability and Pollen Viability Under Natural and Seminatural Conditions. **Sugar Tech**, n. 21, pg. 83–92, 2019.
- Ageitec. Variedades: Cana-de-açúcar, 2020. Disponível em: <http://www.agencia.cnptia.embrapa.br/gestor/cana-deacucar/arvore/CONTAG01_42_1110200717570.html>. Acesso em: <9 de outubro de 2020>.
- ALENCAR, L.P., SEDIYAMA, G.C., MANTOVANI, E.C. Estimation of reference evapotranspiration (ET_o) under FAO standards with missing climatic data in Minas Gerais, Brazil. **Engenharia Agrícola**, Jaboticabal, vol. 35, n. 1, pg. 39-50, 2015.
- ALLEN, R. G., PEREIRA, L. S., RAES, D., SMITH, M. Crop evapotranspiration: Guidelines for computing crop water requirements. Rome: FAO. **Irrigation and Drainage Paper**, n. 56, 300 pg., 1998.
- ALTHOFF, D., DIAS, S.H.B., FILGUEIRAS, R., NEIVA RODRIGUES, L. ET_o-Brazil: a daily gridded reference evapotranspiration dataset for Brazil (2000–2018). **Water Resour. Res.**, vol. 56, pg. 1-16, 2020.
- ALVES, A. O florescimento da cana-de-açúcar tem época certa pra ser evitado. Canal Jornal da Bioenergia, 2016. Disponível em: <<https://www.canalbioenergia.com.br/como-inibir-o-florescimento-da-cana-de-acucar/>>. Acesso em: <21 de junho de 2021>.
- ARALDI, R., SILVA, F.M.L., ONO, E.O., RODRIGUES, J.D. Flowering in sugarcane. **Ciência Rural**, Santa Maria (RS), vol. 40, n. 3, pg. 694-702, 2010.
- AYYANATH, M.M., SHUKLA, M.R., SAXENA, P.K. Role of water percolation in reproductive physiology of hazelnut (*Corylus* spp.). **Environmental and Experimental Botany**, vol. 182, 2021.

- BABAR, B., GRAVERSEN, R., BOSTROM, T. Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5. **Sol. Energy**, vol. 182, pg. 397–411, 2019.
- BARBIERI, V., BACHI, O.O.S., VILLA NOVA, N.A. Espaçamento em cana-de-açúcar. *In*: Congresso Nacional da Stab, Rio de Janeiro (RJ), vol. 3/4, pg. 512-522, 1981.
- Google Maps. Barra de São Miguel, AL, 2021. Disponível em: <<https://www.google.com.br/maps/dir/Barra+de+S%C3%A3o+Miguel,+AL/Coruripe,+AL/@-9.9729332,-36.1664487,11z/data=!3m1!4b1!4m13!4m12!1m5!1m1!1s0x70155645aaa0e8f:0x98ece37939ddd0d8!2m2!1d-35.885903!2d-9.8237028!1m5!1m1!1s0x7046c83e7065787:0x3d7a5648494485f3!2m2!1d-36.1761498!2d-10.1249995>>. Acesso em: <25 de Janeiro de 2021>.
- BARROS, A.H.C., DE ANDRADE JÚNIOR, A.S., BASTOS, E. A., DA SILVA, A. B. Simulação do crescimento e da produtividade da cana-de-açúcar em Teresina, PI. Embrapa Solos, Rio de Janeiro (RJ), doc. 204, 29 pg., 2018.
- BARROS, A.H.C., FILHO, J.C.A., DA SILVA, A.B., SANTIAGO, G.A.C.F. Climatologia do Estado de Alagoas. Embrapa Solos, Recife (PE), doc. 211, 32 pg., 2012.
- BERTOLLO, L.E. **Efeito do Ethephon na maturação e na inibição da floração sob condições naturais de indução fotoperiódica em cana-de-açúcar.** Dissertação (Mestrado em Agronomia) - Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal, 55f. 2020.
- BIELENKI JUNIOR, C., SANTOS, F. M. D., POVINELLI, S. C. S., MAUAD, F. F. Alternative methodology to gap filling for generation of monthly rainfall series with GIS approach. **RBRH**, vol. 23, pg. 1-10, 2018.
- BLAINSKI, E., GARBOSSA, L. H. P., ANTUNES, E. N. Estações hidrometeorológicas automáticas: recomendações técnicas para instalação. **Epagri**, doc. 240, 43 pg., 2012.

- BOGAWSKI, P., GREWLING, Ł., JACKOWIAK, B. Predicting the onset of *Betula pendula* flowering in Poznań (Poland) using remote sensing thermal data. **Science of the Total Environment**, vol. 658, pg. 1485–1499, 2019.
- BOZDONGAN. H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, vol. 52, n. 3, pg. 345-370, 1987.
- CARABALLOSO, V., JORGE, H., GARCÍA, H., GONZÁLEZ, A., BERNAL, N., CÉSPEDES, A., RODRÍGUEZ, R., PUCHADES, Y., ARENCIBIA, A.D. Management Flowering Ability to Increase Efficiency in the Sugarcane Breeding Program. **Sugar Tech**, vol. 14, n. 1, pg. 47–52, 2012.
- CARDOZO, A.C.C. **Florescimento da cana-de-açúcar: efeitos genotípicos, climáticos, perdas e estratégias de controle**. Tese (Doutorado em Agricultura) - Faculdade de Ciências Agrônômicas – UNESP, Botucatu, 36f. 2017.
- CHAVES, L.H.G., FARIAS, C.H.A. Variabilidade espacial do estoque de carbono nos Tabuleiros Costeiros da Paraíba: Solo cultivado com cana-de-açúcar. **Revista Brasileira de Ciências Agrárias**, Recife (PE), vol. 3, n. 1, pg. 20-25, 2008.
- CHEN, M., SHI, W., XIE, P., SILVA, V.B.S., KOUSKY, R., HIGGINS, R.W., JANOWIAK, J.E. Assessing objective techniques for gauge-based analyses of global daily precipitation. **Journal of Geophysical Research Atmospheres**, vol. 113, pg. 1-13, 2008.
- Climatempo. Climatologia do município de Barra de São Miguel (AL), 2020. Disponível em: <<https://www.climatempo.com.br/climatologia/5692/barradesaomiguel-al>>. Acesso em: <25 de outubro de 2020>.
- COELHO, C.P., NETTO, A.P.C., COLASANTI, J., CHALFUN-JUNIOR A. A proposed model for the flowering-signaling pathway of sugarcane under photoperiodic control. **Genet Mol Res Genet Mol Res.**, vol. 12, n. 2, pg. 1347–1359, 2013.
- COLETI, J.T. Inhibition of flowering by ethephon and its influence on sugarcane quality in Brazil. **Proc int Soc Sug Cane Technol**, vol. 19, pg. 298-304, 1986.

- Conab. Acompanhamento da safra brasileira de cana-de-açúcar, Safra 2017/18, vol. 4, n. 3, pg. 1-77, 2017.
- Conab. Acompanhamento da safra brasileira de cana-de-açúcar, Safra 2020/21, vol. 4, n. 4, pg. 1-62, 2021.
- COSTA, J.A., RODRIGUES, G.P., DA SILVA, N.D., SOBRINHO, O.P.L., COSTA, L.D.A. Avaliação de métodos de estimativa da evapotranspiração de referência para Alagoas. **Agrometeoros**, vol. 25, n.1, pg.173-179, 2017.
- CRISOSTO, C. H., GRANTZ, D. A., MEINZER, F. C. Effects of water deficit on flower opening in coffee (*Coffea arabica L.*). **Tree Physiology**, vol. 10, n. 2, pg. 127–139, 1992.
- CRUZ, M. A. S., ROCHA, R. O. F., SILVA, A. G., AMORIM, J. R. A., ARAGÃO, R. GeoTAB: Climate, water resources and watersheds in the work region of the Embrapa Tabuleiros Costeiros. **Scientia Plena**, vol. 13, n. 10, pg. 1-11, 2017.
- DALCHIAVON, F.C., CARVALHO, M.P., DE ANDRADE, F.G., MONTANARI, R., ANDREOTTI, M. Productivity of the sugarcane and definition of specific zones of soil management. *Semina: Ciências Agrárias*, Londrina (SC), vol. 34, n. 5, pg. 2077-2088, 2013.
- DALLAL, G.E., WILKINSON, L. An analytic approximation to the distribution of Lilliefors's test for normality. **The American Statistician**, vol. 40, pg. 291–296, 1986.
- DE ABREU, M.L., SILVA, M.A., TEODORO, I., DE HOLANDA, L.A., NETO, G.D.S. Growth and productivity of sugarcane varieties as affected by water availability in the Coastal Tablelands of the Alagoas State, Brazil. **Rev. Bragantia**, Campinas (SP), vol. 72, n. 3, pg. 262-270, 2013.
- DE ASSIS, F.N., DE ARRUDA, H.V., PEREIRA, A.R. **Aplicações de estatística à climatologia: Teoria e prática**. Ed. Universitária, UFPel, 1996.
- DONALDSON, R.A., SINGELS, A. Yields and estimated economic returns from using ethephon to suppress flowering in annually harvested sugarcane. **South African Sugar Association Experiment Station**, vol. 78, pg. 123-136, 2004.

- DURAI, A.A., GOVINDARAJ, P., PAZHANY, A.S. Flowering Behaviour of Sugarcane Genotypes from Different Agro Climatic Zones of India. **Sugar Tech**, vol. 16, n. 2, pg. 157–163, 2014.
- ECMWF. European Centre for Medium-Range Weather Forecasts, 2021. Disponível em:<www.ecmwf.int>. Acesso em:< 15 de fevereiro 2019>.
- Embrapa. Delimitação da área de atuação da Embrapa Tabuleiros Costeiros. Embrapa Tabuleiros Costeiros, doc. 194, 61 p., 2015.
- Embrapa. Mapa de solos do Brasil, 2021. Disponível em: <http://geoinfo.cnps.embrapa.br/layers/geonode%3Abrasil_solos_5m_20201104> . Acesso em: <15 de fevereiro de 2021>.
- FAWCETT, T. An introduction to ROC analysis. **Pattern Recognition Letters**, vol. 27, pg. 861–874, 2006.
- FERREIRA, R. O., COSTA, O. D. V., SOUZA, L., JACOMINE, P. K. T. Intermittent pond areas in coastal tablelands of the Reconcavo Region of Bahia, Brazil: genesis of areas, characterization, and soil classification. **Revista Brasileira de Ciência do Solo**, vol. 39, n. 6, pg. 1513-1523, 2015.
- GEETA R., DAVALOS L.M., LEVY A., BOHS L., LAVIN M., MUMMENHOFF K., SINHA N., WOJCIECHOWSKI M.F. Keeping it simple: flowering plants tend to retain, and revert to, simple leaves. **New Phytol.**, vol. 193, pg. 481–493, 2012.
- GLASSOP, D., RAE, A., BONNETT, G. Sugarcane flowering genes and pathways in relation to vegetative regression. **Sugar Tech**, vol. 16, pg. 235–240, 2014.
- GLASSOP, D., RAE, A.L. Expression of sugarcane genes associated with perception of photoperiod and floral induction reveals cycling over a 24-hour period. **Functional Plant Biol.**, vol. 46, pg. 314–327, 2019.
- GOODWIN, B.J., MCALLISTER, A.J., FAHRIG, L. Predicting invasiveness of plant species based on biological information. **Conservation Biology**, vol. 13, pg. 422–426, 1998.

- GOSNELL, J.M. Some factors affecting flowering in sugarcane. **Proceedings of the South African Sugar Technologists Association**, pg. 144–147, 1973.
- GREGO, C.R., DE OLIVEIRA, A., NOGUEIRA, S.F., RODRIGUES, C.A.G., BRANCALIÃO, S.R., FURTADO, A. L.S. Estoque de carbono no solo e produtividade da cana-de-açúcar analisados quanto a variabilidade espacial. *In*: INAMASU, R. Y., NAIME, J. DE M., RESENDE, A. V. DE, BASSOI, L. H., BERNARDI, A. C. DE C. **Agricultura de precisão: um novo olhar**. São Carlos (SP), Embrapa Instrumentação, 2011, pg. 240-244.
- GUJARATI, D. M, PORTER, D.C. Capítulo 15: Modelos de regressão de resposta Qualitativa. *In*:_____. **Econometria básica**, 5ª edição, São Paulo: Campus/Elsevier, 2011, pg. 538-586.
- GUTIERREZ, P. A., LOPEZ-GRANADOS, F., JURADO-EXPOSITO, J. M. P. M., HERVAS-MARTINEZ, C. Logistic regression product-unit neural networks for mapping *Ridolfia segetum* infestations in sunflower crop using multitemporal remote sensed data. **Computers and Electronics in Agriculture**, vol. 64, pg. 293–306, 2008.
- HAIR, J.F., WILLIAM, C.B., BARRY, J.B., RONALD L.T. **Análise Multivariada de Dados**. 6a ed. São Paulo: Bookman. 2009.
- HANAKI, K., PORTUGAL-PEREIRA, J. Chapter 6: The Effect of Biofuel Production on Greenhouse Gas Emission Reductions. *In*: K. Takeuchi et al. (eds.). **Biofuels and Sustainability, Science for Sustainable Societies**. Springer Open. 2019. pg. 53-74.
- HOFFMANN, H. P., SANTOS, E.G.D., BASSINELLO, A.I., VIEIRA, M.A.S. **Variedades RB de Cana-de-açúcar**. CCA/UFSCar. Araras (SP), 1.ed, 30p. 2008.
- HOSMER, D.W., LEMESCHOW, S. **Applied Logistic Regression**. 2 ed. New York: Wiley. 2000.
- HUMBERT, R.P. **El cultivo de la caña de azúcar**. Tradução de Alfonso Gonzalez Gallardo. México: Compañía Editorial Continental, 719 p., 1974.

- IBGE. Geociências - Downloads, 2020a. Disponível em: <<https://www.ibge.gov.br/geociencias/downloads-geociencias.html>>. Acesso em: <20 de setembro de 2020>.
- IBGE. Produção agrícola municipal, 2020b. Disponível em: <<https://sidra.ibge.gov.br/tabela/5457>>. Acesso em: <20 de novembro de 2020>.
- IBGE. Geomorfologia 1:250.000, c2021. Disponível em: <<https://www.ibge.gov.br/geociencias/informacoes-ambientais/geomorfologia/10870-geomorfologia.html?=&t=downloads>>. Acesso em: <15 de fevereiro de 2021>.
- INMET. Balanço hídrico climatológico mensal, 2020. Disponível em: <<http://sisdagro.inmet.gov.br/sisdagro/app/climatologia/bhclimatologicomensal/index>>. Acesso em: <25 de outubro de 2020>.
- IPCC. AR6 Climate Change 2021: The Physical Science Basis, 2021. Disponível em: <<https://www.ipcc.ch/report/ar6/wg1/>>. Acesso em: <26 de março de 2021>.
- ISLER, P.R., RAMOS, R.P., FLORENTINO, H.O., CANTANE, D.R. Otimização aplicada ao planejamento de plantio e colheita da cana-de-açúcar utilizando graus-dias. XLVI Simpósio Brasileiro De Pesquisa Operacional: Pesquisa Operacional na Gestão da Segurança Pública, Salvador (BA), 2014.
- LILHARE, R., D'ERY, S.J., POKORNY, S., STADNYK, T.A., KOENIG, K.A. Intercomparison of multiple hydroclimatic datasets across the lower Nelson River Basin, Manitoba, Canada. **Atmosphere-Ocean**, vol. 57, n. 4, pg. 262-278, 2019.
- LIMA NETO, J. A., RIBEIRO, M. R., CORRÊA, M. M., SOUZA JÚNIOR, V. S., LIMA, J. F. W.F., FERREIRA, R. F. A. L. Characterization and genesis of cohesive horizons in oxisols and ultisols of the coastal tablelands of Alagoas State, Brazil. **Revista Brasileira de Ciência do Solo**, vol. 33, n. 4, pg. 1001-1011, 2009.
- LIMA, L.G., BARBOSA, J.L. From the monoculture of sugar cane to eucalyptus monoculture, two sides of the same tragedy: considerations about the wood-paper-cellulose complex in Alagoas. **Revista Nera**, n. 59, pg. 133-151, 2021.

- MACEDO, I.C., LEAL, M.R.L.V., DA SILVA, J.E.A.R..Assessment of greenhouse gas emissions in the production and use of fuel ethanol in Brazil. Government of the State of São Paulo - **Secretariat of the Environment**, pg. 1-32, 2004.
- MAIA, F. C. O. **Curva de retenção e capacidade de água disponível em latossolos**. Trabalho de conclusão de curso (Bacharelado em Agronomia) - Universidade de Brasília, Brasília, 52 f. 2016.
- MALAKAR, P., KESARKAR, A.P., BHATE, J.N., SINGH, V., DESHAMUKHYA, A. Comparison of reanalysis data sets to comprehend the evolution of tropical cyclones over North Indian Ocean. **Earth Sp. Sci.**, vol. 7, pg. 1–15, 2020.
- MARENGO, J.A., NOBRE, C.A., SELUCHI, M.E., CUARTAS, A., ALVES, L.M., MENDIONDO, E.M., OBREGON, G., SAMPAIO, G. A seca e a crise hídrica de 2014-2015 em São Paulo. **Revista USP**, São Paulo (SP), n.106, pg.31-44, 2015.
- MARIN, F. NASSIF, D.S.P. Mudanças climáticas e a cana-de-açúcar no Brasil: Fisiologia, conjuntura e cenário futuro. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande (PB), vol. 17, n. 2, pg. 232–239, 2013.
- MARIN, F.R, PELLEGRINO, G.Q., ASSAD, E.D., PINTO, H.S., ZULLO JUNIOR, J. Cap. 7 – Cana de açúcar. In: Monteiro, J.E.B.A (Org.). **Agrometeorologia dos Cultivos: O fator meteorológico na produção agrícola**. Brasília, DF: INMET, 2009, pg. 109-131.
- MEDEIROS, A.L.M. **Caracterização e comparação de genes expressos em ápices meristemáticos de cana-de-açúcar cultivados em SP e RN**. Dissertação (Mestrado em Bioquímica; Biologia Molecular) - Universidade Federal do Rio Grande do Norte, Natal, 63 f. 2010.
- MELLO, R. Y., KOHLS, W., OLIVEIRA, T. M. N. Use of different methods of gap filling in the rainfall stations. **Boletim de Geografia**, vol. 35, pg. 112-121, 2017.
- NAGELKERKE, N.J.D. A note on a general definition of the coefficient of determination. **Biometrika**, vol. 78, pg. 691–692, 1991.

- NASCIMENTO, R. C., SANTOS, F. A. A. Evolução da cultura da cana-de-açúcar no Estado de Alagoas em relação aos maiores produtores do Brasil, 1980/1996. **Estação Científica**, Juiz de Fora (MG), vol. 1, n. 2, pg. 1-21, 2005.
- NOAA. NCEP/NCAR Reanalysis 1: Summary, c2019. Disponível em: <<https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html>>. Acesso em: <15 de fevereiro de 2019>.
- OGTR. The Biology of the *Saccharum spp.* (Sugarcane), 2011. Disponível em: <<http://www.ogtr.gov.au>>. Acesso em: <26 de agosto de 2019>.
- OHIGASHI, K., MIZUGUTI, A., NAKATANI, K., YOSHIMURA, Y., MATSUO, K. Modeling the flowering sensitivity of five accessions of wild soybean (*Glycine soja*) to temperature and photoperiod, and its latitudinal cline. **Breed. Sci.**, vol. 69, pg. 84–93, 2019.
- OLIVEIRA, G. P., ANGELOTTI-MENDONÇA, J., TANAKA, F. A. O., DA SILVA, S. R., SCARPARE FILHO, J. A. Origin and development of reproductive buds in jaboticaba cv. Sabará (*Plinia jaboricaba Vell*). **Sci. Horticulturae**, vol. 249, pg. 432–438, 2019.
- PARRA-TABLA V, VARGAS CF, MAGAÑA-RUEDA S, NAVARRO J. Female and male pollination success of *Oncidium ascendens* Lindey (Orchidaceae) in two contrasting habitat patches: forests vs. agricultural field. **Biological Conservation**, vol. 94, pg. 335–340, 2000.
- PATIL, S.B., GUDDADAMATH, S.G., KHADI, B.M. Genetic Enhancement of Sugarcane Productivity Combining Non Flowering Feature. **Sugar Tech**, vol. 17, n. 4, pg. 386–394, 2015.
- PEREIRA, A.R. Previsão do florescimento em cana -de-açúcar. **Comunicação da Pesquisa Agropecuária**, vol. 3, n. 6, p. 15-16, 1985.
- PEREIRA, A.R., BARBIERI, V., VILLA NOVA, N.A. CLIMATIC CONDITIONING OF FLOWERING INDUCTION IN SUGARCANE. **Agricultural Meteorology**, vol. 29, pg. 103-110, 1983.

- PORFIRIO, A.C.S., CEBALLOS, J.C., BRITTO, J.M.S., COSTA, S.M.S. Evaluation of global solar irradiance estimates from GL1.2 satellite-based model over Brazil using an extended radiometric network. **Remote Sens.**, vol.12, n. 8, 20 p., 2020.
- RAMPAZO, N. A. M., PICOLI, M. C. A., CAVALIERO, C. K. N. Comparison of meteorological data obtained by remote sensing (from models and satellites) and weather stations data. **Revista Brasileira de Geografia Física**, vol. 12, n. 2, pg. 412-426, 2019.
- RIDESA - Rede Interuniversitária para o Desenvolvimento do Setor Sucroalcooleiro. Catálogo nacional de variedades “RB” de cana-de-açúcar. Curitiba: RIDESA, 83p. 2010.
- SAHA, S., MOORTHI, S., PAN, H., WU, X., WANG, J. NADIGA, S., TRIPP, P., KISTLER, R., WOOLLEN, J., BEHRINGER, D., LIU, H., STOKES, D., GRUMBINE, R., GAYNO, G., WANG, J., HOU, Y., CHUANG, H., JUANG, H.H., SELA, J., IREDELL, M., TREADON, R., KLEIST, D., DELST, P.V., KEYSER, D., DERBER, J., EK, M., MENG, J., WEI, H., YANG, R., LORD, S., DOOL, H.V.D., KUMAR, A., WANG, W., LONG, C., CHELLIAH, M., XUE, Y., HUANG, B., SCHEMM, K., EBISUZAKI, W., LIN, R., XIE, P., CHEN, M., ZHOU, S., HIGGINS, W., ZOU, C., LIU, Q., CHEN, Y., HAN, Y., CUCURULL, L., REYNOLDS, R.W., RUTLEDGE, G., GOLDBERG, M. The NCEP climate forecast system reanalysis. **Bulletin of the American Meteorological Society**, vol. 91, n. 8, pg. 1015-1058, 2010.
- SANTOS, P.H.S. **Seleção e validação de genes de referência em condições de fotoperíodo artificial para indução do florescimento em cana-de-açúcar.** Dissertação (Mestrado em Agronomia) - Faculdade de Ciências Agrárias e Veterinárias – UNESP, Jaboticabal, 24f. 2020.
- SEGATO, S. V.; PEREIRA, L. L. Cap. 19 - Colheita da Cana-de-Açúcar: corte manual. In: SEGATO, S. V. et al (org.). **Atualização em Produção de Cana-de-Açúcar.** Piracicaba: CP 2, 2006, pg. 319-332.
- SENTELHAS, P.C., ANGELOCCI, L.R. Balanço Hídrico Climatológico Normal e Sequencial, de Cultura e para Manejo da Irrigação. LCE 306: Meteorologia Agrícola - Aula: 9. ESALQ/USP, 2012. Disponível em:

<http://www.leb.esalq.usp.br/leb/aulas/lce306/Aula9_2012.pdf>. Acesso em: <10 de outubro de 2020>.

- SENTELHAS, P.C., GILLESPIE, T.J., SANTOS, E.A. Evaluation of FAO Penman–Monteith and alternative methods for estimating reference evapotranspiration with missing data in Southern Ontario, Canadá. **Agricultural Water Management**, vol. 97, n. 5, pg. 635-644, 2010.
- SHAPIRO, S.S., WILK, M. An analysis of variance test for normality (complete samples). **Biometrika**, vol. 52, pg. 591–611, 1965.
- SILVEIRA, C.S., COSTA, A.A., COUTINHO, M.M., FILHO, F.A.S., JÚNIOR, F.C.V., NORONHA, A.W. Verificação das previsões de tempo para chuva usando ensemble regional para o estado do Ceará em 2009. **Revista Brasileira de Meteorologia**, vol.26, n.4, pg. 609- 618, 2011.
- Sindaçúcar-AL. Desempenho Agroindustrial de variedades de Cana de Açúcar 2007/2008, 2009. Disponível em: <<http://www.sindacucaral.com.br/variedades/>>. Acesso em: <20 de setembro de 2019>.
- Sindaçúcar-AL. Safra 20/21 aqueceu oferta de postos de trabalho em setembro, c2021. Disponível em: <<http://www.sindacucar-al.com.br/2020/11/safra-2021-aqueceu-oferta-de-postos-de-trabalho-em-setembro/>>. Acesso em: <20 de abril de 2021>.
- Sindaçúcar-AL. Sindaçúcar-AL estima safra com até 17,5 mi de toneladas de cana processada, 2021. Disponível em: < <http://www.sindacucar-al.com.br/2021/02/sindacucar-al-estima-safra-com-ate-175-mi-de-toneladas-de-cana-processadas/> >. Acesso em: <20 de abril de 2021>.
- Socicana. Características Agronômicas das Variedades SP, c2020. Disponível em: <<http://socicana.com.br/2.0/wp-content/uploads/Variedades-SP.pdf>>. Acesso em: <9 de outubro de 2020>.
- SOUSA, I.A.L. **Caracterização de um homólogo de HINT1 em cana-de-açúcar encontrado em bibliotecas subtrativas de CDNA para floração**. Dissertação

(Mestrado em Bioquímica) - Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, 78f. 2015.

- SOUZA, V. D. S. **Análise de cDNAs identificados em bibliotecas substrativas de cDNA para floração de variedades de cana-de-açúcar cultivadas no Rio Grande do Norte (RN): fator de transição NAC, Calmodulina e Fosfatidiltransferase.** Dissertação (Mestrado em Biodiversidade; Biologia Estrutural e Funcional.) - Universidade Federal do Rio Grande do Norte, Natal, 80f. 2011.
- SRIVASTAVA, R.P., SINGH, S.P., SINGH, P., SINGH, S.B. Artificial Induction of Flowering in Sugarcane Under Sub-tropical Conditions - A Successful Approach. **Sugar Tech**, vol. 8, pg. 184-186, 2006.
- SRTM. Download Manager, c2021. Disponível em: <<https://srtm.csi.cgiar.org/srtmdata/>>. Acesso em: <15 de fevereiro de 2021>.
- STEINBERG, D., COLLA, P. **Logistic Regression.** SYSTAT, ver. 10, SPSS Inc, Chicago, 2000.
- THORNTHWAITE, C.W., MATHER, J.R. Instructions and tables for computing potencial evapotranspiration and the water balance. **Publications in Climatology**, vol. 10, n. 3, 311p., 1957.
- UDOP. Características Agronômicas das Variedades IAC, c2020. Disponível em: <http://www.portaludop.com.br/index.php?item=variedades_iac>. Acesso em: <9 de outubro de 2020>.
- WALLACH, D., HWANG, C., CORRELL, M. J., JONES, J. W., BOOTE, K., HOOGENBOOM, G., GEZAN, S.A., BHAKTA, M. A dynamic model with QTL covariables for predicting flowering time of common bean (*Phaseolus vulgaris*) genotypes. **Eur. J. Agron.**, vol. 101, pg. 200–209, 2018.
- WANG, X., CHEN, D., PANG, G., OU, T., YANG, M., WANG, M. A climatology of surface–air temperature difference over the Tibetan Plateau: Results from multi-source reanalysis. **International Journal of Climatology**, vol. 40, pg. 1–15, 2020.

- WELLS, T.C.E., ROTHERY, P., COX, R., BAMFORD, S. Flowering dynamics of *Orchis morio* L. and *Herminium monorchis* (L.) R.Br. at two sites in eastern England. **Botanical Journal of the Linnaean Society**, vol. 126, pg. 39–48, 1998.
- XAVIER, A. C., KING, C. W., SCANLON, B. R. Daily gridded meteorological variables in Brazil (1980-2013). **International Journal of Climatology**, vol. 36, pg. 2644–2659, 2016.
- ZHOU, C., WANG, K., MA, Q. Evaluation of eight current reanalysis in simulating land surface temperature from 1979 to 2003 in China. **International Journal of Climatology**, vol. 30, pg. 7379–7398, 2017.
- ZHOU, Z., LIN, A., WANG, L., QIN, W., ZHONG, Y., HE, L. Trends in downward surface shortwave radiation from multi-source data over China during 1984–2015. **International Journal of Climatology**, vol. 40, pg. 1–19, 2019.
- ZULUAGA, C.F., AVILA-DIAZ, A., JUSTINO, F.B., WILSON, A.B. Climatology and trends of downward shortwave radiation over Brazil. **Atmospheric Research**, vol. 250, 18 p., 2021.

ANEXO 1 – ROTEIRO DE CÁLCULO DE SALDO DE RADIAÇÃO (RN) DIÁRIO

1) Para o cálculo da radiação no topo da atmosfera é preciso calcular antes:

- . Correção diária da distância Terra-Sol: $(\bar{D}/D)^2$
- . Declinação Solar (δ)
- . Ângulo Horário do Nascer do Sol (H)

-Cálculo de $(\bar{D}/D)^2$, usando a equação de Spencer:

$$(\bar{D}/D)^2 = 1,000110 + 0,034221 \cos x + 0,001280 \sin x + 0,000719 \cos 2x + 0,000077 \sin 2x$$

Sendo x, em radianos:

$$x = \frac{2\pi(nj-1)}{365}$$

-Cálculo da declinação solar diária (δ):

$$\delta = 23,45 \sin \left[\frac{360}{365} (nj + 284) \right]$$

-Cálculo do Ângulo Horário do Nascer do Sol (H):

$$H = \arccos(-\tan \phi \cdot \tan \delta)$$

2) Cálculo da irradiância solar diária no topo da atmosfera (R_0):

$$R_0 = 37,6 \cdot (\bar{D}/D)^2 \cdot [(H \cdot \pi/180) \sin \phi \cdot \sin \delta + \cos \phi \cdot \cos \delta \cdot \sin H]$$

3) Cálculo da Irradiância global diária (R_g), pela equação de Angström-Prescott:

$$R_g = \left[a + b \left(\frac{n}{N} \right) \right]$$

$$\text{Sendo } \begin{cases} a = 0,29 \cdot \cos \phi \\ b = 0,52 \\ N = \frac{2H}{15} \end{cases}$$

4) Cálculo da pressão de vapor d'água (e), balanço de ondas longas (BOL) e balanço de ondas curtas (BOC):

$$e_s = 6,1078 \cdot 10^{\left(\frac{7,5 \cdot t_{med}}{237,3 + t_{med}} \right)}$$

$$UR = \frac{e}{e_s}$$

$$BOC = R_g \cdot (1 - \alpha)$$

$$BOL = 4,9 \cdot 10^{-9} \cdot T_{med}^4 \cdot (0,1 \sqrt{e} - 0,56) \cdot \left[0,1 + 0,9 \left(\frac{n}{N} \right) \right]$$

5) Cálculo do saldo da Radiação à superfície (R_n):

$$R_n = BOC + BOL$$

PARÂMETRO	SIGNIFICADO	UNIDADE DE MEDIDA
$(\bar{D}/D)^2$	Correção diária da distância Terra-Sol	-
a	Coefficiente da equação de Angström-Prescott	-
b	Coefficiente médio da equação de Angström-Prescott	-
BOC	Balanço de ondas curtas	MJ.m ⁻² .dia ⁻¹
BOL	Balanço de ondas longas	MJ.m ⁻² .dia ⁻¹
e	Pressão real do vapor d'água no ar	hPa
e_s	Pressão de saturação do vapor d'água no ar	hPa
H	Ângulo Horário do Nascer do Sol	(°)
N	Período de brilho solar teórico	h
n_j	Número que representa a posição ordinal de determinado dia do ano, por exemplo, dia 23 de setembro possui $n_j=266$ por ser o 266° dia do ano.	-
R_0	Irradiância solar diária no topo da atmosfera	MJ.m ⁻² .dia ⁻¹
R_g	Irradiância global diária	MJ.m ⁻² .dia ⁻¹
R_n	Saldo de radiação à superfície	MJ.m ⁻² .dia ⁻¹
T_{med}	Temperatura média do ar em graus Kelvin	K
t_{med}	Temperatura média do ar em Celsius	(°)
UR	Umidade Relativa	%
x	Ângulo diário para correção da distância Terra-Sol:	rad
α	Albedo	%
δ	Declinação Solar	(°)
ϕ	Latitude local	(°)
η	Período de brilho solar efetivo	h

ANEXO 2 – ROTEIRO DE CÁLCULO DE BALANÇO HÍDRICO SEQUENCIAL DIÁRIO

1) Cálculo da CAD (Doorenbos e Kassam, 1994):

$$CAD = CAD_{m\u00e9dia} * Z_r$$

Em que:

Z_r , para cana-de-a\u00e7\u00facar, foi adotado o valor de 80 cm.

$CAD_{m\u00e9dia}$, para o solo de Barra de S\u00e3o Miguel, foi calculado em 2,14 mm/cm (Equa\u00e7\u00e3o 8).

2) C\u00e1lculo da subtra\u00e7\u00e3o entre chuva e evapotranspira\u00e7\u00e3o, mantendo-se os sinais positivos (+) e negativos (-).

$$P-ET_o$$

3) Determina\u00e7\u00e3o do NAc e do ARM

-Se $(P-ETP) < 0 \rightarrow$ Calcula-se o NAc, ou seja os valores de $(P-ETP)$ s\u00e3o negativos.

Em seguida, calcula-se o valor da ARM. Em que:

$$NAc = CAD * \ln ARM/CAD$$

$$ARM = CAD * e^{-|NAc/CAD|}$$

-Se $(P-ETP) \geq 0 \rightarrow$ Calcula-se primeiro o ARM* e posteriormente calcula-se o NAc. Nesse caso, o NAc deve ser determinado no caso de haver um pr\u00f3ximo per\u00edodo com $(P-ETP) < 0$.

$$ARM^* = ARM_{anterior} + (P-ETP)$$

5) C\u00e1lculo da ALT

$$ALT = \Delta ARM = ARM_i - ARM_{i-1}$$

Em que:

- . ARM_i \u00e9 a observa\u00e7\u00e3o de armazenamento de \u00e1gua no solo atual (dia atual);
- . ARM_{i-1} \u00e9 a observa\u00e7\u00e3o de armazenamento de \u00e1gua no solo anterior (dia anterior);
- . $ALT > 0 \rightarrow$ reposi\u00e7\u00e3o;
- . $ALT < 0 \rightarrow$ retirada de \u00e1gua do solo.

6) Determinação da ETR

-Se $(P-ET_o) < 0 \rightarrow ETR = P + |ALT|$

-Se $(P-ET_o) \geq 0 \rightarrow ETR = ETP$

7) Determinação da DEF

$DEF = ETP - ETR$

8) Determinação do EXC

-Se $ARM < CAD \rightarrow EXC = 0$

-Se $ARM = CAD \rightarrow EXC = (P-ETP) - ALT$

PARÂMETRO	SIGNIFICADO	UNIDADE DE MEDIDA
CAD_{média}	Capacidade de água disponível média, em mm de água / cm de profundidade de solo	mm de água / cm de profundidade de solo
Z_r	Profundidade específica do sistema radicular	cm
P	Chuva diária	mm.dia ⁻¹
ET_o	Evapotranspiração diária	mm.dia ⁻¹
NAc	Negativo acumulado	mm.dia ⁻¹
ARM	Quantidade de água armazenada	mm.dia ⁻¹
ETR	Evapotranspiração Real	mm.dia ⁻¹
DEF	Deficiência hídrica, equivalendo ao quanto o sistema solo-planta deixou de evapotranspirar	mm.dia ⁻¹
EXC	Excedente hídrico, que corresponde à água que não pode ser retida e drena em profundidade	mm.dia ⁻¹
ALT	Varição na água armazenada no solo (ΔARM)	mm.dia ⁻¹
ET_o	Evapotranspiração Potencial ou de referência	mm.dia ⁻¹
CAD	Capacidade de água disponível	mm.dia ⁻¹
NAc	Negativo acumulado	mm.dia ⁻¹