



**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA – UFRA
PROGRAMA DE PÓS-GRADUAÇÃO EM SAÚDE E PRODUÇÃO
ANIMAL NA AMAZÔNIA – PPGSPAA**

VITOR HUGO MAUÉS MACEDO

**AGROMETEOROLOGIA E PRODUÇÃO DOS CAPINS MASSAI E TANZÂNIA NAS
REGIÕES AMAZÔNICA E SEMIÁRIDA DO BRASIL: UMA ABORDAGEM
ESTATÍSTICA**

**BELÉM
2021**

VITOR HUGO MAUÉS MACEDO

**AGROMETEOROLOGIA E PRODUÇÃO DOS CAPINS MASSAI E TANZÂNIA NAS
REGIÕES AMAZÔNICA E SEMIÁRIDA DO BRASIL: UMA ABORDAGEM
ESTATÍSTICA**

Tese de doutorado apresentada à Universidade Federal Rural da Amazônia, como parte das exigências do Programa de Pós-graduação em Saúde e Produção Animal na Amazônia: área de concentração Produção Animal, para obtenção do título de Doutor.

Orientador: Prof. Dr. Aníbal Coutinho do Rêgo - UFRA
Coorientador: Prof. Dr. Magno José Duarte Cândido – UFC

**BELÉM
2021**

Dados Internacionais de Catalogação na Publicação (CIP)
Bibliotecas da Universidade Federal Rural da Amazônia
Gerada automaticamente mediante os dados fornecidos pelo(a) autor(a)

M141a Macedo, Vitor Hugo Maués

Agrometeorologia e produção dos capins Massai e Tanzânia nas regiões amazônica e semiárida do Brasil: uma abordagem estatística / Vitor Hugo Maués Macedo. - 2021.
52 f. : il. color.

Tese (Doutorado) - Programa de Pós-Graduação em Saúde e Produção Animal na Amazônia (PPGSPAA),
Campus Universitário de Belém, Universidade Federal Rural Da Amazônia, Belém, 2021.

Orientador: Prof. Dr. Aníbal Coutinho do Rêgo

Coorientador: Prof. Dr. Magno José Duarte Cândido.

1. Agrometeorologia. 2. Análise multivariada. 3. Forragem. 4. Megathyrus. 5. Modelagem. I.
Rêgo, Aníbal Coutinho do, *orient.* II. Título

VITOR HUGO MAUÉS MACEDO

AGROMETEOROLOGIA E PRODUÇÃO DOS CAPINS MASSAI E TANZÂNIA NAS
REGIÕES AMAZÔNICA E SEMIÁRIDA DO BRASIL: UMA ABORDAGEM
ESTATÍSTICA

Tese apresentada à Universidade Federal Rural da Amazônia, como parte das exigências do Programa de Pós-Graduação em Saúde e Produção Animal na Amazônia: área de concentração Produção Animal, para obtenção do título de Doutor.

Orientador: Prof. Dr. Aníbal Coutinho do Rêgo

Aprovado em 31 de agosto de 2021.

BANCA EXAMINADORA



Dr. Aníbal Coutinho do Rêgo – Orientador
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA – UFRA



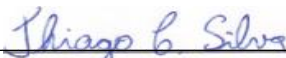
Dr. Márcio André Stefanelli Lara – 1º Examinador
UNIVERSIDADE FEDERAL DE LAVRAS – UFLA



Dr. Marcos Neves Lopes – 2º Examinador
INSTITUTO FEDERAL DO PIAUI – IFPI



Dr. Rodrigo Gregório da Silva – 3º Examinador
INSTITUTO FEDERAL DO CEARÁ – IFCE



Dr. Thiago Carvalho da Silva – 4º Examinador
UNIVERSIDADE FEDERAL RURAL DA AMAZÔNIA – UFRA

RESUMO

A produção de forragem, resulta da ação sistêmica de fatores ambientais que atuam nos ecossistemas de pastagens. O estudo dos processos biofísicos envolvidos e a tomada de decisões a partir de modelos preditivos, deve levar em consideração as condições do meio em que a planta se desenvolve. Na busca do entendimento de como as variáveis ambientais agem na produção de forragem, e obtenção de modelos aplicados a regiões equatoriais, objetivou-se avaliar a produção de biomassa dos cultivares Tanzânia e Massai, em experimentos realizados tanto sob influência do bioma amazônico como do clima semiárido brasileiro, sob a ótica de análises multivariadas e geração de modelos empíricos baseados em aprendizado de máquina. Dados provenientes de seis experimentos: dois com capim Tanzânia e um com capim Massai, realizados no município de Igarapé-Açu-PA; mais dois experimentos: um com capim Tanzânia realizado do município de Pentecoste-CE e outro com capim Massai realizado no município de Fortaleza-CE; foram submetidos a análise de correlação canônica (ACC), função discriminante (FD), e componentes principais (ACP). Utilizou-se, medições do acúmulo de forragem total (AcFT), lâmina foliar (AcLF), hastes (AcH), densidade populacional de perfilhos e incremento em altura do dossel como variáveis agronômicas; e um conjunto de variáveis agrometeorológicas. Os dados de experimentos realizados na região amazônica foram utilizados para a geração de modelos por meio de métodos aprendizado de máquina, posteriormente validados com dados provenientes da região semiárida. O AcFT foi considerado variável resposta, enquanto que as variáveis agrometeorológicas e combinações com variáveis agronômicas foram consideradas variáveis independentes. Os métodos estatísticos multivariados foram capazes de caracterizar os experimentos sob a ótica de variabilidade de condições ambientais e de cultivares estudadas. Essa caracterização demonstrou uma maior variabilidade dos experimentos realizados na região norte em relação aos experimentos realizados na região nordeste. A análise de correlação canônica pode correlacionar bem variáveis agrometeorológicas e variáveis produtivas. E os experimentos foram separados por duas funções discriminantes, uma que discriminou os experimentos quanto as regiões e outra quanto aos cultivares. Quanto a modelagem de previsão de forragem, métodos de aprendizado de máquina em particular o método de regressão por mínimos quadrados parciais foi capaz de gerar modelos de previsão a partir de dados da região amazônica para predição de produção de biomassa de forragem na região semiárida.

Palavras-chave: Agrometeorologia. Análise multivariada. Forragem. *Megathyrsus*. Modelagem. Ruminantes.

ABSTRACT

Forage production results from the systemic action of environmental factors that act on pasture ecosystems. The study of the biophysical processes involved and the decision making based on predictive models, must take into account the conditions of the environment in which the plant grows. In the search to understand how environmental variables act in the production of forage, and to obtain models applied to equatorial regions, the objective in this study was to evaluate the biomass production of the cultivars Guinea and Massai, in studies carried out under the influence of both the Amazon biome and the Brazilian semiarid climate, from the perspective of multivariate analysis and generation of empirical models. Data from six experiments: two with Guinea grass and one with Massai grass, carried out in Igarapé-Açu - PA; and two other experiments: two with Tanzânia grass carried out in Pentecoste - CE and the other with Massai grass in Fortaleza - CE; were subjected to analyses of canonical correlation (CCA), discriminant function (DFA), and principal components (PCA). Records of accumulation of total forage (TFAc), leaf blade (LBAc), stems (SAc), tiler population density, and canopy height were used as agronomical variables; and a set of agrometeorological variables. Data from experiments carried out in the Amazon region were used to generate models using machine learning methods, subsequently validated with data from the semiarid region. The TFAc was considered the response variable, whereas the agrometeorological variables were considered the independent variables. Multivariate statistical methods were able to characterize the experiments from the perspective of the variability of environmental conditions and cultivars studied. This characterization demonstrated a greater variability of the experiments carried out in the north region than those carried out in the northeast region. The canonical correlation analysis can correlate well with the agrometeorological variables and productive variables. And the experiments were separated by two discriminate functions, one that discriminated the experiments in terms of regions and the other in terms of cultivars. As for forage prediction modeling, machine learning methods, particularly the partial least squares regression method, were able to generate prediction models from data from the Amazon region to predict forage biomass production in the semiarid region.

Keywords: Agrometeorology. Forage. Modeling. Multivariate Analysis. *Megathyrsus*. Ruminants.

SUMÁRIO

1. CONTEXTUALIZAÇÃO	6
2. GRAMÍNEAS DA ESPÉCIE <i>Megathyrsus maximus</i> E POTENCIAIS USOS PARA ESTUDO ENVOLVENDO FERRAMENTAS ESTATÍSTICAS	7
2.1 Introdução	8
2.2 A espécie <i>Megathyrsus maximus</i> (syn <i>Panicum maximum</i>)	9
2.3 Estudos de relação entre variáveis climáticas e produtivas: modelagem, métodos multivariados e aprendizado de máquina	10
2.4 Considerações finais	13
2.5 Referências	13
3. RELAÇÃO ENTRE VARIÁVEIS AGROMETEOROLÓGICAS E AGRONÔMICAS DE CAPINS DO GÊNERO <i>Megathyrsus</i> CULTIVADOS NA REGIÃO NORTE E NORDESTE DO BRASIL: UMA ABORDAGEM MULTIVARIADA	16
3.1 Introdução	17
3.2 Material e métodos	19
3.2.1 Locais experimentais.....	19
3.2.2 Descrição dos Experimentos	21
3.2.3 Variáveis agronômicas mensuradas	22
3.2.4 Variáveis agrometeorológicas	23
3.2.5 Análises estatísticas	24
3.3 Resultados	24
3.3.1 Como que os diferentes experimentos se caracterizam em relação aos índices que agrupam variáveis agronômicas e agrometeorológicas?	24
3.3.2 Como que se relacionam as variáveis agrometeorológicas com as variáveis agronômicas?.....	26
3.3.3 Como podem ser descritas funções discriminantes que separam experimentos com capim Tanzânia e Massai conduzidos em regiões distintas mas de latitude semelhante?.....	27
3.4 Discussão	28
3.5 Conclusão	32
3.6 Referências	33
4. O USO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA MELHORA A PREDIÇÃO DA BIOMASSA DE FORRAGEM EM REGIÕES COM DIFERENTES CONDIÇÕES AMBIENTAIS?	36
4.1 Introdução	37
4.2 Material e Métodos	38
4.2.1 Dados de geração dos modelos	38
4.2.2 Dados de validação dos modelos	39
4.2.3 Variáveis agronômicas utilizadas na parametrização dos modelos	39
4.2.4 Variáveis agrometeorológicas utilizadas na parametrização dos modelos	40
4.2.5 Análises estatísticas.....	42
4.3 Resultados	43
4.4 Discussão	48
4.5 Conclusão	50
4.6 Referências	50

1. CONTEXTUALIZAÇÃO

A relação entre a produção de forragem e variáveis climáticas é de extrema importância. Grande parte dos estudos que envolvem essa relação são baseados em estudos de modelagem de produção. As aplicações da modelagem são diversas como: auxiliar no planejamento da produção e orçamentação forrageira; compreensão dos fatores que levam a respostas na produção de forragem; simulação de experimentos; zoneamento agroclimático de regiões e propriedades; e estudos de cenários futuros e indicação de possíveis impactos em sistemas de produção.

Hoje métodos de modelagem são aplicados utilizando principalmente modelos empíricos e mecanicistas. Os modelos mecanicistas são mais robustos porém demandam de informações mais detalhadas para os cenários de modelagem. Já os modelos empíricos são mais simples, porém possuem baixa capacidade de generalização em diferentes locais. Contudo, supostamente, métodos avançados de modelagem empírica, baseados em aprendizado de máquina podem ser capazes de melhorar esse processo de generalização dos modelos. Mas antes de realizar a geração dos modelos, é valoroso entender como que essas variáveis provenientes de um sistema complexo se relacionam, e como elas podem caracterizar as observações a campo por meio de métodos multivariados.

A descrição das relações entre essas variáveis por meio de métodos estatísticos multivariados pode fornecer informações importantes para o processo de modelagem da produção de forragem. E dessa forma, é possível hipotetizar que essa análise mais aprofundada em conjunto com o uso de algoritmos de aprendizado de máquina gera modelos de previsão de forragem da espécie *Megathyrus maximus* com melhor capacidade de generalização.

2. GRAMÍNEAS DA ESPÉCIE *Megathyrsus maximus* E POTENCIAIS USOS PARA ESTUDO ENVOLVENDO FERRAMENTAS ESTATÍSTICAS

RESUMO

A relação entre a produção de forragem e variáveis climáticas é de extrema importância. Grande parte dos estudos que envolvem essa relação são baseados em estudos de modelagem de produção. As aplicações da modelagem são diversas como: auxiliar no planejamento da produção e orçamentação forrageira; compreensão dos fatores que levam a respostas na produção de forragem; simulação de experimentos; zoneamento agroclimático de regiões e propriedades; e estudos de cenários futuros e indicação de possíveis impactos em sistemas de produção. Hoje métodos de modelagem são aplicados utilizando principalmente modelos empíricos e mecanicistas. Os modelos mecanicistas são mais robustos, porém demandam de informações mais detalhadas para os cenários de modelagem. Já os modelos empíricos são mais simples, porém possuem baixa capacidade de generalização em diferentes locais. Contudo, supostamente, métodos avançados de modelagem empírica, baseados em aprendizado de máquina podem ser capazes de melhorar esse processo de generalização dos modelos. Mas antes de realizar a geração dos modelos, é valoroso entender como que essas variáveis provenientes de um sistema complexo se relacionam, e como elas podem caracterizar as observações a campo por meio de métodos multivariados.

Palavras-chave: Agrometeorologia. Análise multivariada. Forragem. Modelagem. Ruminantes.

ABSTRACT

The relationship between forage production and climatic variables is extremely important. Most of the studies involving this relationship are based on production modeling studies. Modeling applications are diverse, such as: assisting in forage production planning and budgeting; understanding the factors that lead to responses in forage production; simulation of experiments; agroclimatic zoning of regions and properties; and studies of future scenarios and an indication of possible impacts on production systems. Today modeling methods are applied using mainly empirical and mechanistic models. Mechanistic models are more robust but require more detailed information for modeling scenarios. Empirical models, on the other hand, are simpler but have low generalizability in different places. However, supposedly, advanced empirical modeling methods, based on machine learning, may be able to improve this model

generalization process. But before generating the models, it is important to understand how these variables from a complex system are related, and how they can characterize field observations using multivariate methods.

Keywords: Agrometeorology. Multivariate analysis. Forage. Modeling. Ruminants.

2.1 Introdução

Atualmente, estudos que envolvam a adaptabilidade, crescimento e desenvolvimento de gramíneas tropicais estão altamente difundidos na literatura. Esses estudos avaliam a aplicação de diversos tratamentos baseados em manejo (MACEDO et al. 2021) e nutrição (CANTO et al. 2009) de plantas forrageiras de importância econômica para a produção de herbívoros. Porém muitos desses estudos são baseados em análises univariadas.

Uma abordagem importante que vem crescendo muito nos últimos anos é a utilização de técnicas de modelagem de produção de forragem, que auxiliam na previsão de produção de biomassa para planejamentos e orçamentação forrageira. Grande parte desses estudos são baseados em trabalhos de modelagem mecanicista. E os poucos trabalhos com modelos empíricos usam métodos de análise pouco avançadas diante das disponibilidades de métodos estatísticos mais robustos disponíveis atualmente.

Os estudos com plantas forrageiras são aplicados em espécies que apresentam certo potencial produtivo. Dentre elas destacam-se as gramíneas da espécie *Megathyrsus maximus* (Jacq.) BK Simon e Jacobs. Essa espécie é um valioso capim de pasto reconhecido por altos rendimentos de ganho para a produção animal (MACIEL et al., 2018). Os principais fatores abióticos limitantes da faixa potencial da espécie *Megathyrsus maximus* podem ser as temperaturas mínimas e a precipitação (RHODES et al., 2021). Portanto é importante que estudos levem em consideração uma análise da relação existente entre o cultivo e potencial de produção de gramíneas da espécie *Megathyrsus maximus* com as variabilidades climáticas existentes em diferentes regiões.

Esses estudos podem levar em consideração primeiro uma caracterização baseada em abordagens multivariadas. Pois esses métodos de análise são valiosos para o entendimento de como um grande conjunto de variáveis se relaciona em sistemas complexos (YATER et al. 2015). Um segundo passo seria a aplicação de métodos de modelagem matemática que possam auxiliar no entendimento e previsão da produção de biomassa de gramíneas com potencial produtivo, para fins de planejamento e orçamentação em sistemas de produção de pastagens.

Com isso o presente objetivo dessa revisão é trazer informações importantes acerca de pontos a serem considerados no processo de estudos que levem em conta a utilização de métodos estatísticos para estudo da relação da produção de capins da espécie *Megathyrsus maximus*, e que são apresentados nos próximos capítulo dessa tese.

2.2 A espécie *Megathyrsus maximus* (syn *Panicum maximum*)

A espécie forrageira *Megathyrsus maximus* é conhecida mundialmente pela capacidade produtividade, qualidade e adaptação às diferentes condições climáticas tropicais. A espécie é a forrageira tropical mais produtiva propagada por sementes e tem despertado a atenção de pecuaristas também por pela abundante produção de folhas longas, porte elevado e alta aceitabilidade pelos animais das mais variadas categorias e espécies de ruminantes (FONSECA; MARTUSCELLO, 2011).

Os cultivares da espécie *Megathyrsus maximus* tiveram grande difusão no Brasil Central na década de 80. Posteriormente, nas regiões norte e nordeste, a necessidade de forrageiras com maior potencial produtivo gerou uma busca por forrageiras com essas características, proporcionando a entrada das cultivares como Tanzânia e Massai na região (MACEDO et al. 2021; SILVA et al. 2007; LOPES et al. 2016).

O cultivar Tanzânia foi lançado em 1990 pela EMBRAPA CNPQC, sendo considerada uma gramínea bem adaptada as condições brasileiras, pois apesar de ser exigente em fertilidade do solo respondia a fertilização. Além disso era relativamente adaptada a solos bem drenados e se desenvolvia em regiões com precipitação próximo a 1000 mm ano⁻¹. É uma gramínea com folhas largas e hábito de crescimento cespitoso (GOMES et al. 2011).

O capim Tanzânia é uma gramínea tropical que responde a adubação nitrogenada e, portanto, responde a esse manejo em termos de produção de forragem resultando em melhorias no ganho animal por área. Quando adubado com 100 kg de N ha⁻¹ ano⁻¹, o capim Tanzânia apresenta uma produção média de 6,6 ton de MS ha⁻¹ ano⁻¹, porém, ao se aplicar 400 kg de N ha⁻¹ ano⁻¹, a produção aumenta em aproximadamente 50%. Essa resposta na produtividade influencia diretamente no ganho por área, que chega a aumentar significativamente a magnitude do ganho quando comparado a pastos adubados com apenas 100 kg de N ha⁻¹ ano⁻¹ (CANTO et al. 2009).

Em resposta ao manejo de desfolhação, o capim Tanzânia apresenta características estruturais distintas. Quando a frequência de desfolhação é curta, a estrutura do dossel é alterada, em especial a densidade populacional de perfilhos (DPP), uma vez que a pastagem desfolhada mais frequentemente apresenta maior DPP. Com isso esse dossel atingirá a altura

pré-pastejo mais rapidamente, devido uma maior cobertura vegetal. Além disso, uma maior massa de forragem total é observada nessas pastagens, que possuem ciclos de pastejo mais curtos (MACEDO et al. 2021).

O cultivar Massai é um híbrido de *Megathyrsus* que foi lançado oficialmente pela EMBRAPA em 2001. É caracterizado por ser uma gramínea de porte baixo, com folha finas (aproximadamente 0,85 cm) e um hábito de crescimento cespitoso, mas que proporciona boa cobertura do solo (GOMES et al. 2011). Dentre os cultivares de *Megathyrsus*, o Massai é um capim menos exigente em fertilidade do que o capim Tanzânia, e possui uma exigência em precipitação de 700 mm ano⁻¹ (ALMEIDA et al. 2000). Esse capim tem respostas rápidas a adubação nitrogenada, uma vez que o nível de exigência da cultura é médio. Doses de 150 kg de N ha⁻¹ ano⁻¹ proporcionam incrementos de aproximadamente cinco vezes mais do que quando não adubado, ainda mantendo uma proporção considerável de lâmina foliar (PACIULLO et al. 2016).

Quando manejado sob lotação intermitente em pastejo por novilho, com 7 dias de ocupação e 35 dias de descanso, são observados valores de pré-pastejo com de massa seca total 3,5 t.ha⁻¹ (EUCLIDES et al. 2008). Em relação a exigência hídrica, o capim-Massai necessita de pelo menos 30% em relação a evapotranspiração da planta (MARANHÃO et al. 2021).

2.3 Estudos de relação entre variáveis climáticas e produtivas: modelagem, métodos multivariados e aprendizado de máquina

Atualmente no Brasil foram conduzidos diversos estudos baseados em modelagem de produção de forragem. Pedreira et al. (2011) observaram que o modelo CROPGRO integra bem os aspectos fisiológicos do capim Xaraés, sendo uma ferramenta para previsão de produtividade da gramínea. Em estudo com capim-Tanzânia, que objetivou simular o crescimento da espécie utilizando o modelo CROPGRO, Lara et al. (2012) obtiveram resultados satisfatórios, integrando aspectos fisiológicos ao crescimento da cultura. Pequeno et al. (2014) realizaram a calibração do modelo CROPGRO, para prever o crescimento do capim Marandu em condições de irrigação e sequeiro. Nesse estudo, é visto que a calibração foi eficiente, sendo possível a sua adaptação para o capim Marandu. Em trabalho que objetivou calibrar modelo APSIM, para estimar a produção de capim Marandu em pastagem sob lotação contínua com taxa de lotação variável, em sistemas a pleno sol e sombreados, Gomes et al. (2020) observaram que o modelo se adequa a sistemas de lotação variável, sendo necessário melhorias para prever os efeitos do nitrogênio no pasto. Já em sistemas sombreados, se faz necessário a inclusão de competição por umidade do solo, no modelo, uma vez que é um fato importante em projetos

silvipastoris. Bosi et al. (2020) ao realizarem a calibração do CROPGRO com capim-Piatã, observaram que mesmo obtendo bons resultados de predição de massa de forragem, alguns ajustes são necessários no modelo para simular o efeito da floração no crescimento da pastagem, mudanças no dossel da pastagem em decorrência do pisoteio de animais e competições dentro do sistema silvipastoril.

Muitos trabalhos com modelagem de produção de forragem realizados no Brasil são realizados por meio da calibração de modelos mecanicistas. Os modelos empíricos baseados em análise de regressão possuem a desvantagem de explicar pouco ou praticamente nada dos processos biofísicos envolvidos na relação de variáveis de entrada e saída de uma função, portanto eles têm baixa capacidade de generalização, por isso eles são modelos voltados para a utilização nos locais aos quais os dados para a geração dos modelos foram obtidos (ANDRADE et al. 2016).

Muitos estudos agrícolas continuam a ser abordados dentro de uma perspectiva univariada, com análises estatísticas sendo realizadas em uma variável por vez. No entanto, os sistemas agrícolas são complexos, e compreender a complexidade dos sistemas bióticos e abióticos é mais bem realizado com o pensamento multivariado, permitindo que sejam abordadas questões que de outra forma seriam difíceis de conceber por abordagens univariadas.

Apesar de menos difundidos que os estudos com análises univariadas, muitos estudos com abordagens multivariadas vem sendo realizados. Para aumentar a eficiência de seleção de cultivares de milho Apraku et al. (2011), elucidaram a interrelação entre o rendimento de grãos e outras características agronômicas com variáveis ambientais utilizando análise discriminante canônica (CDA), análise de componentes principais (PCA) e análise de agrupamento. Em um estudo de campo de dois anos Riedell et al. (2007) determinaram como o estresse das plantas causado por pulgões e vírus afetam as características agronômicas, os nutrientes da folha e a clorofila de aveia e trigo por meio de análise de correlação canônica. Aplicando análises de componentes principais Lupwayi et al. (2011), compararam a aplicação de diferentes faixas de N aplicado no semeio sobre a biomassa microbiana do solo e investigou se o aumento da taxa de semeadura da cevada amenizaria quaisquer efeitos negativos.

Em uma abordagem sobre o uso de métodos estatísticos multivariados Yeater et al. (2015), destacaram a importância de reconhecer que tempo e esforço são necessários para entender as relações dos conjuntos de dados multidimensionais, mas as visualizações das relações entre variáveis e as conclusões nos estudos de sistemas complexos pode valer a pena o esforço. Muitos métodos estatísticos multivariados são utilizados atualmente em abordagem computacionais que envolvem o aprendizado de máquina.

Os métodos de aprendizado de máquina foram desenvolvidos para solucionar problemas que demandam o uso de muitos dados (big data) a partir do uso de ferramentas computacionais de alto desempenho. Esses métodos são geralmente classificados em duas categorias principais: a aprendizagem supervisionada e não supervisionada. A aprendizagem supervisionada funciona a partir de informações já fornecidas sobre os dados, é fornecido aos algoritmos informações prévias que auxiliam no processo de aprendizagem da máquina, e dessa forma os processos de regressão e classificação são realizados. Já a aprendizagem não supervisionada é quando o algoritmo de modelagem não possui as informações prévias durante o processo de análise dos dados, o algoritmo faz então as classificações e agrupamentos baseados em similaridade dos dados fornecidos (LIAKOS et al. 2018).

Os métodos de aprendizado supervisionado são usados para resolver principalmente os problemas de classificação e de regressão. Os problemas de regressão são usados basicamente quando o objetivo é prever valores específicos, como dados numéricos. Existem muitos algoritmos de aprendizado de máquina que atuam para resolver problemas de regressão. O mais conhecido e simples deles é o método da regressão linear simples. Porém outros métodos são bem difundidos, como métodos de regressão polinomial, regressão não linear, regressão múltipla e suas variações como regressão por seleção de variáveis, regressão penalizada, regressão por fatores, e regressão baseadas em árvores de decisão. Ransom et al. (2019) usaram métodos de regressão para estimativas de recomendação de doses nitrogenadas para o milho, esse autor demonstrou a capacidade de melhoras de algoritmos de aprendizado de máquina principalmente relacionado ao algoritmo de floresta aleatória. Em estudos com pastagens muitos trabalhos usando métodos de aprendizado de máquina são usados para monitoramento da biomassa acima do solo usando dados de espectrometria (2021).

Os problemas de classificação são usados para prever categorias e rótulos. Essas classes podem ser de natureza binomial ou multinomial. O algoritmo mais difundido de classificação é a regressão logística, e muitos outros algoritmos que também fazem regressão também podem fazer classificação, como os algoritmos baseados em árvores de decisão. Amatya et al. (2016) usou algoritmos de classificação para determinar a colheita de frutos de café baseado em três categorias: frutos colhíveis, não colhíveis e estado de maturação desconsiderado. Maione et al. (2016) trabalharam com a previsão e classificação a origem geográfica de uma amostra de arroz, baseado em análise de vinte componentes químicos na composição de amostras com espectrometria de massa utilizando o algoritmo de floresta aleatória.

2.4 Considerações finais

O uso de técnicas estatísticas multivariadas e de aprendizado de máquina para a descrição, caracterização e previsão de variáveis produtivas de cultura relacionadas com as características ambientais ao quais estão incluídas, é de grande importância para o avanço de tecnologias no setor agropecuário. Dessa forma, é interessante fomentar pesquisas que visem verificar as possíveis influências do clima nas características agronômicas de plantas forrageiras.

2.5 Referências

ALMEIDA AAS, MONTEIRO FA, JANK L. Avaliação de *Panicum maximum* Jacq. para tolerância ao alumínio em solução nutritiva. **Rev. Bras. de Ciên. do Solo**. 2000; 24 (2):339-334.

AMATYA S, KARKEE M, GONGAL A, ZHANG Q, WHITING MD. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting. **Biosystems engineering**. 2016;146:3-15.

ANDRADE AS, SANTOS PM, PEZZOPANE JRM, ARAUJO LC, PEDREIRA BC, PEDREIRA CGS, MARIN FR, LARA MAS. Simulating tropical forage growth and biomass accumulation: An overview of model development and application. **Grass and Forage Science**, 2016; 71(1):54–65.

APRAKU BB, AKINWALE RO, AJALA SO, MENKIR A, FAKOREDE MAB, OYEKUNIE M. Relationships among traits of tropical early maize cultivars in contrasting environments. **Agronomy Journal**. 2011;103(3):717-729.

BOSI C, SENTELHAS PC, PEZZOPANE JRM, SANTOS PM. CROPGRO-Perennial Forage model parameterization for simulating Piatã palisade grass growth in monoculture and in a silvopastoral system. **Agricultural Systems**. 2020;177:102724.

CANTO MW, BONA FILHO A, MORAES A, HOESCHL AR, GASPARINO E. Animal Production in Tanzania grass swards fertilizes with nitrogen. **R. Bras. Zootec**. 2009;38:1176-1182.

CUTRIM JUNIOR JAA, CÂNDIDO MJD, VALENTE BSM, CARNEIRO MSD, CARNEIRO HAV. Características estruturais do dossel de capim-tanzânia submetido a três frequências de desfolhação e dois resíduos pós-pastejo. **R. Bras. de Zootec**. 2011;40:489-497.

EUCLIDES VP, MACEDO MCM, ZIMMER AH, JANK L; OLIVEIRA, M. P. Avaliação dos capins mombaça e massai sob pastejo. **R. Bras. de Zootec**. 2008;37(1):18-26.

FONSECA DM, MARTUSCELLO JA. Plantas Forrageiras. 1st ed. Viçosa: Editora UFV, 537p.

GOMES FJ, BOSI C, PEDRIERA BC, SANTOS PM, PEDREIRA CGS. Parameterization of the APSIM model for simulating palisadegrass growth under continuous stocking in monoculture and in a silvopastoral system. **Agricultural System**. 2020;184:102876.

GOMES RA, LEMPP B, JANK L, CARPEJANI, G. C., MORAIS, M. G. Características anatômicas e morfofisiológicas de lâminas foliares de genótipos de *Panicum maximum*. **Pesq. Agropec. Bras**. 2011; 46(2):205-211.

LARA MAS, PEDREIRA CGS, BOOTE KJ, PEDREIRA BC, MORENO LSB, ALDERMAN PD. Predicting Growth of *Panicum maximum*: An Adaptation of the CROPGRO–Perennial Forage Model. **Agronomy Journal**. 2012;104(3):600-611.

LIAKOS KG, BUSATO P, MOSHOU D, PEARSON S, BOCHTIS D. Machine learning in agriculture: A review. **Sensors**. 2018;18(8):2674.

LOPES MN, CÂNDIDO MJD, POMPEU RCFF, SILVA RG, MORAIS NETO LB, CARNEIRO MSS. Tillering dynamics in massai grass fertilized with nitrogen and grazed by sheep. **Bioscience Journal**. 2016; 32:446–454.

LUPWAYI NZ, CLAYTON GW, O'DONOVAN JT, GRANT CA. Soil microbial response to nitrogen rate and placement and barley seeding rate under no till. **Agronomy Journal**. 2011;103(4):1064-1071.

MACEDO VHM, CUNHA AMQ, CÂNDIDO EP, DOMINGUES FN, SILVA WL, LARA MAS, RÊGO AC. Canopy structural variations affect the relationship between height and light interception in Guinea Grass. **Field Crops Research**. 2021;271:108249.

MACIEL GA, BRAGA GJ, GUIMARÃES JÚNIOR R, RAMOS AKB, CARVALHO MA, FERNANDES FD, FONSECA CEL, JANK, L. Seasonal Liveweight Gain of Beef Cattle on Guineagrass Pastures in the Brazilian Cerrados. **Agronomy journal**. 2018;110:01-8.

MAIONE C, BATISTA BL, CAMPIGLIA AD, BARBOSA JUNIOR F, BARBOSA RM. Classification of geographic origin of rice by data mining and inductively coupled plasma mass spectrometry. **Computers and Electronics in Agriculture**. 2016;121:101-107.

MARANHÃO, SR, POMPEU RCFF, LOPES MN, CÂNDIDO MJD, SOUZA HÁ, CAVALCANTE ACR, FONTINELE RG, ROGÉRIO MCP. Morphophysiology of tropical grasses under different water supply in two growing seasons: II. BRS Massai and BRS Tamani grasses. **Semina: Ciências agrárias**. 2021;42(1):301-318.

MORAIS TG, TEIXEIRA RFM, FIGUEIREDO M, DOMINGOS T. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. **Ecological Indicators**. 2021;130:108081.

PACIULLO DSC, GOMIDE CAM, CASTRO CRT., MAURICIO RM, FERNANDES PB, MORENZ MJF. Morphogenesis, biomass and nutritive value of *Panicum maximum* under different shade levels and fertilizer nitrogen rates. **Grass and Forage Science**. 2016;72:590-600.

PEDREIRA BC, PEDREIRA CGS, BOOTE KJ, LARA MAS, VEREADOR PD. Adapting the CROPGRO perennial forage model to predict growth of *Brachiaria brizantha*. **Field Crops Research**, 2011;120(3):370-379.

PEQUENO DNL, PEDREIRA CGS, BOOTE KJ. Simulating forage production of Marandu palisade grass (*Brachiaria brizantha*) with the CROPGRO-Perennial Forage model. **Crop and Pasture Science**. 2014;65(12):1335-1348.

RANSOM CJ, KITCHEN NR, CAMBERATO JJ, CARTER PR, FERGUSON RB, FERNÁNDEZ FG, FRANZEN DW, LABOSKI, CAM, NAFZIGER ED, SAWYER JE, SHANAHAN JF. Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations. **Computers and Electronics in Agriculture**. 2019;164:104872.

RHODES AC, PLOWES RM, GOOLSBY JA, GASKIN JF, MUSYOKA B, CALATAYUD PA, CRISTOFARO M, GRAHMANN ED, MARTINS DJ, GILBERT LE. The dilemma of Guinea grass (*Megathyrsus maximus*): a valued pasture grass and a highly invasive species. **Biol Invasions**. 2021.

RIEDEL WE, OSBORNE SL, JARADAT AA. Crop mineral nutrient and yield responses to aphids or barley yellow dwarf virus in spring wheat and oat. **Crop Sci**. 2007;47:1553–1560.

SILVA RG, CÂNDIDO MJD, NEIVA, JNM, LÔBO RN, SILVA DS. Características estruturais do dossel de pastagens de capim-tanzânia mantidas sob três períodos de descanso com ovinos. **R. Bras. Zootec**. 2007;36:1255-1265.

YEATER KM, DUKE SE, RIEDEL WE. Multivariate analysis: greater insights into complex systems. *Agronomy Journal*. 2015;107(2):799-810.

3. RELAÇÃO ENTRE VARIÁVEIS AGROMETEOROLÓGICAS E AGRONÔMICAS DE CAPINS DO GÊNERO *Megathyrsus* CULTIVADOS NA REGIÃO NORTE E NORDESTE DO BRASIL: UMA ABORDAGEM MULTIVARIADA

RESUMO

O objetivo deste estudo foi caracterizar dados experimentais obtidos a partir de pesquisas realizadas com os capins Tanzânia e Massai, na região norte e nordeste do Brasil por meio de variáveis agronômicas e agrometeorológicas, usando a abordagem de análises multivariada. Foram utilizados dados obtidos a partir de quatro estudos com capim Tanzânia, e dois com capim Massai. Foram utilizadas as seguintes variáveis agronômicas: acúmulo de forragem total (AcFT), acúmulo de lâmina foliar (AcLF), acúmulo de haste (AcH), acúmulo de material morto (AcMM), densidade populacional de perfilhos (DPP) e incremento em altura (InAlt) do dossel, densidade populacional de perfilhos e dias de rebrotação (DR). As variáveis agrometeorológicas foram: temperatura média, precipitação (Prec), radiação solar (Rs), evapotranspiração real e de referência, nitrogênio fornecido (NF) e índice hídrico. Foi realizada a análise de componentes principais (ACP), correlação canônica (ACC) e função discriminante (FD). A ACP pôde caracterizar as observações com base no ambiente por meio do componente principal que reflete o contraste entre variáveis de disponibilidade hídrica e o NF com variáveis de fornecimento energético. Enquanto a caracterização com base nas variáveis agronômicas se deu por meio do componente que reflete a distinção entre a DPP, AcFT e InAlt, com o AcH e AcMM. A ACC gerou um coeficiente de correlação de 0,84 entre a variável canônica 1 (Can1) e a variável canônica 2 (Can2). A Can1 representou um contraste entre as variáveis relacionadas disponibilidade hídrica mais o NF com as variáveis relacionadas ao fornecimento energético em termos de radiação e temperatura. A Can2 refletiu um contraste entre o AcMM com as demais variáveis (AcFT, AcLF, DPP e InAlt). A função discriminante 1, com 70,36% do poder de separação discriminou os experimentos com base nos locais de estudo. Os grupos de experimentos foram separados, com base no cultivar estudado, por meio da função discriminante 2, com 19,20% de poder de separação. É possível considerar que apesar das semelhanças quanto ao fornecimento de energia em termos de radiação e temperatura, a disponibilidade de água e o fornecimento de nutrientes é um fator determinante que pode auxiliar estudos futuros que envolvam o uso de modelagem aplicada a produção de forragem.

Palavras-chave: agrometeorologia, componentes principais, correlação canônica, função discriminante, métodos multivariados.

ABSTRACT

The objective of this study is to characterize experimental data obtained from research carried out with the Tanzania and Massai grasses, in the North and Northeast regions of Brazil, through agronomic and agrometeorological variables, using a multivariate analysis approach. Data obtained from four studies with Tanzania grass and two with Massai grass were used. The following agronomic variables were used: total forage accumulation (TFAc), leaf blade accumulation (LBAc), stem accumulation (SAc), dead material accumulation (DMAc), tiller population density (TPD) and increase in height (IH), and regrowth days (RD). The agrometeorological variables were: mean temperature, precipitation (Prec), solar radiation (SR), real and reference evapotranspiration, supplied nitrogen (SN), and hydric index. Principal component analysis (PCA), canonical correlation (ACC), and discriminant function (DF) were performed. The PCA was able to characterize the observations based on the environment through the main component that reflects the contrast between hydric availability variables and the NF with energy supply variables. While the characterization based on agronomic variables took place through the component that reflects the distinction between TPD, TFAc, and IH, with SAc and DMAc. ACC generated a correlation coefficient of 0.84 between canonical variable 1 (Can1) and canonical variable 2 (Can2). The Can1 represented a contrast between variables related to hydric availability plus SN with variables related to energy supply in terms of radiation and temperature. The Can2 reflected a contrast between DMAc and the other variables (TFAc, LBAc, TDP, and IH). The discriminant function 1, with 70.36% of the separation power, discriminated the experiments based on the study locations. The experimental groups were separated, based on the cultivar studied, through the discriminant function 2, with 19.20% of separation power. It is possible to consider that despite the similarities in terms of energy supply in terms of radiation and temperature, the availability of water and the supply of nutrients is a determining factor that may help future studies involving the use of modeling applied to forage production.

Key words: agrometeorology, principal components, canonical correlation, discriminant function, multivariate methods

3.1 Introdução

A grande extensão territorial brasileira proporciona condições climáticas contrastantes (ALVARES et al., 2013). Por exemplo, em contraste com as regiões sul e sudeste do país, as

regiões norte e nordeste são caracterizadas por uma estacionalidade que não depende tanto do fotoperíodo e da temperatura (ALVARES et al., 2013). Isso está diretamente ligado ao estudo das influências ambientais sobre o crescimento e desenvolvimento de gramíneas forrageiras altamente responsivas às condições, como as gramíneas do gênero *Megathyrus*.

Atualmente, essa influência ambiental sobre as plantas está sendo muito explorada devido a elaboração de modelos que objetivam entender melhor a relação entre as características ambientais e produtivas de culturas agrícolas nos locais onde elas estão inseridas. Considerando tal ponto, muitos estudos têm sido conduzidos com culturas agrícolas de exploração industrial, como milho (PUNTEL et al., 2018), soja (MARTIN et al., 2012) e mandioca (CADENA et al., 2021). Por outro lado, estudos com esse viés envolvendo o crescimento de plantas forrageiras em pastagens, ainda tem sido restrito.

O entendimento da relação entre as características produtivas de capins tropicais, e os fatores agrometeorológicos obtidos no local em que essas estão inseridas, necessita ser analisado por meio de uma visão sistêmica do ecossistema de pastagens. Para isso, análises que envolvam uma abordagem multivariada dos fatores relacionados às influências climáticas exercidas sobre as plantas, em conjunto com a dinâmica de crescimento e produção de biomassa, podem oferecer importante contribuição para esse processo de entendimento. (YATER et al., 2014)

Mesmo regiões situadas próximas da linha do equador, como as regiões amazônica e semiárida, e que teoricamente estão sob a mesma influência climática em termos de radiação e temperatura, supostamente possuem fatores climáticos que possam distingui-las, como por exemplo, variáveis ligadas a umidade ao regime hídrico. Dessa forma, a existência de variáveis que representam a influência de fatores agrometeorológicos, e que estejam relacionadas a fatores produtivos, podem caracterizar a distinção entre a produção de espécies do mesmo gênero cultivadas nessas regiões.

Nesse contexto, o objetivo desse estudo foi caracterizar experimentos realizados em regiões distintas do Brasil e responder as seguintes perguntas: Como que se relacionam as variáveis agrometeorológicas com as variáveis agronômicas medidas nesses experimentos? Como podem ser descritas funções discriminantes que possam distinguir os experimentos quanto as espécies de capim Tanzânia e Massai nas regiões amazônica e semiárida do Brasil?

3.2 Material e métodos

3.2.1 Locais experimentais

Foram utilizados dados de quatro experimentos com capim Tanzânia [*Megathyrsus maximus* (syn. *Panicum maximum*) (Jacq.) B.K.Simon & S.W.L.Jacobs cv. Tanzânia], dois ensaios desses experimentos foram conduzidos no município de Igarapé-Açu (01°07'S, 47°36'O, 47m), estado do Pará, e dois em Pentecoste (03°48'S, 49°19'O, 71m), estado do Ceará. Outros dois experimentos com capim Massai [*Megathyrsus maximus* x *Megathyrsus infestus* (Peters) B.K.Simon & S.W.L.Jacobs cv. Massai], foram conduzidos em Igarapé-Açu e em Fortaleza (03°44'S, 38°34'O, 20m), estado do Ceará. O município de Igarapé-Açu localiza-se na região oriental do bioma amazônica, apresenta clima chuvoso com pequena estação seca e enquadra-se no tipo Am da classificação de Köppen (ALVARES et al., 2013). O campo experimental de Pentecoste localiza-se na região semiárida do nordeste brasileiro com tipo climático BSw'h' da classificação de Köppen, clima seco com pequena temporada úmida. Fortaleza é uma região litorânea do nordeste brasileiro que apresenta clima tropical chuvoso, do tipo Aw' segundo classificação de Köppen, e localiza-se próximo ao limite da região semiárida brasileira. Os experimentos foram nomeados como: MS.IGA.15 (experimento que avaliou o capim Massai em Igarapé-Açu durante o ano de 2015), TZ.IGA.15 (experimento que avaliou o capim Tanzânia em Igarapé-Açu durante o ano de 2015), TZ.IGA.17-18 (experimento que avaliou o capim Tanzânia em Igarapé-Açu durante os anos de 2017 e 2018), MS.FOR.09 (experimento que avaliou o capim Massai em Fortaleza durante o ano de 2009), TZ.PEN.03 (experimento que avaliou o capim Tanzânia em Pentecoste durante o ano de 2003) e TZ.PEN.05-06 (experimento que avaliou o capim Tanzânia em Pentecoste durante o ano de 2005 e 2006). Um resumo das características climáticas durante os períodos experimentais de cada estudo é apresentado (Tabela 1).

Tabela 1 – Médias e desvios padrões (DP) das condições agrometeorológicas e caracterização das condições de cultivo durante a de coleta dos dados dos ensaios experimentais.

Identificação dos experimentos	Tipo de Solo	Fotoperíodo (horas)	Radiação Solar ($W m^{-2}$)	Temperatura média ($^{\circ}C$)	Evapotranspiração de Referência	Precipitação	Capacidade de água disponível	Evapotranspiração Real
TZ.IGA.15	Latossolo	12,00	461,93	27,21	8,40	4,95	100	3,78
	Amarelo	$\pm 0,02$	$\pm 76,08$	$\pm 1,33$	$\pm 1,37$	$\pm 8,86$		$\pm 3,45$
TZ.IGA.17-18	Latossolo	12,00	443,57	26,56	8,01	6,40	100	4,26
	Amarelo	$\pm 0,02$	$\pm 90,32$	$\pm 1,24$	$\pm 1,59$	$\pm 13,88$		$\pm 3,14$
TZ.PEN.03	Fluvissolo	12,00	215,26	26,78	4,51	12,60	56	4,18
		$\pm 0,05$	$\pm 25,60$	$\pm 0,79$	$\pm 0,53$	$\pm 20,46$		$\pm 0,57$
TZ.PEN.05-06	Fluvissolo	12,08	200,53	30,30	5,15	12,50	56	4,69
		$\pm 0,02$	$\pm 34,04$	$\pm 1,11$	$\pm 0,98$	$\pm 20,80$		$\pm 0,89$
MS.IGA.15	Latossolo	11,98	433,61	26,00	7,79	10,02	100	6,81
	Amarelo	$\pm 0,01$	$\pm 89,16$	$\pm 0,85$	$\pm 1,56$	$\pm 11,76$		$\pm 1,53$
MS.FOR.09	Argissolo	11,97	284,60	26,76	5,82	8,82	32	5,22
	Amarelo	$\pm 0,04$	$\pm 40,39$	$\pm 0,82$	$\pm 0,89$	$\pm 12,56$		$\pm 0,88$

3.2.2 Descrição dos Experimentos

O ensaio experimental TZ.IGA.15, foi conduzido na Fazenda Experimental de Igarapé-Açu (FEIGA) da Universidade Federal Rural da Amazônia (UFRA). Nesse ensaio, foi testado diferentes frequências de desfolhação em capim Tanzânia baseadas em dias fixos (14; 21; 28; 35; 42 e 49 dias) de período de descanso. Nós utilizamos os dados coletados entre o período de 14 março de 2015 a 02 de janeiro de 2016, esse intervalo de tempo abrangeu coletas realizadas tanto no período chuvoso quanto no período seco de 2015. Para mais detalhe sobre este estudo consultar Macedo et al. (2021).

O ensaio experimental TZ.IGA.17-18 utilizou as mesmas unidades experimentais do ensaio TZ.IGA.15. Esse estudo avaliou o efeito de diferentes intensidades de desfolhação com base na altura de resíduo (15; 25; 35; 45 e 55 cm) desfolhadas sempre que o dossel de capim Tanzânia atingia 95% de interceptação de luz. Os dados desse estudo compreenderam as coletas realizadas de 2 de setembro de 2017 a 12 de setembro de 2018, e abrangeu coletas realizadas durante o período seco e chuvoso. Para mais detalhe sobre este estudo consultar Lage Filho et al. (2021). Tanto o ensaio TZ.IGA.15 quanto o TZ.IGA.17-18 foram conduzidos em um delineamento em blocos aleatorizados, com as desfolhações realizadas por meio de corte mecânico.

O ensaio experimental TZ.PEN.03 foi realizado no campo avançado Núcleo de Ensino e Estudos em Forragicultura (NEEF), em Pentecoste-CE. O capim Tanzânia foi submetido a três períodos de descanso definidos em função do tempo necessário para a expansão de 1,5; 2,5 e 3,5 novas folhas por perfilho. Os dados desse estudo compreenderam as coletas realizadas de 3 de agosto de 2003 a 8 de novembro de 2003 e abrangeu o período seco da região. A desfolhação foi realizada por ovinos em uma área sob irrigação por aspersão com lâmina de água em torno de 11,4 mm dia⁻¹ e turno de rega de quatro dias. O delineamento foi o inteiramente casualizado com duas repetições. Para mais detalhe sobre este estudo consultar Silva et al. (2007).

O ensaio experimental TZ.PEN.05-06 foi realizado no NEEF, em Pentecoste-CE. O capim Tanzânia foi submetido a três períodos de descanso baseados no tempo necessário para o dossel atingir 85; 95 e 97% de interceptação de luz, em combinação com dois resíduos pós-pastejo baseado no índice de área foliar residual de 1,0 ou 1,8. O delineamento foi o inteiramente casualizado com duas repetições. Os dados desse estudo compreenderam as coletas realizadas de 25 de outubro de 2005 a 7 de março de 2006. A desfolhação foi realizada por ovinos em uma área sob irrigação fixa de aspersão com lâmina de água em torno de 11,4

mm dia-1 e turno de rega de quatro dias. Para mais detalhe sobre este estudo consultar Cutrim Junior et al. (2011).

No ensaio experimental MS.IGA.15, o capim Massai foi submetido a 5 doses de adubação nitrogenada (100; 200; 300; 400 e 500 kg ha-1 ano-1), realizadas em seis aplicações fixas ao longo do período experimental, e um tratamento controle (sem nitrogênio), com 5 repetições em um delineamento inteiramente casualizado. As desfolhações foram realizadas por meio de corte mecanizado quando a IRFA atingia 95%. O experimento foi realizado na FEIGA, durante o período de 14 de fevereiro de 2015 a 5 de agosto de 2015. Para mais detalhe sobre este estudo consultar Cunha et al. (no prelo).

O ensaio experimental MS.FOR.09 foi realizado no campo experimental do NEEF, em Fortaleza - CE. O capim Massai foi submetido a crescentes doses de nitrogênio (0-controle; 400; 800 e 1200 kg ha-1 ano-1) sob área irrigada por aspersão fixa com lâmina líquida de 7,0 mm dia-1 e turno de rega de três dias, em um sistema de lotação intermitente pastejado por ovinos. O delineamento experimental foi inteiramente casualizado com duas repetições e avaliações realizadas a cada ciclo de rebrotação com duração de 22; 18; 16 e 13 dias para os tratamentos controle (sem adubação), 400; 800 e 1200 kg de adubação nitrogenada, respectivamente. Para mais detalhes sobre este estudo consultar Lopes et al. (2016).

3.2.3 Variáveis agronômicas mensuradas

Em todos os ensaios anteriormente citados, as variáveis comuns mensuradas na pré-desfolhação foram: acúmulo de biomassa, componentes morfológicos, densidade populacional de perfilhos (DPP) e incremento em altura do dossel (InAlt). Também foi determinada o número de dias de rebrotação de cada período de descanso.

O acúmulo de biomassa foi tomado por duas coletas destrutivas com o uso de molduras de área conhecida que foram convertidas para hectares, considerou-se a forragem acima da altura de resíduo particular a cada ensaio. Uma das amostras foi utilizada para determinação do acúmulo de forragem total (AcFT), e a outra para separação morfológica e determinação do acúmulo de lâmina foliar (AcLF), acúmulo de haste (colmo + bainha) (AcH) e acúmulo de material morto (AcMM). Cada amostra foi levada à estufa de ventilação forçada (55 °C até peso constante), para determinação do acúmulo de biomassa em termos de matéria pré-seca. O InAltr foi obtido pela diferença do dossel antes e após a desfolhação medida com auxílio de régua graduada em centímetros. A DPP foi estimada contando-se os perfilhos vivos dentro de uma moldura de área conhecida e posteriormente convertida em número de perfilhos por metro quadrado.

3.2.4 Variáveis agrometeorológicas

As variáveis de temperatura média (T_{med}) e precipitação ($Prec$) para o município de Igarapé-Açu foram obtidas por meio da estação meteorológica convencional da Embrapa Amazônia Oriental, distante 900 metros dos experimentos. Os mesmos dados meteorológicos, para o município de Pentecoste, foram obtidos pela estação meteorológica automática da UFC a 800 metros da área experimental. Para os municípios de Igarapé-Açu, Pentecoste e Fortaleza os dados de radiação solar global (R_s) foram provenientes da empresa Solcast (SOLCAST, 2019).

A evapotranspiração de referência (ET_o) dos municípios de Fortaleza e Pentecoste, foi calculada por meio do método de Penman-Monteith FAO 56 (ALLEN et al., 1998):

$$ET_o = \frac{0,408\Delta(R_n - G) + \gamma \frac{900}{T_{med} + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (1)$$

onde, Δ é a inclinação da saturação de pressão de vapor como função da temperatura ($kPa \cdot ^\circ C^{-1}$); R_n , a radiação líquida ($MJ \cdot m^{-2} \cdot dia^{-1}$); G , o fluxo de calor do solo ($MJ \cdot m^{-2} \cdot dia^{-1}$); γ , constante psicométrica ($kPa \cdot ^\circ C^{-1}$); T_{med} , média entre a T_{max} e T_{min} ($^\circ C$); u_2 ; velocidade do vento ($m \cdot s^{-1}$); $(e_s - e_a)$, déficit de pressão de vapor ($kPa \cdot ^\circ C^{-1}$).

Em Igarapé-Açu, a ET_o foi calculada de acordo com o método de Turc (1961), que representa uma boa estimativa de ET_o para a região (DA SILVA JÚNIOR et al., 2017; FARIAS et al., 2019):

$$ET_o = \frac{0,013T_{med}}{T_{med} + 15} (23,9R_s + 50) \quad (2)$$

O índice hídrico (IH), foi obtido por meio da relação entre a evapotranspiração real (ETR) e a de referência (ET_o):

$$IH = \frac{\text{Evapotranspiração real}}{\text{Evapotranspiração de referência}} \quad (3)$$

A ETR foi calculada a partir da elaboração do balanço hídrico sequencial em escala diária de acordo com Thornthwaite e Mather (1955) com base na capacidade de água disponível (CAD), que foi diferente para cada localidade (Tabela 1).

Também foi considerado como variável agrometeorológica o nitrogênio fornecido (NF) na forma de adubação.

2.2.5 Análises estatísticas

A média e/ou somatório dos valores das variáveis agrometeorológicas e agronômicas, referente ao período de rebrotação de cada ciclo de produção de forragem, foram utilizados para formar esses dois grupos de variáveis.

A caracterização das observações pertencentes a cada grupo de experimento, verificadas por meio das variáveis produtivas e agrometeorológicas, foi gerada a partir da análise de componentes principais. Para verificar de que forma as variáveis produtivas se relacionam com as variáveis agrometeorológicas, os dados destas variáveis foram submetidos a análise de correlação canônica. A análise de função discriminante foi aplicada para verificar as funções responsáveis pela maximização da diferença entre os grupos de experimentos. Foi utilizado o software R (R CORE TEAM, 2019) como ferramenta para processamento dos dados.

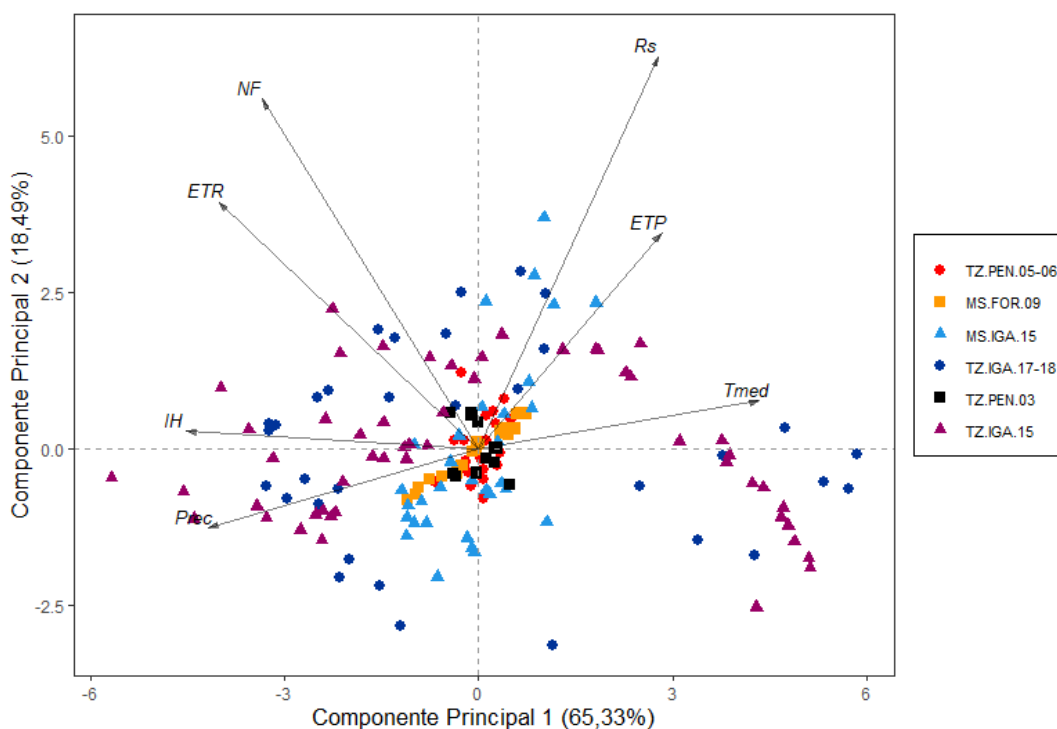
3.3 Resultados

3.3.1 Como que os diferentes experimentos se caracterizam em relação aos índices que agrupam variáveis agronômicas e agrometeorológicas?

A caracterização das observações referente aos seis experimentos, por meio das variáveis agrometeorológicas, é influenciada praticamente pelo primeiro componente principal (CP1, Figura 1), que responde por 65,33% da variação dos dados. O CP1 (Figura 1) demonstra um contraste entre as variáveis relacionadas ao ciclo hidrológico mais o NF, com variáveis relacionadas ao fornecimento energético e perda potencial de água, principalmente em relação a radiação solar e temperatura média. Sobre esse componente, verifica-se uma maior variabilidade dos experimentos que utilizaram o capim Tanzânia no município de Igarapé-Açu em relação aos demais experimentos (Figura 1).

O componente principal 2 (CP2, Figura 1) representa uma proporção menor da variação dos dados, e está relacionado a uma maior variabilidade dos experimentos realizados na região amazônica, tanto de capim Massai quanto Tanzânia em relação aos experimentos realizados na região nordeste (Figura 1).

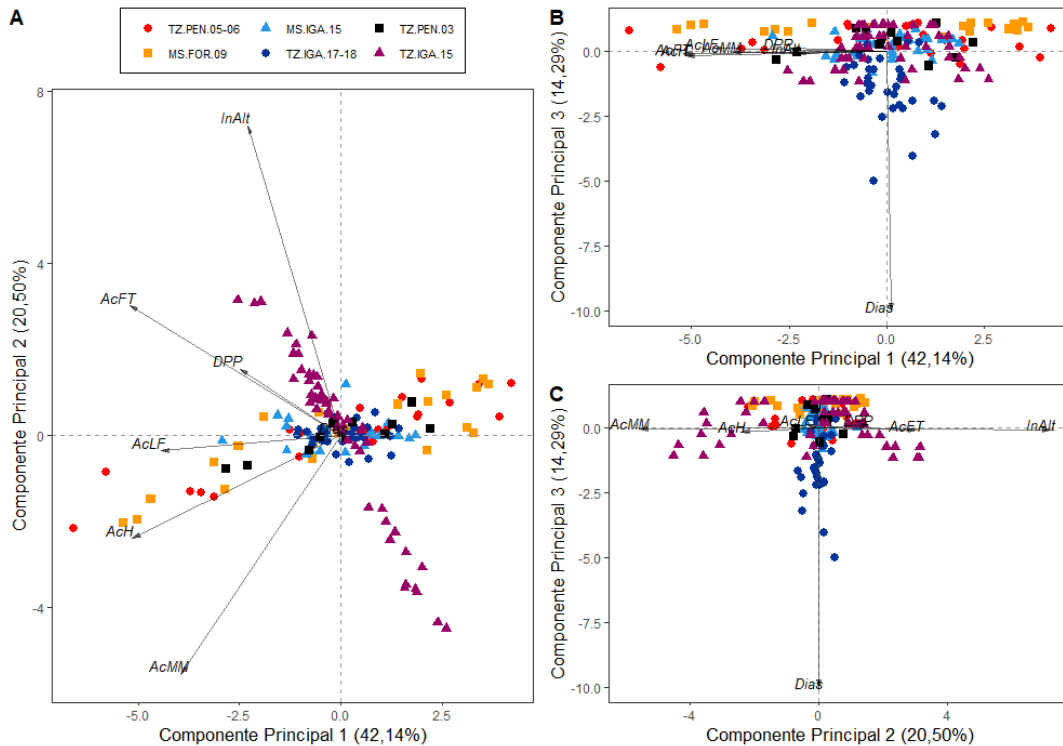
Figura 1 – Dois primeiros componentes principais obtidos a partir das variáveis agrometeorológicas, com observações provenientes dos 6 tipos de experimentos utilizados na análise.



Sob o ponto de vista das variáveis agrônômicas, considerou-se os três primeiros componentes principais, que juntos são responsáveis por 76,92% da variabilidade dos dados. O componente principal 1 (CP1, Figura 2), responsável pela maior parte da variação (42,14%), representa o crescimento dos capins, por meio de um índice que considera praticamente todas as variáveis agrônômicas, exceto os dias de rebrotação. Nele, todos os experimentos são caracterizados por uma variabilidade em relação a altos e baixos valores de variáveis agrônômicas (Figura 2 – A e B).

O segundo componente principal (CP2, Figura 2), que deteve 20,50% da variabilidade dos dados, está relacionado a um contraste principalmente entre AcH e AcMM versus InAlt e AcFT. Neste componente, observa-se uma maior variabilidade dos dados provenientes do experimento com capim Tanzânia ocorrido no ano de 2015 (Figura 2 – A e C). O componente principal 3 (CP3, Figura 2) é um componente que destaca a importância dos dias de rebrotação dos capins. Ele representou 14,29% da variabilidade dos dados, e mostra que essa variação no período de rebrotação entre os experimentos não mudou muito, exceto para alguns valores no experimento com Tanzânia em Igarapé-Açu durante os anos de 2017 e 2018 (Figura 2 – B e C).

Figura 2 – Três primeiros componentes principais obtidos a partir das variáveis agronômicas, com observações provenientes dos seis tipos de experimentos utilizados na análise.



3.3.2 Como que se relacionam as variáveis agrometeorológicas com as variáveis agronômicas?

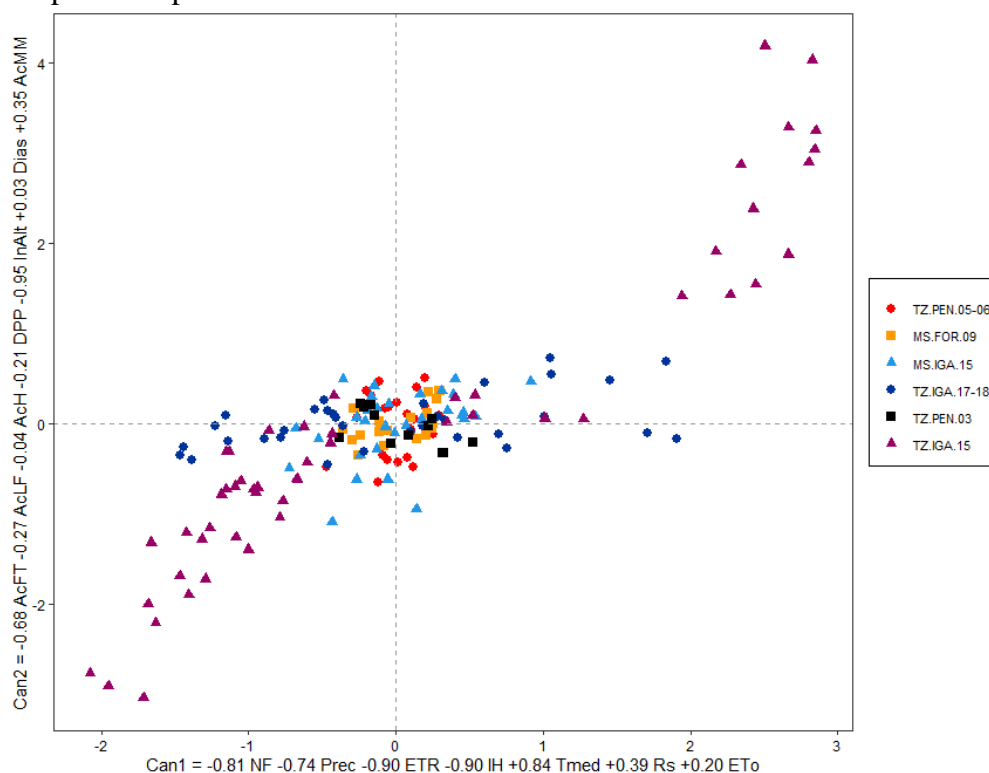
A correlação entre a variável canônica 1 (Can1, gerada a partir das variáveis agrometeorológicas) e as variáveis que a originaram, mostra um contraste entre variáveis relacionadas a disponibilidade de água (Prec, ETR e IH) mais o NF versus as variáveis relacionadas ao fornecimento de energia em termos de temperatura e radiação (ET_o , R_s , e T_{med}) (Figura 3).

Por meio da correlação entre a variável canônica 2 (Can2) e as variáveis produtivas, verifica-se que a variável Can2 representou um contraste entre a variável AcMM com as demais variáveis (AcFT, AcLF, DPP e InAlt), principalmente em relação ao incremento em altura (Figura 3). As variáveis AcH e dias de rebrotação apresentaram pouca participação nessa variável canônica.

O coeficiente de correlação entre a variável Can1 e Can2 foi de 0,84, isso representa 76,94% da variação tomada pelo primeiro par de variáveis canônicas em relação às demais. Observa-se que os dados oriundos do experimento com capim Tanzânia em Igarapé-Açu durante o ano de 2015 foram os mais heterogêneos em relação aos dos demais experimentos, quando observados sob a influência tanto da variável Can1 quanto da variável Can2. O

experimento com o capim Tanzânia em Igarapé-Açu durante os anos de 2017 e 2018 apresentou uma certa variabilidade em relação à variável Can1, porém com resultados mais homogêneos em relação à variável Can2 (Figura 3).

Figura 3 – Representação gráfica dos valores da variável Can1 em relação à Can2 e a distinção entre os 6 tipos de experimentos utilizados na análise.



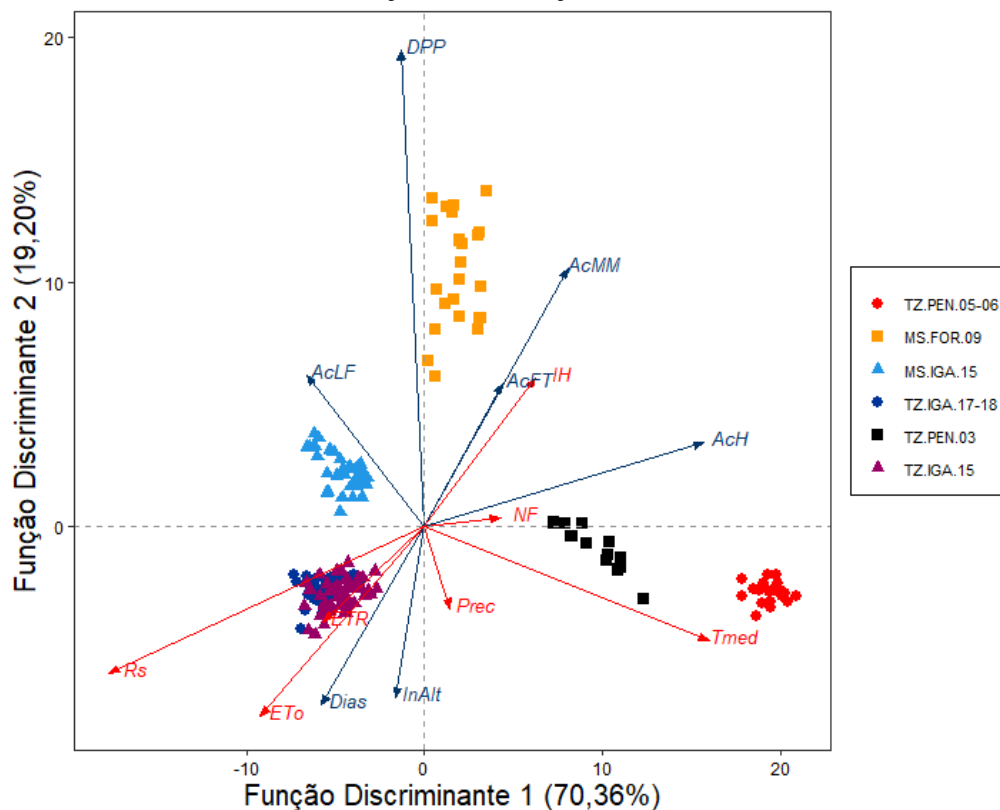
3.3.3 Como podem ser descritas funções discriminantes que separam experimentos com capim Tanzânia e Massai conduzidos em regiões distintas mas de latitude semelhante?

A separação entre os grupos de experimentos pode ser visualizada pela influência de duas funções discriminantes, que juntas, representam 89,56% da variação responsável pela máxima separação dos grupos (Figura 4). A função discriminante 1 (FD1), no que se refere às variáveis produtivas, mostra um contraste entre a dias de rebrotação e AcLF versus o AcFT, AcFM e AcH. Essa função discriminante separa os experimentos, principalmente em relação as regiões onde eles foram realizados. Ao lado esquerdo estão representados os experimentos realizados em Igarapé-Açu, ao lado direito a região de Fortaleza e mais aos extremos, os experimentos realizados na região de Pentecoste. Os dados provenientes dos experimentos na região sob influência do clima semiárido estão caracterizados por maior produção de biomassa e seus componentes de colmo e material morto, em relação aos dados de experimentos

realizados na região amazônica, com destaque para os experimentos com capim Tanzânia. Sob a ótica das variáveis agrometeorológicas, a separação das regiões ocorre por meio de um contraste principalmente entre ETR, ETo e Rs de um lado, e T_{med} e IH do outro (Figura 4).

A separação entre os diferentes cultivares de capim, ocorreu por meio da função discriminante 2 (FD2). Essa função, separa os experimentos que utilizaram o capim Massai, caracterizados por maior DPP, acúmulo de biomassa e seus componentes morfológicos, dos experimentos que avaliaram o capim Tanzânia, com maior incremento em altura e dias de rebrotação. Com relação as variáveis agrometeorológicas, ela é influenciada pelo T_{med} , Rs e ETo em contraste com o IH (Figura 4).

Figura 4 – Representação dos valores da FD1 em relação a FD2. Os pontos no gráfico mostram a distinção dos seis tipos de experimentos utilizados na análise. As setas em vermelho e em azul mostram a participação de cada variável agrometeorológica e produtiva, respectivamente, na formação da função discriminante. Quanto maior o comprimento da seta em relação ao eixo da função discriminante, maior a correlação desta função com a variável.



3.4 Discussão

Com relação as variáveis agrometeorológicas, tanto pela análise de componentes principais quanto pela análise de correlação canônica, verificou-se a formação de índices que resumem a informação das variáveis em basicamente dois grupos. Um formado por variáveis

relacionadas ao fornecimento energético em termos de radiação e temperatura, e outra relacionada a disponibilidade de água e nitrogênio. Porém o que pode distinguir mais os experimentos realizados em diferentes locais foi as características relacionadas a disponibilidade de água. Isso se corrobora com a nossa hipótese de que mesmo regiões situadas próximas a linha do equador, como as regiões amazônica e semiárida, essas possuem fatores climáticos que possam distingui-las (SILVA et al., 2019).

A construção dos componentes principais, se baseia na obtenção de combinações lineares entre as variáveis originais a partir da matriz de variância e covariância entre essas variáveis. Isso resulta em uma redução do número de variáveis em componentes principais que melhor caracterizam a forma como as observações podem ser visualizadas a partir da combinação das variáveis estudadas (YATER et al., 2014).

O CP1 (Figura 1) que separa as variáveis relacionadas ao ciclo hidrológico de um lado e as variáveis relacionadas ao fornecimento energético de outro, mostra que os experimentos realizados em Igarapé-Açu, sejam eles com capim Tanzânia ou Massai, possuem maior dispersão dos dados em relação aos experimentos realizados na região nordeste, seja em Fortaleza ou em Pentecoste. Isso se deve principalmente ao fato dos experimentos realizados na região nordeste levarem em consideração um maior controle das condições climáticas. Pois não se observa muita variação em relação aos dados de fotoperíodo, radiação e temperatura média (Tabela 1), característica essa das regiões localizadas próxima a linha do equador. Além disso, também não se observou variabilidade alta na evapotranspiração real, que está relacionada ao fornecimento hídrico constante para as forrageiras nesses experimentos, fornecida por meio da irrigação durante todo o período experimental (GRAHAM et al., 2016). A região de Igarapé-Açu, também se localiza próxima a linha do equador, conseqüentemente os dados de fotoperíodo, radiação solar e temperatura, também não variaram muito ao longo do período experimental. Contudo, os dados de ETR apresentaram uma maior variabilidade nessa região, principalmente em relação aos estudos TZ.IGA.15, e TZ.IGA.17-18, que foram experimentos em que houve avaliação de dados tanto no período chuvoso quanto no período seco. Já no estudo MS.IGA.15, as avaliações abrangeram somente o período chuvoso e o período de transição chuva-seca. O CP1 (Figura 1) também leva em consideração o fornecimento de nitrogênio, junto com as variáveis do ciclo hidrológico. Isso porque essa variável nos experimentos com capim Tanzânia em Igarapé-Açu também apresentou maior variabilidade pelo fato de que durante o período seco, não há o fornecimento de adubação nitrogenada na forma de ureia devido a falta de umidade no solo.

O CP1 (Figura 1) demonstra claramente um contraste entre variáveis que apresentaram maior variação em relação as que apresentaram menor variação dos dados principalmente em relação aos experimentos realizados em Igarapé-Açu. Isso nos fornece indícios de considerar o uso de métodos estatísticos baseados em regressão por componentes principais para a estimação de valores em variáveis produtivas, a partir de variáveis agrometeorológicas. Uma maior dispersão das observações dos experimentos conduzidos na região amazônica também é observada em relação ao segundo componente principal (CP2, Figura 1), porém essa dispersão é menor em relação ao CP1 (Figura 1). O CP2 (Figura 1) representou 18,49% da variação dos dados, e representa um índice de maior importância para as variáveis Rs, NF, ETR e ETo.

Nas variáveis agronômicas, a escolha do número de componentes baseou-se nos fatores que obtivessem autovalores maior que 1, de acordo com o critério de Kaiser (KAISER, 1958). Nesse caso é possível a obtenção de três componentes principais. Diferente do que aconteceu em relação as variáveis agrometeorológicas, o primeiro componente (Figura 2), que considera basicamente todas as variáveis agronômicas, mostra como praticamente todos os experimentos variaram em relação a ele. Isso se deve basicamente ao fato de que a variação nas observações seja em decorrência do efeito dos tratamentos adotados nos experimentos, sejam eles pelo uso de diferentes doses de adubação nitrogenada ou diferentes técnicas de manejo. É comum em sistemas de produção a pasto, a adoção de diferentes técnicas de manejo baseadas em metas de pastejo, que irão depender do nível tecnológico, sistemas de produção e possíveis variações edafoclimáticas da região (MACEDO et al., 2021; SILVA et al., 2019; SILVA et al., 2019). Esses fatores são então indispensáveis quando se pensa em variabilidade de variáveis resposta que sejam interessantes no monitoramento da estrutura e produtividade dos pastos. Provavelmente, não somente variáveis agrometeorológicas sejam importantes no entendimento e previsões de variáveis produtivas em pastagens. Um exemplo disso é a necessidade e inclusão de fatores de manejo em modelos mecanicistas que estimam a produção de biomassa, como o CROPGRO (BOSI et al., 2020; BRUNETTI et al., 2020) e APSIM (BOSI et al., 2020; GOMES et al., 2020).

O segundo componente principal (Figura 2), mostra principalmente uma maior variabilidade em relação ao estudo TZ.IGA.15, isso porque os dados observados desse experimento refletem uma ampla condição estrutural de dosséis com altos e baixos valores de variáveis agronômicas principalmente devido ao uso de dias de descanso que variaram de 14 a 49 dias (MACEDO et al., 2021). O terceiro componente principal (Figura 2), que contribui com 14,29% da variação em relação aos três primeiros componentes, destaca principalmente um efeito dos dias de rebrotação nas observações. Com exceção de alguns dados para o estudo

TZ.IGA.17-18, os experimentos apresentaram distribuição das observações similares em relação a esse componente. O fato de o experimento TZ.IGA.17-18 não ter considerado períodos de rebrotação fixos e possuir avaliações durante o período seco, a condição para que o dossel atingisse a interceptação de luz de 95% (meta do experimento) nessas condições aumentou bastante o período de rebrotação de algumas observações. Isso pode estar relacionado ao aumento da variabilidade das observações desse experimento, em relação a esse componente.

A análise de correlação canônica (Figura 3) tem como objetivo principal o estudo das relações lineares existentes entre dois conjuntos de variáveis, nesse caso, as variáveis agrometeorológicas e agronômicas. Cada conjunto de variáveis serão representados por uma variável canônica, de forma que a correlação entre essas duas variáveis canônicas seja a máxima possível. Portanto, a alta correlação canônica (0,84) entre as variáveis agrometeorológicas e agronômicas demonstra a influência próxima existente entre esses dois grupos de variáveis. Essa influência está relacionada ao fato da disponibilidade hídrica junto com o NF contrastantes com os fatores relacionados a perda potencial de água (E_{To} , R_s e T_{med}), afetarem diretamente o conjunto de variáveis ligadas ao crescimento do dossel forrageiro (Ac_{FT} , Ac_{LF} , DPP e In_{Alt}) que são fatores importantes em sistemas de produção de herbívoros. Esses fatores importantes contrastam com o Ac_{MM} que é o resultado da perda de tecido vegetal, e é visto de forma negativa em sistemas de produção intensivos.

A análise de regressão múltipla é um caso particular de correlação canônica, esses resultados podem ser interessantes para o entendimento de quais variáveis agrometeorológicas são importantes na composição de modelos empíricos para estimação de variáveis produtivas, como a produção de biomassa, densidade de perfilhos, dentre outras. Nesse caso, essas variáveis produtivas estariam mais relacionadas aos dados que envolvem a disponibilidade hídrica, principalmente a ETR e o IH , mais o NF . Isso pode ser observado principalmente a partir dos dados em que há uma maior variabilidade nas condições ambientais durante a condução do experimento, que foram os estudos TZ.IGA.15 e TZ.IGA.17-18 (Figura 3). Os aspectos relacionados ao manejo do estudo TZ.IGA.15, considerando períodos de rebrotação que variaram de 14 a 49 dias, fez com que esse experimento também apresentasse uma alta variabilidade em relação a variável Can_2 , o que leva a resultados provenientes de uma alta variabilidade de condições estruturais do dossel forrageiro (MACEDO et al., 2021).

A discriminação dos experimentos quanto aos locais onde eles foram realizados ocorre pela função discriminante 1. Essa função é a que tem o maior poder de separação entre os grupos. Isso demonstra que a influência relacionada aos diferentes locais é mais acentuada que aquela relacionada a diferença entre os capins. Contudo, mesmo considerando locais iguais,

ainda é possível visualizar a diferença existente entre os experimentos TZ.PEN.03 e TZ.PEN.05-06. Isso provavelmente se deve ao fato de existirem diferenças climáticas entre a época em que os experimentos foram conduzidos, mais precisamente na temperatura, onde no período de 2005 a 2006 essa foi 3,52°C superior em relação a 2003 (Tabela 1). Por sua vez, essa diferença na temperatura foi menor quando observada entre os experimentos TZ.IGA.15 e TZ.IGA.17-18, de 0,65 °C. Como ambos os experimentos de Pentecoste foram supridos adequadamente por água através da irrigação, esse efeito da temperatura pode ter influenciado a aceleração dos processos fisiológicos da planta, resultando em maior acúmulo de forragem, alongamento de colmo e material morto (IVORY; WHITEMAN., 1978).

A discriminação dos experimentos quanto a espécie do capim estudado ocorre pela função discriminante 2. Essa função é a que tem o segundo maior poder de separação dos grupos. É possível observar a forte influência que as variáveis agronômicas promovem em relação a essa função discriminante, principalmente em relação ao efeito da DPP, AcFT, AcMM e AcLF. O capim Massai se destaca em relação a essas variáveis, justamente por apresentar maior DPP e produção de biomassa em relação ao capim Tanzânia (VERAS et al., 2020). Já o capim Tanzânia apresenta maiores valores em incremento de altura, pois esse cultivar apresenta porte maior em relação ao capim Massai. É possível constatar aqui a relação negativa existente em termos de altura do dossel e DPP para gramíneas tropicais observados em estudos recentes (MACEDO et al., 2021; XILIANG et al., 2021), no qual dosséis com maior altura tendem a ter menor DPP.

3.5 Conclusão

Por meio de uma abordagem multivariada com dados experimentais, nós mostramos que em sistemas de produção baseado no cultivo de gramíneas forrageiras, em regiões localizadas próxima a linha do equador, é possível considerar que apesar das semelhanças quanto ao fornecimento de energia em termos de radiação e temperatura, a disponibilidade de água e o fornecimento de nutrientes é um fator determinante que pode auxiliar estudos futuros que envolvam o uso de modelagem aplicada a produção de forragem. Porém, em sistemas de produção em pastagens, o manejo aplicado ao dossel também tem um efeito determinante na produção de biomassa e seus componentes morfológicos e estruturais, por isso deve ser considerado nesse processo que estuda as relações entre variáveis e caracterização produtivas das pastagens.

3.6 Referências

- ALLEN, R.G., PEREIRA, L.S., RAES, D., SMITH ROME, M. **Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56**. Fao 300, D05109, 1998.
- ALVARES, C. A., STAPE, J. L., SENTELHAS, P. C., GONÇALVES, J. L. M., SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 2013.
- BOSI, C., SENTELHAS, P. C., HUTH, N. I., PEZZOPANE, J. R. M., ANDREUCCI, M. P., SANTOS, P. M. APSIM-Tropical Pasture: A model for simulating perennial tropical grass growth and its parameterisation for palisade grass (*Brachiaria brizantha*). **Agricultural Systems**. v. 184, p. 102917. 2020.
- BOSI, C., SENTELHAS, P. C., PEZZOPANE, J. R. M., SANTOS, P. M. CROPGRO-Perennial Forage model parameterization for simulating Piatã palisade grass growth in monoculture and in a silvopastoral system. **Agricultural Systems** , v. 177, p. 102724, 2020.
- BRUNETTI, H. B., BOOTE, K. J., SANTOS, P. M., PEZZOPANE, J. R. M., PEDREIRA, PEDREIRA, C. G. S., LARA, M. A. S., MORENO, L. S. B., HOOGENBOOM, G. Improving the CROPGRO Perennial Forage Model for simulating growth and biomass partitioning of guineagrass. **Agronomy Journal**, v. 113, p. 3299-3314, 2021.
- CUTRIM JUNIOR, J. A. A. CÂNDIDO, M. J. D., VAKENTE, B. S. M., CARNEIRO, M. S. D., CARNEIRO, H. A. V. Características estruturais do dossel de capim-tanzânia submetido a três frequências de desfolhação e dois resíduos pós-pastejo. **R. Bras. de Zootec.**, v. 40, p. 489-497, 2011.
- DA SILVA JÚNIOR, R. O., SOUZA, E. B., TAVARES, A. L., MOTA, J. A., FERREIRA, D. B. S., SOUZA FILHO, P. W. M., ROCHA Three decades of reference evapotranspiration estimates for a tropical watershed in the eastern Amazon. **Anais da Academia Brasileira de Ciências**, v. 89, n. 3, p. 1985–2002, 2017.
- FARIAS, V. D. S., COSTA, D. L. P., PINTO, J. V. N., SOUZA, P. J. O. P., SOUZ, E. B., FARIAS, S. O. Calibration of reference evapotranspiration models in Pará. **Acta Scientiarum. Agronomy**, v. 42, p. e42475, 2019.
- FITZPATRICK, E. A.; NIX, H. A. The climatic factor in Australian grassland ecology. In: MOORE, R. M. (Ed.). **Australian grasslands**. Canberra, ACT, Australia: Australian National University Press, 1970. p. 3–26.
- GOMES, F. J., BOSI, C., PEDRIERA, B. C., SANTOS, P. M., PEDREIRA, C. G. S., Parameterization of the APSIM model for simulating palisadegrass growth under continuous stocking in monoculture and in a silvopastoral system. **Agricultural System**. v. 184, p. 102876, 2020.
- GRAHAM, S. L., KOCHENDORFER, J., MCMILLAN, A. M. S., DUNCAN, M. J., SRINIVASAN, M. S., HERTZOG, G. Effects of agricultural management on measurements, prediction, and partitioning of evapotranspiration in irrigated grasslands. **Agricultural Water**

Management, v. 177, p. 340-347, 2016.

IVORY, D.; WHITEMAN, P. Effect of Temperature on Growth of Five Subtropical Grasses. I. Effect of Day and Night Temperature on Growth and Morphological Development. **Functional Plant Biology**, v. 5, n. 2, p. 131, 1978.

KAISER, H. F. The varimax criteria for analytical rotation in factor analysis. **Psychometrika**, v. 23, p. 141-51, 1958.

LOPES, M. N., CÂNDIDO, M. J. D., POMPEU, R. C. F. F., SILVA, R. G., MORAIS NETO, L. B., CARNEIRO, M. S. S. Tillering dynamics in massai grass fertilized with nitrogen and grazed by sheep. **Bioscience Journal**, v. 32, p. 446-454, 2016.

MACEDO, V. H. M., CUNHA, A. M. Q., CÂNDIDO, E. P., DOMINGUES, F. N., SILVA, W. L., LARA, M. A. S., RÊGO, A. C. Canopy structural variations affect the relationship between height and light interception in Guinea Grass. **Field Crops Research**, v. 271, p. 108249, 2021.

MARTIN, T. N., DOURADO NETO, D., VIEIRA JUNIOR, P. A., PEREIRA, A. R., MANFRON, P. A., CHRISTOFFOLETI, P. J. Modified crop model estimation of depleted and pterial soybean yield. **Acta Scientiarum**.v. 34, p. 369-378. 2012.

MCMASTER, G. Growing degree-days: one equation, two interpretations. **Agricultural and Forest Meteorology**, v. 87, n. 4, p. 291-300, 1997.

MORENO, L. S. B., PEDREIRA, C. G. S., BOOTE, K. J., ALVES, R. R. Base temperature determination of tropical Panicum spp. grasses and its effects on degree-day-based models. **Agricultural and Forest Meteorology**, v. 186, p. 26-33, 2014.

PUNTEL, L. A., SAWYER, J. E., BARKER, D. W., THORBURN, P. J., CASTELLANO, M. J., MOORE, K. J., VANLOOCKE, A., HEATON, E. A., ARCHONTOULIS, S. V. A systems modelong approach to forecast corn economic optimum nitrogen rate. **Frot. Plant. Sci.**, v. 9, p. 436, 2018.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundationfor Statistical Computing, Vienna, Austria. 2019. Disponível em: <https://www.R-project.org/>. Acesso em 05 set. 2019.

SILVA, P. E., SILVA, C. M. S., SPYRIDES, M. H. C., ANDRADE, L. M. B. Precipitation and air temperature extremes in the Amazon and northeast Brazil. **International Journal of Climatology**, v. 39, n. 2, p. 579-595, 2019.

SILVA, R. G. CÂNDIDO, M. J. D., NEIVA, J. N. M., LÔBO, R. N. ., SILVA, D. S. Características estruturais do dossel de pastagens de capim-tanzânia mantidas sob três períodos de descanso com ovinos. **R. Bras. Zootec.**, v. 36, p. 1255-1265, 2007.

SILVA, S. C., BUENO, A. A. O., CARNEVALLI, R. A., SILVA, G. P., CHIAVEGATO, M. B. Nutritive value and morphological characteristics of Mombaça grass managed with different rotational grazinff strategie. **The Journal of Agricultural Science**. v. 157, p. 592-598. 2019

SILVA, W. L., COSTA, J.P.R., CPUTTI, G .P., LAGE FILHO, N. M., RUGGIERI, A. C.,

REIS, R. A. Effects of grazing intensity and supplementation strategies on Tifton 85 production and on sheep performance. **Small Ruminant Research**. V. 174, p. 118-124, 2019.

SOLCAST. **Global solar irradiance data and PV system power output data**. Disponível em: <<https://solcast.com/>>. Acesso em: 15 jan. 2020.

THORNTHWAITE, C. W.; MATHER, J. R. **The water balance**. New Jersey: Publications in climatology, 1955. v. 8

TURC, L. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. **Ann. Agron**, v. 12, n. 1, p. 13-49, 1961.

VERAS, E. L. L., DIFANTE, G. S., GURGEL, A. L. C., COSTA, A. B. G., RODRIGUES, J. G., COSTA, C. M., EMERENCIANO NETO, J. V., PEREIRA, M. G., COSTA, P. R. Tillerling and Structural Characteristics of Panicum Cultivars in the Brazilian Semiarid Region. **Sustainability**. V. 12, p. 3849, 2020

VILLA NOVA, N. A., TONATO, F., PEDREIRA, C. G. S., MEDEIROS, H. R. Método alternativo para cálculo da temperatura base de gramíneas forrageiras. **Ciência Rural**, v. 37, n. 2, p. 545-549, 2007.

WANG, E.; ENGEL, T. Simulation of phenological development of wheat crops. **Agricultural Systems**, v. 58, n. 1, p. 1-24, 1998.

XILIANG, L., PNG, G. K., LI, Y., JIMOH, S. O., DING, Y., LI, F., SUN, S. Leaf plasticity contributes to plant anti-herbivore defenses and indicates selective foraging: Implications for sustainable grazing. **Ecol. Ind**, v. 122, p. 107273. 2021.

4. O USO DE ALGORITMOS DE APRENDIZADO DE MÁQUINA MELHORA A PREDIÇÃO DA BIOMASSA DE FORRAGEM EM REGIÕES COM DIFERENTES CONDIÇÕES AMBIENTAIS?

RESUMO

Os estudos com predição de forragem em sistemas pastoris vêm crescendo no mundo, porém, com pouca expressão em regiões tropicais, e sem o uso de abordagem empírica por meio da utilização de métodos avançados de análise, como o uso de algoritmos de aprendizado de máquina. Com isso, objetivou-se no presente estudo, gerar modelos empíricos provenientes de dados da região amazônica brasileira e verificar quais métodos baseados em aprendizado de máquina podem ser utilizados para prever a produção de biomassa do capim Tanzânia na região semiárida do Brasil. Utilizou-se dados provenientes de dois estudos com capim Tanzânia realizados em Igarapé-Açu (01°07'S, 47°36'O, 47m), para a geração dos modelos, e dados proveniente de um experimento em Pentecoste (03°49'S, 39°20'O, 47m) para validação dos modelos. Para realizar a parametrização dos modelos, foram utilizadas variáveis agronômicas e agrometeorológicas. Dentre os algoritmos utilizados para geração dos modelos de previsão, o menor erro do quadrado médio foram os gerados a partir do método dos mínimos quadrados parciais. Dentre os modelos gerados por esse método, o melhor modelo baseado na porcentagem de viés foi o modelo mais simples, que usou somente variáveis agrometeorológicas e sem interação bidirecional como fatores preditivos. No geral, modelos em que foram incluídas variáveis com interação bidirecional ou variáveis agronômicas não melhoraram a capacidade de predição do modelo. A radiação solar, e as variáveis relacionadas a disponibilidade hídrica foram as que apresentaram maior importância na geração do modelo de predição mais acurado. Portanto, é possível utilizar modelos gerados a partir de métodos de aprendizado de máquinas para predição da produção de biomassa em regiões distintas.

Palavras-chave: Amazônia, aprendizado de máquina, modelo matemático, semiárido, validação.

ABSTRACT

Studies with forage prediction in pastoral systems have been growing in Brazil, however, with little expression in the North and Northeast regions of the country, and without the use of an empirical approach using advanced analysis methods, such as the use of machine learning algorithms. Thus, this study aimed to generate empirical models from Amazon region data and verify which methods based on machine learning can be used to predict the biomass production of Tanzania grass in the semiarid region of Brazil. Data from two studies with

Tanzania grass conducted in Igarapé-Açu (PA) were used to generate the models, and data from an experiment in Pentecoste (CE) to validate the models. To carry out the parameterization of the models, agronomic and agrometeorological variables were used. Among the algorithms used to generate the forecast models, the smallest mean square error was generated using the partial least squares method. Among the models generated by this method, the best model based on the percentage of bias was the simplest model, which used only agrometeorological variables and no bidirectional interaction as predictive factors. In general, models in which variables with bidirectional interaction or agronomic variables were included did not improve the predictive capacity of the model. Solar radiation and the variables related to water availability were the ones that were most important in generating the most accurate prediction model. Therefore, it is possible to use models generated from machine learning methods to predict biomass production in different regions.

Keywords: amazon, machine learning, mathematical model, semiarid, validation.

4.1 Introdução

Dentro da modelagem matemática, que envolve os estudos relacionados a previsão do rendimento de culturas, os modelos empíricos, que são modelos gerados a partir de análises de regressão, explicam pouco ou praticamente nada dos processos biofísicos envolvidos na relação de variáveis de entrada e saída de uma função. Porém, esses modelos são de aplicação relativamente simples, e de uso valioso em situações que demandam uma quantidade menor de informações para o processo de simulação (ANDRADE et al., 2016).

Um dos problemas mais importantes relacionados a análise de regressão, é a multicolineariedade de variáveis preditoras que fazem parte do modelo. Os métodos tradicionais de regressão múltipla e seleção de variáveis como *forward*, *backward* e *stepwise*, podem apresentar baixo poder de previsão principalmente quando o número de variáveis preditoras é alto e essas variáveis são correlacionadas (RANSOM et al., 2019). O surgimento de métodos de regressão mais avançados, baseados em aprendizado de máquina, podem ser uma alternativa na geração de modelos empíricos que garantam melhor previsão.

Muitos dos modelos empíricos gerados para simulação de produção de forragem no Brasil, são provenientes, em sua maioria, de dados oriundos da região sudeste onde a variabilidade de variáveis climáticas são maiores ao longo do ano, como temperatura e fotoperíodo (ARAUJO et al., 2013; CRUZ et al., 2011; PEZZOPANE et al., 2012; 2013; 2017;

2018; TONATO et al., 2010). Devido a variação climática existente no Brasil, há uma necessidade de estudos e geração de modelos que considere as condições de regiões de mais baixa latitude no país (ANDRADE et al., 2016), como na região norte e nordeste.

Uma das limitações do uso de modelos empíricos é a capacidade de generalização para diferentes condições ambientais (ASHOOR et al., 2019). Modelos empíricos geralmente são usados nos mesmos locais de onde os dados para a sua geração e calibração foram obtidos. Uma questão importante é: diferentes regiões que apresentam uma certa similaridade em relação a algumas condições climáticas podem ser usadas para a geração e utilização de modelos empíricos de produção agrícola? Um exemplo, é a região amazônica e a região semiárida do Brasil, em que ambas estão localizadas próximas a linha do equador e portanto não apresentam tanta variação em relação ao fotoperíodo, radiação e temperatura. O que mais distingue essas regiões são o regime de chuvas e conseqüentemente a disponibilidade hídrica, relacionadas a capacidade de armazenamento de água no solo (cap 2).

Apesar dos modelos empíricos serem gerados para aplicação no local ao qual os dados para sua criação foram coletados, métodos de regressão mais avançados supostamente podem gerar esses modelos que possam ser aplicados tanto na região amazônica quanto na região semiárida, tendo em vista que ambas as regiões estão em latitudes próximas e apresentam basicamente o mesmo fator crítico climático, que é a disponibilidade hídrica. Portanto, nosso objetivo com o presente estudo é a partir da geração de modelos empíricos provenientes de dados da região amazônica, verificar qual combinação de variáveis preditoras e quais métodos de regressão baseado em algoritmos de aprendizado de máquina são capazes de gerar modelo para a previsão de biomassa de capim Tanzânia para a região semiárida do Brasil.

4.2 Material e Métodos

4.2.1 Dados de geração dos modelos

Para a geração dos modelos de previsão do acúmulo de forragem do capim Tanzânia, utilizou-se dados de dois experimentos realizados na Fazenda Escola de Igarapé-Açu (FEIGA) da Universidade Federal Rural da Amazônia (UFRA), no município de Igarapé-Açu (01°07'S, 47°36'O, 47m), estado do Pará, Brasil. O local faz parte da região oriental do bioma amazônico, apresentando clima chuvoso com pequena estação seca e enquadrasse no tipo Am da classificação de Köppen (ALVARES et al., 2013). Para o primeiro experimento, utilizou-se dados referente ao período de março de 2015 a janeiro de 2016. Nesse experimento testou diferentes períodos de rebrotação entre desfolhações com base em dias fixos (14, 21, 28, 35, 42 e 49 dias) em um delineamento em blocos casualizados (MACEDO et al., 2021). O segundo

experimento foi conduzido na mesma área do primeiro, avaliando-se o efeito de diferentes alturas de resíduo (15, 25, 35, 45 e 55 cm) em um delineamento em blocos casualizados, durante o período de setembro de 2017 a setembro de 2018. Ambos os experimentos contêm dados provenientes de avaliações realizadas no período chuvoso e período seco.

4.2.2 Dados de validação dos modelos

Para a validação dos modelos, utilizou-se dados de um experimento conduzido Núcleo de Ensino e Estudos em Forragicultura (NEEF) do Departamento de Zootecnia do Centro de Ciências Agrárias da Universidade Federal do Ceará – DZ/CCA/UFC, em Pentecoste – CE (03°49'S, 39°20'O, 47m). O município está localizado na região semiárida do nordeste brasileiro, com clima BSw'h' da classificação de Köppen (ALVARES et al., 2013), caracterizado como clima seco com pequena temporada úmida. Nesse experimento, o capim Tanzânia foi submetido a três períodos de descanso definidos em função do tempo necessário para a expansão de 1,5; 2,5 e 3,5 novas folhas por perfilho. A desfolhação foi realizada por ovinos em uma área sob irrigação fixa de aspersão com lâmina de água em torno de 11,4 mm.dia⁻¹ e turno de rega de quatro dias. O delineamento foi o inteiramente aleatorizado com duas repetições e avaliações realizadas a cada ciclo de rebrotação. Para mais detalhe sobre este estudo consultar Silva et al. (2007).

Tanto dos experimentos usados para o processo de geração dos modelos quanto para a validação, foram obtidos dados de dois grupos de variáveis preditoras. O primeiro grupo é referente as variáveis agronômicas e o segundo grupo são de variáveis agrometeorológicas.

4.2.3 Variáveis agronômicas utilizadas na parametrização dos modelos

Para compor o grupo das variáveis agronômicas, foram utilizados os dados de acúmulo de biomassa acima do resíduo, colhida por meio de corte mecanizado em moldura de área conhecida e convertida em kg ha⁻¹. Altura de resíduo (AltResid) pós-desfolhação do dossel e incremento altura (InAlt) do dossel, obtida pela diferença entre a altura pré-desfolhação e pós-desfolhação, a altura foi mensurada com o uso de uma régua graduada em centímetros. A densidade populacional de perfilhos (DPP), obtida a partir da contagem de perfilhos vivos em moldura de área conhecida e convertida para perfilhos por m². E o número de dias de rebrotação (DR).

4.2.4 Variáveis agrometeorológicas utilizadas na parametrização dos modelos

As variáveis de temperatura média e precipitação (Prec) para o município de Igarapé-Açu foram obtidas por meio da estação meteorológica convencional da Embrapa Amazônia Oriental, distante 900 metros dos experimentos. Os mesmos dados meteorológicos, para o município de Pentecoste, foram obtidos pela estação meteorológica automática da UFC à 800 metros da área experimental. Tanto para o município de Igarapé-Açu quanto de Pentecoste os dados de radiação solar global (R_s) são provenientes da empresa Solcast (SOLCAST, 2019).

A variável graus dias de crescimento (GDC) acumulado em cada ciclo de produção foi determinado pela soma dos graus-dia (GD) dentro de cada ciclo. Em todas as condições experimentais a temperatura mínima do dia não foi menor que a temperatura base (T_b) das espécies, de 14,3 °C e 16,5 °C para o capim Tanzânia (PEZZOPANE et al., 2017) e Massai (MORENO et al., 2014), respectivamente. Tão pouco superior a temperatura limite de 40 °C. Dessa forma, o GD foi calculado de acordo com McMaster (1997):

$$GD = \frac{T_{\max} + T_{\min}}{2} - T_b \quad (1)$$

A unidade fototérmica (UF) é uma variável que combina o efeito da temperatura e do fotoperíodo (VILLA NOVA et al., 2007). A UF é calculada como segue:

$$UF = \frac{\left(\frac{n_d}{2} \overline{GDC}\right)^{\frac{N_f}{N_i} + 1}}{\frac{N_f}{N_i} + 1} \quad (2)$$

onde \overline{GDC} é a média do GDC acumulado durante o período de crescimento, n_d é o número de dias do ciclo de produção, N_i e N_f são o fotoperíodo no início do ciclo e no final do ciclo, respectivamente.

Para o município de Pentecoste, a evapotranspiração de referência (ET_o) foi calculada por meio do método de Penman-Monteih FAO 56 (ALLEN et al., 1998):

$$ET_o = \frac{0,408\Delta(R_n - G) + \gamma \frac{900}{T_{\text{med}} + 273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (3)$$

onde, Δ é a inclinação da saturação de pressão de vapor como função da temperatura ($\text{kPa} \cdot ^\circ\text{C}^{-1}$); R_n , a radiação líquida ($\text{MJ} \cdot \text{m}^{-2} \cdot \text{dia}^{-1}$); G , o fluxo de calor do solo ($\text{MJ} \cdot \text{m}^{-2} \cdot \text{dia}^{-1}$); γ , constante psicométrica ($\text{kPa} \cdot ^\circ\text{C}^{-1}$); T_{med} , média entre a T_{\max} e T_{\min} ($^\circ\text{C}$); u_2 ; velocidade do vento ($\text{m} \cdot \text{s}^{-1}$); $(e_s - e_a)$, déficit de pressão de vapor ($\text{kPa} \cdot ^\circ\text{C}^{-1}$).

Em Igarapé-Açu, a ET_o foi calculada de acordo com o método de Turc (1961), que representa uma boa estimativa de ET_o para a região (DA SILVA JÚNIOR et al., 2017; FARIAS et al., 2019):

$$ET_o = \frac{0,013T_{med}}{T_{med} + 15} (23,9R_s + 50) \quad (4)$$

O índice climático de crescimento (ICC) proposto de acordo com metodologia descrita por Fitzpatrick e Nix (1970), obtido por meio da multiplicação de três diferentes índices que variam de zero a um, foi calculado como segue:

$$ICC = IL \times IT \times IH \quad (5)$$

onde IL é o índice de luz, obtido pela seguinte equação:

$$IL = 1 - e^{(-3,5\frac{R_s}{750})} \quad (6)$$

com $e = 2,71829$, e a R_s em $\text{cal.cm}^{-2}.\text{dia}^{-1}$. O índice térmico (IT), é derivado de curvas que relacionam a temperatura do ar com a produção de matéria seca, neste estudo o IT foi obtido por meio da relação com a T_{med} utilizando uma função do tipo beta desenvolvida por Wang e Engel (1998), que considera a temperatura ótima (T_{otm}), a T_b e a temperatura limite superior (T_{ls}) para a espécie em estudo por meio da equação:

$$IT = \begin{cases} \frac{2(T_{med} - T_b)^\alpha (T_{otm} - T_b)^\alpha - (T_{med} - T_b)^{2\alpha}}{(T_{otm} - T_b)^{2\alpha}} & ; \text{ se } T_b \leq T_{med} \leq T_{ls} \\ 0 & ; \text{ se } T_{med} < T_b \text{ ou } T_{med} > T_{ls} \end{cases} \quad (7)$$

foi considerado T_{otm} igual a $30,1^\circ\text{C}$ e T_{ls} igual a 40°C de acordo com estudo que utilizou gramíneas da espécie *Panicum maximum* (IVORY; WHITEMAN, 1978), o parâmetro α é o fator de forma da função beta, tomado por:

$$\alpha = \frac{\ln 2}{\ln\left(\frac{T_{ls} - T_b}{T_{otm} - T_b}\right)} \quad (8)$$

O índice hídrico (IH), é obtido por meio da relação entre a evapotranspiração real (ETR) e a de referência (ET_o):

$$IH = \frac{\text{Evapotranspiração real}}{\text{Evapotranspiração de referência}} \quad (9)$$

A evapotranspiração real (ETR) calculada a partir da elaboração do balanço hídrico sequencial em escala diária de acordo com Thornthwaite e Mather (1955), assim como a porcentagem de água armazenada no solo (ARM) com base na capacidade de água disponível (CAD). A CAD para o experimento conduzido em Igarapé-Açu foi de 100 mm e para o município de Pentecoste foi de 56 mm.

Também foi considerado como uma variável agrometeorológicas, a quantidade de nitrogênio fornecida (NF) nos experimentos, por meio da adubação nitrogenada aplicada em cada ciclo de produção.

4.2.5 Análises estatísticas

Antes da geração dos modelos realizou-se análise de correlação de Spearman (ρ) para avaliar de forma não paramétrica o grau de correlação entre as variáveis preditoras a serem usadas na geração do modelo e verificar a possível presença de multicolinearidade.

4.2.5.1 Geração dos modelos

Os modelos de regressão foram gerados utilizando os dados de acúmulo de biomassa como variável dependente. Foi considerado a geração de modelos com e sem a inclusão de variáveis agrônômicas usando oito algoritmos de aprendizado de máquina baseados em análise de regressão.

Os métodos de aprendizado de máquina baseados em regressão penalizada foram: método LASSO de Tibshirani (1996), que reduz os parâmetros do coeficiente a zero, selecionando assim variáveis essenciais e reduzindo o número de parâmetros do modelo simultaneamente. O método LASSO adaptativo, refere-se a uma adaptação do método LASSO, que considera além da restrição aplicada aos valores dos parâmetros, uma diminuição maior no número de fatores por desconsiderar parâmetros que apresentam valores muito próximos de zero. O método de rede elástica (ZOU; HASTIE, 2005), faz uma melhor seleção de variáveis por considerar a escolha de variáveis que fazem parte de um grupo de variáveis que são correlacionadas. Os métodos de regressão penalizada foram realizados utilizando o procedimento GLMSELECT do software SAS (SAS Institute Inc., 2018).

Além dos métodos de penalização, também foram considerados os métodos de regressão por componentes principais e regressão de mínimos quadrados parciais. A regressão por componentes principais transforma grupos de variáveis explicativas em vetores de base ortogonal ou componentes principais que melhor capturam a variação das variáveis originais. A regressão de mínimos quadrados parciais funciona encontrando iterativamente os vetores projetados com uma variância mínima entre as variáveis dependentes e independentes. O conjunto de vetores resultante é a melhor relação entre as variáveis explicativas e de resposta que explica a estrutura básica dos dados. Essas análises foram realizadas por meio do procedimento PLS do software SAS.

Também foram considerados os algoritmos de árvore de decisão, árvores ensacadas e floresta aleatória. Esses métodos funcionam dividindo continuamente um conjunto de dados com base em algumas estatísticas de variáveis explicativas, criando assim um fluxograma de

decisões para prever uma variável resposta (QUINLAN, 1986). Os algoritmos baseados em árvores de decisão foram aplicados por meio dos pacotes *tree* e *randomForest* do software R (R CORE TEAM, 2019).

Com exceção dos modelos gerados a partir dos métodos baseados em árvore de decisão, para os demais métodos também foi considerado no modelo a ausência e presença de interações bidirecionais. Ao todo foram obtidos 26 cenários de modelagem.

4.2.5.2 Validação dos modelos

Os valores de produção de biomassa preditos em cada cenário de modelagem foram comparados com os dados de validação para a obtenção do erro do quadrado médio (EQM). Os modelos que apresentaram os menores valores de EQM foram escolhidos e então submetidos ao cálculo do índice de Willmott (d) (WILLMOTT, 1981) e a porcentagem de viés (PV).

O índice d varia de zero a um, no qual o valor de 1 indica uma acurácia perfeita, e 0 indica não haver qualquer exatidão. A PV possui valor ótimo igual a zero com valores de baixa magnitude indicando boa acurácia, valores positivos indicam superestimação e negativos indicam subestimação do modelo. Obtidos de acordo com as seguintes fórmulas:

$$d - index = 1 - \frac{\sum_{i=1}^n (O_i - S_i)^2}{\sum_{i=1}^n (|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad (10)$$

$$PV = 100 \times \frac{\sum_{i=1}^n (O_i - S_i)}{\sum_{i=1}^n O_i} \quad (11)$$

onde S_i são os valores simulados pelo modelo; O_i , os valores observados; \bar{O} ; a média dos valores observados; e n , o número de observações.

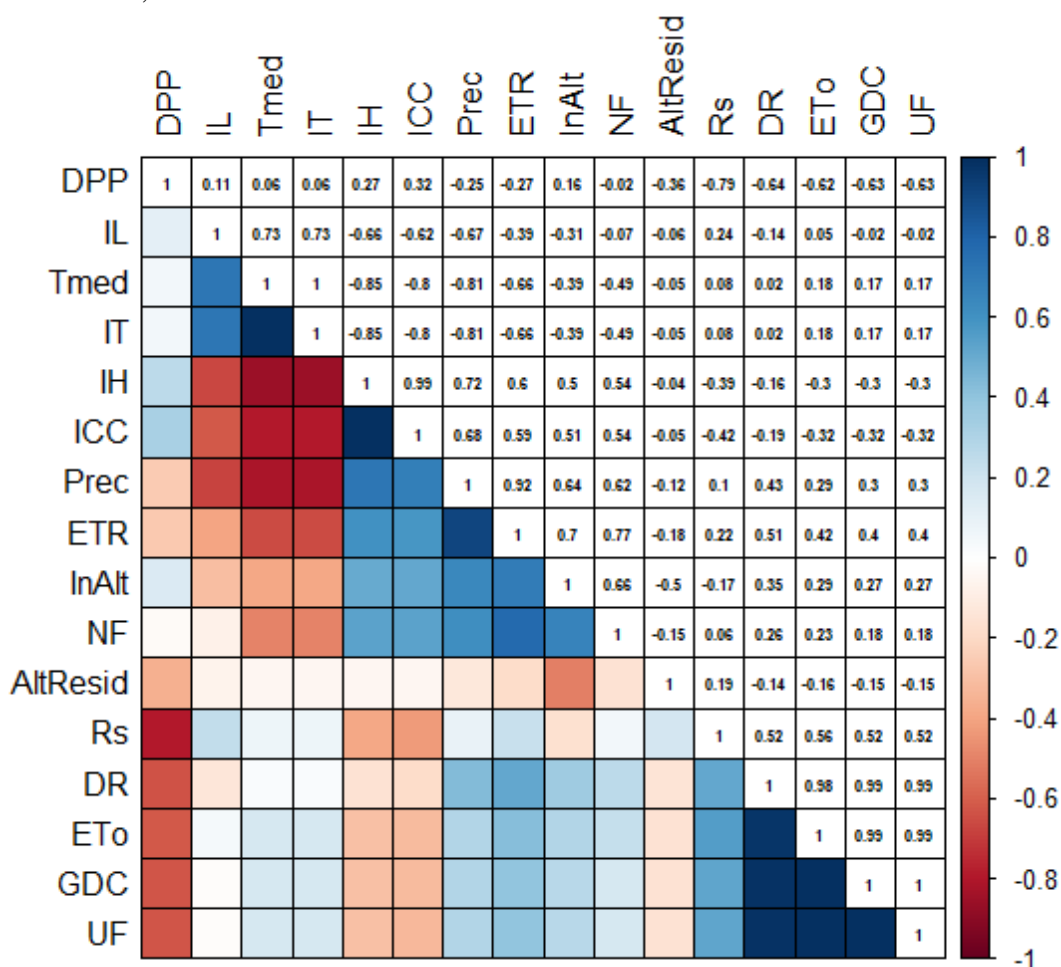
Além disso, foi efetuada a análise de regressão entre os valores observados e simulados com o objetivo de verificar a precisão do modelo por meio do cálculo do coeficiente de determinação e a acurácia do modelo por meio da verificação dos intervalos de confiança dos coeficientes angulares e lineares da equação linear de regressão (TEDESCHI, 2006). Os índices para o processo de avaliação foram obtidos por meio do pacote *hydroGOF* do software R (R CORE TEAM, 2019).

4.3 Resultados

A matriz de correlação de Spearman entre as variáveis (Figura 1) mostra grupos de variáveis que apresentaram forte correlação (acima de 0,80) entre si. As variáveis *Prec*, *IH* e

ICC apresentaram forte correlação negativa com as variáveis IL, Tmed e IT, e a DPP com a Rs. Fortes correlações positivas foram observadas entre as variáveis Prec e ETR, Tmed e IT, IH e ICC, e GDC, UF e DR. Os resultados mostram alguns indícios de multicolinearidade entre as variáveis.

Figura 1 – Matriz de correlação de Spearman entre as variáveis. Abaixo da diagonal principal, quanto mais intensa a coloração vermelha e azul, maior a correlação negativa e positiva, respectivamente, entre as variáveis.



Dentre os algoritmos de aprendizado de máquina usados para a geração dos modelos de previsão, o que apresentou menores valores de EQM foram aqueles gerados a partir do método de mínimos quadrados parciais. Esse método teve uma média da raiz do erro do quadrado médio (REQM) de 542,11 kg de MS ha⁻¹. Depois dele, o método do LASSO adaptativo, com uma REQM de 726,94 kg de MS ha⁻¹, e o método de regressão por componentes principais, com uma REQM de 830,37 kg de MS ha⁻¹. Para o método LASSO adaptativo e de mínimos quadrados parciais, os modelos que consideraram a inclusão das interações bidirecionais apresentaram menor EQM, independente se foi usado ou não as variáveis agrônomicas no

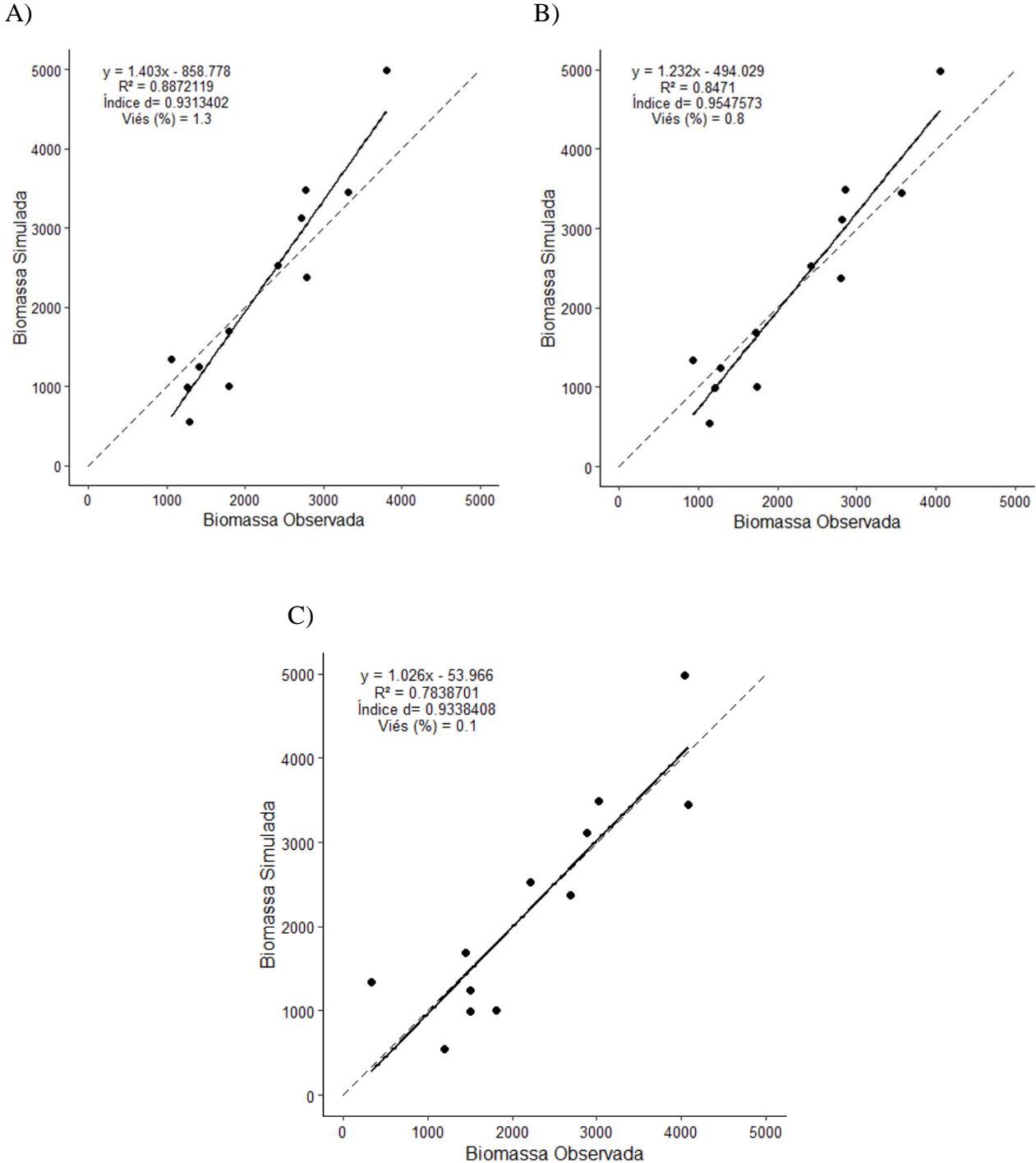
modelo. O método que apresentou maior valor de EQM foi referente aos modelos baseados em árvore de decisão, com uma REQM de 1.290,87 kg de MS ha⁻¹. O melhor método baseado em árvores de decisão foi o algoritmo de árvores ensacadas, que apresentou uma REQM de 953,11 kg de MS ha⁻¹. Considerando os modelos em que se incluiu as variáveis com interação bidirecional, não foi observada uma diferença muito grande em relação aos modelos com e sem interação, com valores de REQM 805,48 e 874,02 kg de MS ha⁻¹, respectivamente. Também não foi observado diferença entre os modelos com e sem a inclusão de variáveis agronômicas nos modelos, com valores de REQM 894,49 e 915,11 kg de MS ha⁻¹, respectivamente (Tabela 1).

Tabela 1 – Valores do erro do quadrado médio do processo de validação dos modelos gerados a partir dos diferentes algoritmos com ou sem o uso de interação bidirecional das variáveis e inclusão de variáveis agronômicas.

Algoritmo	Com variáveis agronômicas		Sem variáveis agronômicas		Média
	Com interação	Sem Interação	Com interação	Sem Interação	
LASSO	880673,00	838579,00	554087,00	653410,00	731687,25
LASSO adaptativo	430313,00	614973,00	446546,00	621941,00	528443,25
Rede Elástica	1096215,00	1012412,00	946682,00	1395784,00	1112773,25
Componentes Principais	926089,80	562590,10	705071,60	1285214,00	689523,95
Mínimos quadrados parciais	223339,10	302628,50	278958,80	351560,00	289121,60
Árvore de decisão	-	1480047,00	-	1852663,00	1666355,00
Árvores ensacadas	-	961980,20	-	854872,90	908426,55
Floresta aleatória	-	1071701,00	-	939840,00	1005770,50

De todos os 26 cenários, os modelos que apresentaram menor EQM, são provenientes do método de mínimos quadrados parciais. Destes, o que apresentou melhor precisão, baseado no coeficiente de determinação entre os valores simulados e observados foi o modelo em que houve a inclusão de variáveis agronômicas sem o uso de interação bidirecional, com valor de R² igual a 0,89 (Figura 2 – A). O modelo mais complexo, que é aquele em que há a inclusão de variáveis agronômicas e efeito da interação foi o que apresentou maior valor do índice de Willmott, igual a 0,95 (Figura 2 – B). O modelo que apresentou menor porcentagem de viés foi o modelo mais simples, aquele em que não há a inclusão de variáveis agronômicas e nem efeito de interação no modelo, a porcentagem de viés foi igual a 0,1 (Figura 2 – C).

Figura 2 – Relação entre os valores de biomassa observados e simulados baseados nos modelos de predição pelo método de mínimos quadrados parciais. A) Modelo com a inclusão de variáveis agronômicas sem o uso de interação bidirecional. B) Modelo com a inclusão de variáveis agronômicas e com o efeito da interação bidirecional. C) Modelo sem a inclusão de variáveis agronômicas e sem o efeito de interação bidirecional.



O modelo em que houve a inclusão de variáveis agronômicas sem o uso de interação bidirecional, apresentou um intervalo de confiança de 95% para o intercepto menor que zero e maior que 1. Já o modelo mais simples e o modelo mais complexo apresentaram os valores zero

para o intercepto e 1 para o coeficiente angular dentro do intervalo de confiança de 95% (Tabela 2). Esses dois modelos apresentam, portanto, valores de parâmetro iguais aos preconizados para um bom processo de validação baseado na análise de regressão entre os valores observados e estimados ($p < 0,05$).

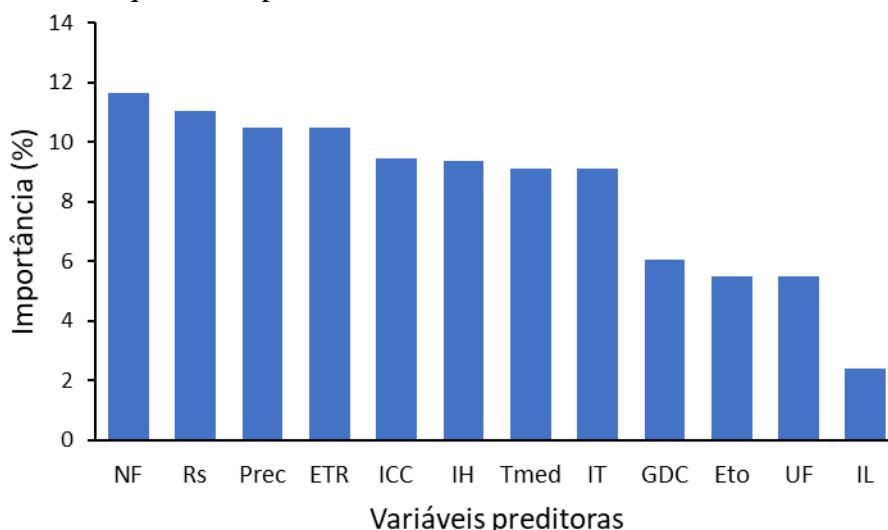
Tabela 2 – Intervalos de confiança de 95% para os valores dos parâmetros da regressão linear entre os valores de biomassa observados e simulados pelos modelos baseadas no método de mínimos quadrados parciais.

Modelo	Parâmetros da regressão de validação	Intervalo de confiança de 95%	
		Inferior	Superior
Modelo 1	Intercepto	-1691,04	-26,51
	Coeficiente angular	1,05	1,75
Modelo 2	Intercepto	-1215,71	227,65
	Coeficiente angular	0,93	1,53
Modelo 3	Intercepto	-996,11	888,18
	Coeficiente angular	0,65	1,40

Modelo 1: Modelo com a inclusão de variáveis agrônômicas sem o uso de interação bidirecional. Modelo 2: Modelo com a inclusão de variáveis agrônômicas e com o efeito da interação bidirecional. Modelo 3: Modelo sem a inclusão de variáveis agrônômicas e sem o efeito de interação bidirecional.

O modelo mais simples, possui vantagem baseado no princípio da parcimônia de modelos. Para esse modelo, com exceção da radiação solar, as variáveis de maior importância são aquelas baseadas na disponibilidade hídrica (Prec, ETR, ICC e IH) mais o NF. As demais variáveis, baseadas no fornecimento energético em termos de radiação e temperatura (Tmed, IT, GDC, Eto, UF e IL) foram menos importantes para a geração do modelo (Figura 3).

Figura 3 – Importância das variáveis preditoras utilizadas na geração do modelo de previsão de biomassa gerado a partir de variáveis agrometeorológicas sem efeito de interação pelo método dos mínimos quadrados parciais.



4.4 Discussão

Nesse estudo demonstrou-se que a abordagem de modelagem empírica baseada em métodos de aprendizagem de máquina é uma opção viável para a geração de modelos de predição de biomassa de forragem usando variáveis agrometeorológicas como preditores, nas condições do estudo. Confirmou-se a hipótese de que modelos empíricos podem ser usados para aplicação em locais de onde os dados não foram usados para a geração do modelo. Os motivos que levam a esses resultados são discutidos a diante.

A forte correlação negativa observada entre as variáveis Prec, IH e ICC em relação as variáveis IL, Tmed e IT, demonstra a caracterização de contrastes entre as variáveis baseadas em recursos hídricos em relação as variáveis baseadas em fornecimento energético via radiação e temperatura sob as condições climáticas de regiões de baixa latitude (cap1). Apesar do ICC ser proveniente do produto de três outros índices (IL, IT e IH), o fato de não haver tanta variação em relação a temperatura e radiação na região norte, demonstra a forte correlação positiva observada entre ICC e IH, o faz com que esse índice faça parte do grupo de variáveis hídricas. A forte correlação positiva também observada entre Prec e ETR é característico das condições de variabilidade em relação a disponibilidade hídrica nos experimentos realizados na região amazônica. Com a diminuição das chuvas, há uma tendência também à diminuição da perda de água por evapotranspiração devido a capacidade de retenção de água pelas plantas e pelo solo (ALLEN, 1998). Por outro lado, variáveis que não apresentaram tanta variação ao longo do período experimental e que são baseadas em soma térmica como o GDC e a UF apresentaram forte correlação positiva com os DR. As altas correlações observadas entre algumas variáveis predictoras (Figura 1) mostra que os métodos de regressão baseados em metodologias que evitem o problema de multicolinearidade são ferramentas de importante aplicação no processo de modelagem baseada em análise de modelos empíricos.

O melhor método para a geração de modelos de previsão foi o método de mínimos quadrados parciais (MQP). Esse método é similar ao método de regressão por componentes principais (RCP). Ambos os métodos são interessantes em casos em que há a existência de multicolinearidade das variáveis predictoras (MARABEL et al., 2013), pois eles são capazes de reduzir o grande número de variáveis colineares medidas a alguns fatores não correlacionados. A principal diferença é que, a RCP decompõe primeiro as variáveis em um conjunto de autovetores e pontuações e, em seguida, faz a regressão em relação à variável resposta como uma etapa separada. Enquanto a MQP usa as informações da variável resposta durante o processo de decomposição (CHO et al., 2007). Por esse motivo, a regressão por MQP tem

apresentado bons resultados de previsão, principalmente em estudos que envolvam a predição de biomassa de forragem que utilizam índices de sensoriamento remoto como variáveis preditoras (MORAIS et al., 2021), pois esses índices muitas vezes são altamente correlacionados (VERRELST et al., 2019).

Em estudos com estimação de biomassa acima do solo em pastagens, geralmente valor médio de R^2 da regressão por MQP é maior que métodos baseados em árvores de decisão, como floresta aleatória (MORAIS, 2021). Como ocorreu com esse estudo, os métodos de modelagem baseados em árvores aleatórias apresentaram piores resultados de previsão do que os demais métodos. Embora os métodos baseados em árvore de decisão sejam diretos e fáceis de interpretar, eles geralmente têm um desempenho insatisfatório em conjuntos de dados independentes, pois superajustam os conjuntos de dados de treinamento usados para criar o modelo (RANSOM et al 2019).

O uso de interações bidirecionais não afetou a precisão dos modelos de previsão. A interação de variáveis nos modelos é um método que melhora os ajustes de modelos nos dados de treinamento, porém não melhoram a precisão de modelos considerando dados de validação. Isso nos dá indícios de que para esse processo de modelagem o uso de modelos mais simples é melhor para previsão. De forma geral, a inclusão de variáveis agronômicas nos modelos não melhorou a precisão no processo de validação. Possivelmente as altas correlações entre as variáveis agronômicas e agrometeorológicas (cap 1), fez com que a inclusão dessas variáveis não gerasse ganhos de previsão dos modelos devido a redundância de informações na variabilidade das variáveis preditoras como um todo. Talvez por isso, o modelo mais simples baseado no método de regressão por MQP resultou na menor porcentagem de viés, o que demonstra sua forte acurácia, apesar de ser o modelo com menor precisão (Figura 2).

É interessante considerar a importância das variáveis preditoras no processo de geração do modelo com menor viés. As variáveis baseadas na disponibilidade hídrica mais a radiação e o NF foram consideradas mais importantes na geração do modelo (Figura 3). Mas isso não significa que variáveis baseadas em temperatura não sejam importantes, pois considerou-se aqui dois locais (região amazônica e semiárida) que apesar de suas diferenças em relação ao solo e regime de chuvas, são regiões que apresentam certa similaridade quanto a ausência de variabilidade em termos de temperatura, radiação e fotoperíodo. Estudos adicionais que levem em consideração regiões com características mais contrastantes, que apresentem maior variação em termos de radiação e temperatura são necessárias para testar e melhorar a capacidade de generalização de métodos de modelagem empírica, baseada em algoritmos de aprendizado de máquina para estimativas de produção de forragem.

4.5 Conclusão

Utilizando dados da região amazônica na geração de modelos empíricos, um modelo simples com a utilização de variáveis agrometeorológicas gerado a partir do método de mínimos quadrados parciais pode ser utilizado na predição de biomassa de capim Tanzânia para a região semiárida.

4.6 Referências

ALLEN, R.G., PEREIRA, L.S., RAES, D., SMITH ROME, M. **Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56**. Fao 300, D05109, 1998.

ALVARES, C. A. et al. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711–728, 1 dez. 2013.

ANDRADE, A. S. et al. Simulating tropical forage growth and biomass accumulation: An overview of model development and application. **Grass and Forage Science**, v. 71, n. 1, p. 54–65, mar. 2016.

ARAÚJO, L. C. et al. Simulating guinea grass production: Empirical and mechanistic approaches. **Agronomy Journal**, v. 105, n. 1, p. 61–69, 2013.

ASHOOR, B. B.; GIWA, A.; HASAN, S. W. **Full-Scale Membrane Distillation Systems and Performance Improvement Through Modeling: A Review**. BASILE, A.; CURCIO, E.; INAMUDDIN (Org.). In: Current Trends and Future Developments on (Bio-) Membranes 1. Ed. Elsevier, 2019, p. 105-140.

BORRA, S.; DI CIACCIO, A. Measuring the prediction error. A comparison of cross-validation, bootstrap and covariance penalty methods. **Computational Statistics & Data Analysis**, v. 54, n. 12, p. 2976–2989, dez. 2010.

CHO, Moses Azong et al. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. **International journal of applied Earth observation and geoinformation**, v. 9, n. 4, p. 414-424, 2007.

CRUZ, P. G. DA et al. Modelos empíricos para estimar o acúmulo de matéria seca de capim-marandu com variáveis agrometeorológicas. **Pesquisa Agropecuária Brasileira**, v. 46, n. 7, p. 675–681, jul. 2011.

DA SILVA JÚNIOR, R. O. et al. Three decades of reference evapotranspiration estimates for a tropical watershed in the eastern Amazon. **Anais da Academia Brasileira de Ciências**, v. 89, n. 3, p. 1985–2002, 2017.

FARIAS, V. D. DA S. et al. Calibration of reference evapotranspiration models in Pará. **Acta Scientiarum. Agronomy**, v. 42, p. e42475, 20 nov. 2019.

FITZPATRICK, E. A.; NIX, H. A. The climatic factor in Australian grassland ecology. In: MOORE, R. M. (Ed.). **Australian grasslands**. Canberra, ACT, Australia: Australian National

University Press, 1970. p. 3–26.

IVORY, D.; WHITEMAN, P. Effect of Temperature on Growth of Five Subtropical Grasses. I. Effect of Day and Night Temperature on Growth and Morphological Development. **Functional Plant Biology**, v. 5, n. 2, p. 131, 1978.

LOPES, M. N. et al. Tillering dynamics in massai grass fertilized with nitrogen and grazed by sheep. **Bioscience Journal**, p. 446–454, 2016.

MARABEL, Miguel; ALVAREZ-TABOADA, Flor. Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. **Sensors**, v. 13, n. 8, p. 10027-10051, 2013.

MCMASTER, G. Growing degree-days: one equation, two interpretations. **Agricultural and Forest Meteorology**, v. 87, n. 4, p. 291–300, dez. 1997.

MORAIS, T. G., TEIXEIRA, R. F. M., FIGUEIREDO, M., DOMINGOS, T. The use of machine learning methods to estimate aboveground biomass of grasslands: A review. **Ecological Indicators**. v. 130, p. 108081, 2021.

MORENO, L. S. B. et al. Base temperature determination of tropical Panicum spp. grasses and its effects on degree-day-based models. **Agricultural and Forest Meteorology**, v. 186, p. 26–33, mar. 2014.

PEZZOPANE, J. R. M. ACED. et al. Simple agrometeorological models for estimating Guineagrass yield in Southeast Brazil. **International journal of biometeorology**, v. 58, n. 7, p. 1479–1487, 2013.

PEZZOPANE, J. R. M. et al. An integrated agrometeorological model to simulate Marandu palisade grass productivity. **Field Crops Research**, v. 224, n. May, p. 13–21, jul. 2018.

PEZZOPANE, J. R. M. et al. Dry matter production of Tanzania grass as a function of agrometeorological variables. **Pesquisa Agropecuaria Brasileira**, v. 47, n. 4, p. 471–477, 2012.

PEZZOPANE, J. R. M. et al. Panicum maximum cv. Tanzânia: climate trends and regional pasture production in Brazil. **Grass and Forage Science**, v. 72, n. 1, p. 104–117, 2017.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. 2019. Disponível em: <https://www.R-project.org/>. Acesso em 05 set. 2019.

RANSOM, Curtis J. et al. Statistical and machine learning methods evaluated for incorporating soil and weather into corn nitrogen recommendations. **Computers and Electronics in Agriculture**, v. 164, p. 104872, 2019.

SAS Institute Inc. 2018. **SAS/STAT® 15.1 User's Guide**. Cary, NC: SAS Institute Inc.

SOLCAST. **Global solar irradiance data and PV system power output data**. Disponível em: <<https://solcast.com/>>. Acesso em: 15 jan. 2020.

THORNTHWAITE, C. W.; MATHER, J. R. **The water balance**. New Jersey: Publications in climatology, 1955. v. 8

TIBSHIRANI, R. Regression Shrinkage and Selection Via the Lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 58, n. 1, p. 267–288, jan. 1996.

TONATO, F. et al. Desenvolvimento de modelos preditores de acúmulo de forragem em pastagens tropicais. **Pesq. agropec. bras.**, v. 45, n. 5, p. 522–529, 2010.

TURC, L. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. **Ann. Agron**, v. 12, n. 1, p. 13–49, 1961.

VERRELST, Jochem et al. Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. **Surveys in Geophysics**, v. 40, n. 3, p. 589–629, 2019.

VILLA NOVA, N. A., TONATO, F., PEDREIRA, C. G. S., MEDEIROS, H. R. Método alternativo para cálculo da temperatura base de gramíneas forrageiras. **Ciência Rural**, v. 37, n. 2, p. 545–549, 2007.

WANG, E.; ENGEL, T. Simulation of phenological development of wheat crops. **Agricultural Systems**, v. 58, n. 1, p. 1–24, set. 1998.

WILLMOTT, C. J. On The Validation of Models. **Physical Geography**, v. 2, n. 2, p. 184–194, 15 jul. 1981.

ZOU, H. The Adaptive Lasso and Its Oracle Properties. **Journal of the American Statistical Association**, v. 101, n. 476, p. 1418–1429, dez. 2006.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 67, n. 5, p. 768–768, nov. 2005.