

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**APRENDIZADO DE MÁQUINA NA PREDIÇÃO DA
PRODUTIVIDADE E MAPEAMENTO DA QUALIDADE DA
FIBRA EM CAMPOS COMERCIAIS DE ALGODÃO**

**Francielle Morelli Ferreira
Engenheira Agrícola**

2022

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**APRENDIZADO DE MÁQUINA NA PREDIÇÃO DA
PRODUTIVIDADE E MAPEAMENTO DA QUALIDADE DA
FIBRA EM CAMPOS COMERCIAIS DE ALGODÃO**

Francielle Morelli Ferreira

Orientador: Prof. Dr. Rouverson Pereira da Silva

Coorientadores: Prof. Dr. Glauco de Souza Rolim

Prof. Dr. Luciano Shozo Shiratsuchi

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências do título de Doutora em Agronomia (Ciência do Solo).

2022

F383a Ferreira, Francielle Morelli
Aprendizado de máquina na predição da produtividade e mapeamento da qualidade da fibra em campos comerciais de algodão / Francielle Morelli Ferreira. -- Jaboticabal, 2022
114 p. : il., tabs., fotos, mapas

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Rouverson Pereira da Silva

1. Agricultura de precisão. 2. Sensoriamento remoto por satélite. 3. Algoritmos de computador. 4. Rastreabilidade. 5. Classificação da Fibra. I. Título.

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

Essa ficha não pode ser modificada.

REGISTRO DE IMPACTO

A presente tese aborda a difusão de tecnologia aplicada à agricultura sustentável, apresentando resultados inovadores e satisfatórios que poderão ser replicados. A internacionalização (doutorado sanduíche) com coleta e processamento de dados nos EUA, a participações em eventos com apresentações orais de trabalhos em inglês contribuirão nas publicações em revistas internacionais de alto impacto.

IMPACT RECORD


This thesis approaches the diffusion of technology applied to sustainable agriculture, presenting innovative and satisfactory results that can be replicated. Internationalization (sandwich Ph.D. program) with data collection and processing in the US and participation in conferences with English oral presentations will contribute to publications in high-impact international journals.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: APRENDIZADO DE MÁQUINA NA PREDIÇÃO DA PRODUTIVIDADE E MAPEAMENTO DA QUALIDADE DA FIBRA EM CAMPOS COMERCIAIS DE ALGODÃO

AUTORA: FRANCIELLE MORELLI FERREIRA
ORIENTADOR: ROUVERSON PEREIRA DA SILVA
COORIENTADOR: GLAUCO DE SOUZA ROLIM
COORIENTADOR: LUCIANO SHOZO SHIRATSUCHI

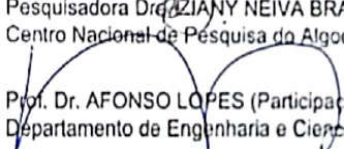
Aprovada como parte das exigências para obtenção do Título de Doutora em Agronomia (Ciência do Solo), pela Comissão Examinadora:



Prof. Dr. ROUVERSON PEREIRA DA SILVA (Participação Virtual)
Departamento de Engenharia e Ciências Exatas DECEX / FCAV UNESP Jaboticabal



Pesquisadora Dra. ZIANY NEIVA BRANDÃO (Participação Virtual)
Centro Nacional de Pesquisa do Algodão / EMBRAPA - Campina Grande/PB



Prof. Dr. AFONSO LOPES (Participação Virtual)
Departamento de Engenharia e Ciências Exatas DECEX / FCAV UNESP Jaboticabal

Prof. Dr. RENILDO LUIZ MION (Participação Virtual)
Universidade Federal de Rondonópolis / Rondonópolis/MT

gov.br

Documento assinado digitalmente

RENILDO LUIZ MION

Data: 30/01/2023 10:49:53:0399

Verifique em <https://verificador.iti.br>

Mailson Freire de Oliveira
Pós-doutorando MAILSON FREIRE DE OLIVEIRA (Participação Virtual)
Auburn University-College of Agriculture / Estados Unidos

Jaboticabal, 09 de dezembro de 2022

DADOS CURRICULARES DA AUTORA

FRANCIELLE MORELLI FERREIRA – nascida em Cuiabá, Mato Grosso, Brasil, no dia 13 de março de 1989. Ingressou no ensino superior em agosto de 2006, sendo aprovada na primeira turma do curso de Engenharia Agrícola e Ambiental da Universidade Federal de Mato Grosso, Câmpus de Rondonópolis. Durante a graduação foi monitora na disciplina de estatística e de agosto de 2009 a julho de 2011 foi bolsista de iniciação científica pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), em projeto na área de Máquinas e Implementos Agrícolas, sob a orientação do Prof. Dr. Antonio Renan Berchol da Silva. Em agosto de 2011 iniciou o curso de Mestrado no Programa de Pós-graduação em Agronomia (Energia na Agricultura) na Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), em Botucatu - SP, sendo bolsista CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), sob a orientação do Prof. Dr. Sérgio Hugo Benez, e coorientação do Prof. Dr. Paulo Arbex e Prof. Dr. Antonio Renan Berchol da Silva. Foi integrante do Grupo de Plantio Direto – GPD, obtendo o título de Mestra em Agronomia em julho de 2013, com dissertação na área de Máquinas Agrícolas. Em dezembro de 2013 foi aprovada em primeiro lugar no Concurso Público de Provas e Títulos para Provimento do Cargo de Docente da Educação Superior na UNEMAT - Universidade do Estado de Mato Grosso, Câmpus de Alta Floresta, área de Engenharia Agrícola. Em março de 2014 iniciou a carreira Docente junto ao curso de Agronomia com as disciplinas de Desenho Técnico, Máquinas e Mecanização Agrícola, e Agricultura de Precisão. Em agosto de 2017 foi removida a pedido para a UNEMAT - Câmpus Universitário de Nova Mutum, sendo responsável pelas mesmas disciplinas já mencionadas. Durante o ano de 2018 atuou como Coordenadora do Curso de Agronomia. Em 2019 conseguiu junto à UNEMAT o afastamento para qualificação docente e iniciou o curso de Doutorado no Programa de Pós-Graduação em Agronomia (Ciência do Solo) na UNESP, em Jaboticabal - SP, sob a orientação do Prof. Dr. Rouverson Pereira da Silva. Foi integrante do Laboratório de Máquinas e Mecanização Agrícola (LAMMA) e do Rouverson Silva Research Group (RSRG). De setembro de 2021 a abril de 2022 participou como “J-1 Exchange Visitor” do Programa de Doutorado Sanduíche na Louisiana State University Agricultural Center em Baton Rouge, LA, EUA sob a supervisão e coorientação do Prof. Dr. Luciano Shozo Shiratsuchi. Em 09 de dezembro de 2022, submeteu-se à banca examinadora para obtenção do título de Doutora em Agronomia, com tese na área de Agricultura Digital.

“A tarefa não é tanto ver aquilo que ninguém viu,
mas pensar o que ninguém ainda pensou sobre
aquilo que todo mundo vê.”

Arthur Schopenhauer

“É muito melhor lançar-se em busca de
conquistas grandiosas, mesmo expondo-se ao
fracasso, do que alinhar-se com os pobres de
espírito, que nem gozam muito nem sofrem
muito, porque vivem numa penumbra cinzenta,
onde não conhecem nem vitória, nem derrota.”

Theodore Roosevelt

Em primeiro lugar:

A **Deus** e nosso **Senhor Jesus Cristo**, pela vida, saúde. pelas oportunidades e por suas constantes bênçãos e proteção.

Agradeço!

Ao meu filho **Matheus Morelli Ferreira de Aquino**, que desde 2007 é minha fonte de força e alegria para vencer e continuar caminhando, sem desanimar, em busca do melhor, para que um dia, as minhas vitórias e conquistas possam servir de inspiração, orgulho e exemplo para ele.

Homenageio!

Aos meus pais, **José Francisco Ferreira Filho** e **Marta Morelli Ferreira**, à minha irmã **Michelle Faifer Morelli Ferreira**: Pela confiança, apoio, admiração, ajuda, incentivo, compreensão, dedicação, amor, carinho, cuidado e orgulho, em mim depositados e principalmente por compreenderem a importância dessa conquista e aceitarem a minha ausência em diversos momentos ao longo desses 4 anos.

Dedico!

À **Universidade do Estado de Mato Grosso (UNEMAT)**, instituição a qual pertencço e me dedico desde 2014 com muito orgulho, e que me possibilitou a realização de tantos sonhos, inclusive este “doutorado”, me proporcionando três anos de afastamento remunerado em prol da qualificação do seu corpo docente.

Ofereço!

AGRADECIMENTOS

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências Agrárias e Veterinárias, Câmpus de Jaboticabal, em especial ao Programa de Pós-Graduação em Agronomia (Ciência do Solo).

Ao meu companheiro Getulio de Freitas Seben Junior, por todo o apoio, confiança, paciência, carinho e amor destinados a mim ao longo desses 4 anos. Obrigada por ter sido o esteio do nosso relacionamento, pois enquanto eu “voava”, você “criava raízes”. Gratidão pelo cuidado e atenção com o meu filho nos diversos momentos em que estive ausente.

Ao meu querido orientador Prof. Dr. Rouverson Pereira da Silva, por toda a orientação e confiança depositada em mim. Gratidão por ser quem és, ao qual tenho profundo respeito e admiração. Obrigada pela oportunidade de ter feito parte do seu time e por toda a compreensão para comigo. És um orientador e líder de peso, que levarei por toda a vida, a gratidão será eterna. Sempre que orientar, lembrarei do Senhor. Obrigada por ter sido esse “paizão” durante a “saga doutorado”, tornando-a tão mais leve sempre com palavras de conforto e incentivo!

Ao meu coorientador Prof. Dr. Glauco Rolim que me apresentou o caminho da programação em Python, me ajudou em diversos momentos de reuniões, e hoje busco trilhar uma caminhada como uma jovem cientista de dados da agricultura digital. Uma área pelo qual me apaixonei no doutorado devido ao seu apoio e coorientação.

Ao meu coorientador Prof. Dr. Luciano Shozo Shiratsuchi por ter me oportunizado a realização de um dos meus sonhos, o doutorado sanduíche nos EUA. Obrigada por ter me recebido tão bem em seu grupo de pesquisa na LSU AgCenter, em Baton Rouge, LA, USA. Por todo auxílio prestado enquanto estive nos EUA, pela bolsa concedida, pelos ensinamentos transmitidos. Agradeço o auxílio e oportunidade em ter um capítulo da tese com dados coletados nos EUA. Gratidão por esta oportunidade!

Aos membros da banca examinadora Dra. Ziany Neiva Brandão, Prof. Dr. Renildo Luiz Mion, Prof. Dr. Afonso Lopes e Dr. Mailson Freire de Oliveira pela disposição em participar desse importante momento e por toda contribuição dada para a melhoria deste trabalho.

A todos os Professores que por meio de disciplinas, transmitiram um pouco do seu conhecimento a mim durante o doutorado - à Profa. Dra. Teresa Cristina Tarlé Pissara, ao Prof. Dr. Carlos Eduardo Furlani, ao Prof. Dr. Tiago Rodrigo Francetto, ao Prof. Dr. Walter Maldonado Junior, ao Prof. Dr. José Carlos Barbosa, ao Prof. Dr. Glauco de Souza Rolim, ao Prof. Dr. Murilo Voltarelli e ao Prof. Dr. Rouverson Pereira da Silva.

À Empresa Terra Santa Agro, e a toda a equipe da Fazenda Mãe Margarida (Santa Rita do Trivelato – MT), em especial aos especialistas em Agricultura de Precisão, Davi Besson, Paulino Vinícius S. Gajardo e Renan Tavanti, pela parceria na pesquisa do segundo capítulo da tese, pelo constante pronto atendimento em prol ao desenvolvimento experimental.

Ao grupo GMS Agro e toda a equipe da Fazenda Santa Maria (Sorriso – MT), em especial ao Eng. Agrônomo Evandro Bortolussi responsável pela Agricultura Digital da fazenda. Obrigada por terem acreditado na ideia e pela parceria nas coletas em campo do terceiro capítulo desta tese. Agradeço o apoio no manuseio e envio das amostras para classificação de fibra junto ao Laboratório Kuhlmann de Classificação de Fibra (Sorriso, Brasil).

Às minhas amigas Samira Luns H. Almeida e Franciele Morlin Carneiro que o doutorado me proporcionou conhecer, e à minha amiga Elizabeth Haruna Kazama, que o doutorado proporcionou reencontrar, meu muito obrigada pela amizade, companheirismo, parceria, conversas e mais conversas, muitas risadas, lágrimas, incentivo, apoio, e parceria durante toda essa jornada. Vocês são especiais!

Aos colegas e amigos do Laboratório de Máquinas e Mecanização Agrícola (LAMMA) - Alex, Aline, Armando, Bruno Rocca, Danilo, Edcarla, Elizabeth, Fran Morlin, Gustavo, Jarlysson, Jean, João, Letícia, Marcelo, Mariana, Nivanda, Renata, Rafael, Samira, Titi, Vinícius, meu muito obrigada por terem passado pela minha caminhada!

Aos colegas que fiz na Pós-Graduação, em especial à Nayane Maia e ao Danilo Tedesco, aos quais me ajudaram muito com a construção do conhecimento para alcançar o segundo capítulo desta tese. Gratidão!

Aos colegas e amigos que tive na temporada “doutorado sanduíche” nos Estados Unidos – Leticia, Mari, Melina, Tabata, Ju, Cesar, Samuel, Aline, Murilo,

Philip, Fagner, Mari, Eduardo, Julia, Rodrigo e em especial à Lilian e ao Mauricio pela enorme parceria no último mês.

À Universidade Federal de Mato Grosso, Câmpus Rondonópolis (atual UFR), pela formação em Engenharia Agrícola e Ambiental e à FCA/Unesp de Botucatu – SP, pelo Mestrado em Agronomia (Energia na Agricultura) que me possibilitaram chegar até aqui.

Ao meu orientador de graduação Prof. Dr. Antonio Renan Berchol por todo ensinamento e valores repassados a mim, por quem tenho tamanha admiração e sou extremamente grata por ter me apresentado a área de Máquinas e Mecanização Agrícola. Sou grata também por ter acreditado no meu potencial, e por ter me incentivado e apoiado a ingressar no mestrado na UNESP em Botucatu, ao qual me apaixonei mais ainda pela área, que me possibilitou estar hoje aqui defendendo esta tese.

A todos os meus familiares (avós, tias(os), primas(os)) pertencentes às famílias Morelli e Ferreira que, de alguma forma, torceram para essa conquista.

A todos meus amigos, colegas e conhecidos que fiz “nesta longa estrada da vida” e que de alguma forma torceram para essa conquista.

A todos que contribuíram de forma direta e indiretamente, os meus sinceros agradecimentos.

SUMÁRIO

RESUMO.....	i
ABSTRACT	ii
CAPÍTULO 1 – Considerações Gerais	14
INTRODUÇÃO	14
REVISÃO DE LITERATURA	18
Panorama da cultura do algodão.....	18
Colheita do algodão.....	20
A qualidade da fibra do algodão	23
Rastreabilidade da qualidade da fibra	27
Agricultura de Precisão e Agricultura Digital.....	30
Aprendizado de máquina na agricultura	32
Sensoriamento remoto na agricultura.....	34
Controle estatístico do processo.....	37
Referências	38
CAPÍTULO 2 - Comparação de algoritmos de aprendizado de máquina na previsão de produtividade do algodão em campos comerciais utilizando sensoriamento remoto por satélite.....	47
RESUMO.....	47
ABSTRACT	48
Introdução	49
Material e Métodos.....	52
Campos de estudo.....	52
Aquisição e processamento dos dados das imagens de satélite.....	54
Fenologia e soma térmica do algodão.....	54
Índices de vegetação – preditores de entrada	55

Produtividade observada – variável y que o modelo tentará prever	57
Correlação, seleção e padronização das variáveis preditoras.....	58
Algoritmos de <i>machine learning</i> para predição de produtividade do algodão.....	58
Avaliação do desempenho dos algoritmos	60
Resultados e Discussão.....	61
Produtividade observada x variáveis preditoras	61
Seleção e padronização das variáveis preditoras.....	64
Algoritmos x Predição de produtividade do algodão.....	65
Considerações Finais.....	68
Conclusão	69
Referências	69
CAPÍTULO 3 – É possível mapear os parâmetros de qualidade da fibra de algodão em campos comerciais?.....	74
RESUMO.....	74
ABSTRACT	75
Introdução	76
Material e Métodos.....	80
Etapas e áreas de estudo.....	80
Colheita e Produtividade do algodão	82
Métodos de coleta de algodão em caroço	83
Beneficiamento e classificação da fibra do algodão	84
Mapeamento da qualidade da fibra	85
Qualidade do processo de classificação de fibras	87
Resultados e Discussão.....	88
Produtividade do algodão	88
Indicadores de qualidade da fibra do algodão	89
Mapeamento da qualidade da fibra	96

Conclusões.....	103
Referências	103
CAPÍTULO 4 – Considerações finais	109
APÊNDICES	112
Apêndice A. Valores dos hiperparâmetros utilizados pelos algoritmos avaliados na predição de produtividade em cada cenário e campo de estudo do segundo capítulo desta tese.....	112
Apêndice B. Tabela de ágios e deságios (R\$/kg) em função Da Classificação Universal para comercialização de fibras de algodão upland produzidas no Brasil. Fonte: CONAB/MOC (2022).	113
Apêndice C. Tabela de prêmios e descontos (pontos por libra) em função dos Padrões Universais dos Estados Unidos para comercialização de algodão upland em 2021. Fonte: USDA (2021).....	114

APRENDIZADO DE MÁQUINA NA PREDIÇÃO DA PRODUTIVIDADE E MAPEAMENTO DA QUALIDADE DA FIBRA EM CAMPOS COMERCIAIS DE ALGODÃO

RESUMO - A agricultura digital faz uso intensivo de tecnologias proporcionando a utilização de técnicas de aprendizado de máquina na predição da produtividade de culturas com base em sensoriamento remoto. Nesse contexto, este trabalho foi elaborado com o intuito de auxiliar produtores e gerenciadores de dados na antecipação de informações sobre a produtividade do algodão, bem como explorar tecnologias presentes nas colhedoras de algodão que proporciona a rastreabilidade e o mapeamento da qualidade da fibra do algodão em campos comerciais. Dessa forma, o trabalho foi constituído de quatro fases distintas, sumarizadas em quatro capítulos que originaram a tese. O primeiro se trata de um referencial teórico aos temas abordados. O segundo capítulo apresenta resultados inéditos de comparação do desempenho de algoritmos de aprendizado de máquina (K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Random Forest, Linear Regression, e AutoML) na predição da produtividade do algodão baseado em índices de vegetação (IV) em campos comerciais no Mato Grosso. O terceiro capítulo fundamenta-se na exploração de tecnologia existente nas colhedoras de algodão (Radio Frequency Identification - RFID), porém pouco explorada, com resultados coletados em duas etapas (Brasil e EUA), apresentando abordagens inéditas, que objetiva vincular as informações de localização de colheita dos módulos com as informações de classificação da fibra recebidas na algodoeira e, por meio da geoespacialização das áreas colhidas dos módulos, resultar na geração de mapas de parâmetros de qualidade da fibra. Para alcançar o objetivo do segundo capítulo que foi identificar o quão cedo, qual índice e com qual algoritmo consegue-se prever a produtividade do algodão no âmbito de fazenda, otimizou-se a predição e por seleção de variáveis (stepwise), avaliou-se apenas o melhor IV em cada campo, e detectou-se que dentre os oito IVs, o Simple Ratio, impulsionado pelo algoritmo KNN apresentou a melhor performance, predizendo com apenas 0.26 e 0.28 t ha⁻¹ de RMSE e MAPE de 5.20%; antecipando com baixo erro em ±143 dias a produtividade de algodão. No terceiro capítulo, concluiu-se que é possível mapear os parâmetros de qualidade da fibra em campos comerciais, desde que se vincule os dados de identificação do módulo com a etiqueta criada na algodoeira para os fardos beneficiados. Os parâmetros de qualidade da fibra apresentam variabilidade em cada módulo e em cada método de coleta, para uma mesma área, e a qualidade do processo de classificação da fibra apresentou pontos fora de controle de maneira isolada. O mapeamento da qualidade da fibra auxiliou na visualização em campo dos parâmetros que sofreram ágios e deságios na comercialização devido à sua classificação atender ou não os padrões do mercado nacional ou internacional.

Palavras-chave: algoritmo de regressão, classificação HVI, imagem de satélite, índice de vegetação, módulo de algodão, rastreabilidade

MACHINE LEARNING COTTON YIELD PREDICTION AND FIBER QUALITY MAPPING IN COMMERCIAL FIELDS

ABSTRACT - Digital agriculture intensively uses technologies to improve machine learning techniques on yield predictions based on remote sensing. Following the context, this research was written to inform farmers and data managers about cotton yield information beforehand. Also, to explore embedded technologies such as Radio Frequency Identification (RFID) in cotton harvesters that provide traceability and mapping of cotton fiber quality on commercial fields in Mato Grosso State, Brazil. This thesis was divided into a four-chapter: the first treats a theoretical reference around themes covered by this research. The second chapter presents unprecedented results comparing the machine learning algorithms' performance (k-Nearest-Neighbours - KNN, Multilayer Perceptron - MLP, Random Forest Linear Regression - RF and Automatic Machine Learning - AutoML) in cotton yield prediction based on vegetation indices (VI). The third chapter explored existing technology to analyze data from Brazil and the USA about the rarely used cotton harvesters (RFID), presenting distinctive approaches and linking information from the area harvested per module with the fiber quality from the gin to generate geospatial maps of fiber quality parameters. To achieve the objective in the second chapter, the prediction was optimized by the variables selection (stepwise) using only the most essential VI among eight VI in each field. As a result, it was detected that Simple Ratio, aided by KNN, showed the best metrics predicting with only 0.26 and 0.28 t ha⁻¹ of RMSE and MAPE of 5.20%, anticipating cotton yield by ± 143 days with low error. Finally, in the third chapter, it was observed that it is possible to map the fiber quality parameters in commercial fields if the module identification data is linked to the label created in the cotton plant for the processed bales. However, the fiber quality parameters showed variability in each module and each collection method for the same area, and the quality of the fiber classification process presented points out of control in an isolated way. Nevertheless, quality mapping is possible and assists in visualizing the parameters that suffered premiums and discounts in the commercialization due to its classification meeting or not the national or international market standards.

Keywords: regression algorithm, HVI classification, satellite image, vegetation index, cotton module, traceability

CAPÍTULO 1 – Considerações Gerais

INTRODUÇÃO

O algodão é uma fonte crucial de fibra natural, e os Estados Unidos da América (EUA) é o maior exportador mundial da fibra (USDA, 2022), entretanto, a cotonicultura brasileira vem crescendo e em 2019 posicionou o Brasil como o segundo maior exportador mundial.

Novas técnicas de gerenciamento do campo, investimentos em tecnologias embarcadas, pesquisa e melhor beneficiamento aumentaram a qualidade e a quantidade do algodão brasileiro, adequando-o às necessidades das indústrias internacionais (Severino et al., 2019).

Nos EUA e Brasil a colheita do algodão é totalmente mecanizada e altamente tecnológica, entretanto, mesmo em fazendas que possuem máquinas com sensor de rendimento, ainda se observa a utilização de métodos manuais, considerados morosos, para estimar a produção antes da operação de colheita. Diante disso, é necessário recorrer a métodos alternativos que tornem a atividade mais rápida, como proposto no capítulo 2 do presente trabalho.

A tecnologia na colheita mecanizada do algodão evoluiu de tal forma que as novas máquinas fornecem rastreabilidade e a documentação da colheita em campo estão em ampla adoção. Segundo Franco (2020) os módulos redondos originados pelas colhedoras JD7760, CP690 e JD770 são a maioria do algodão colhido e processado por descaroçadores nas regiões produtoras de algodão dos EUA e do Brasil, e contabilizam cerca de 80% em substituição aos antigos módulos retangulares.

Mesmo com esse nível de tecnologia, a rastreabilidade e o completo gerenciamento dos módulos em campo atrelados à classificação da fibra não vêm sendo realizados. Isso porque, para interrelacionar a qualidade da fibra ao local de colheita desta, as fazendas precisam vincular as informações dos módulos com as etiquetas que as algodozeiras criam no beneficiamento desses módulos.

Além de máquinas agrícolas, plataformas digitais como imagens de satélite, parâmetros climáticos/meteorológicos e séries temporais para previsões de produtividade são formas de monitorar e adquirir dados.

Algoritmos de machine learning (aprendizado de máquina) colocam em prática a ideia de que as máquinas podem aprender parcial ou totalmente de forma autônoma diretamente dos dados de entrada. Como resultado, eles são promissores para analisar dados agrícolas em grande volume (big data), agindo de forma mais rápida, eficiente e precisa.

Na colheita do algodão, a tecnologia evoluiu de tal forma que as colhedoras de algodão dotadas da tecnologia RFID (Radio Frequency Identification) em conjunto com o Rastreador de Colheita (Harvest Identification Cotton), fornece diversas informações como a variedade, coordenadas geográficas do trajeto dos módulos colhidos, umidade, peso, área colhida e produtividade.

Diante do exposto, este trabalho foi elaborado com o intuito de contribuir com a cotonicultura, possibilitando auxiliar produtores e gerenciadores de dados na predição de produtividade e rastreabilidade da qualidade da fibra em campos comerciais de algodão. Como consequência espera-se um ganho de tempo no planejamento, seja na intervenção do manejo à colheita do algodão, e até mesmo na da comercialização.

Dessa forma, o trabalho foi constituído de quatro fases distintas, sumarizadas em quatro capítulos.

No primeiro capítulo foi realizado um referencial teórico abordando a importância de alternativas para estimar a produtividade do algodoeiro utilizando técnicas de aprendizado de máquina com base em sensoriamento remoto por satélite. Além disso, foi abordada a importância da exploração da tecnologia RFID presentes nas recentes colhedoras de algodão, que, com a integração de dados de campo e algodoeira, possibilitam o mapeamento da qualidade da fibra em campos comerciais de algodão.

No segundo capítulo, são apresentados resultados inéditos de comparação do desempenho de algoritmos de machine learning (K-Nearest Neighbors, Multilayer Perceptron, Random Forest, Linear Regression e AutoML) na predição da produtividade do algodão baseado em sensoriamento remoto por satélite. O estudo

foi realizado em 2019 em campos comerciais em propriedade agrícola no estado do Mato Grosso, e utilizou-se imagens do satélite Sentinel 2A adquiridas durante o ciclo do algodoeiro, que originaram índices de vegetação avaliados no algodoeiro das três áreas de estudo (Figura 1).

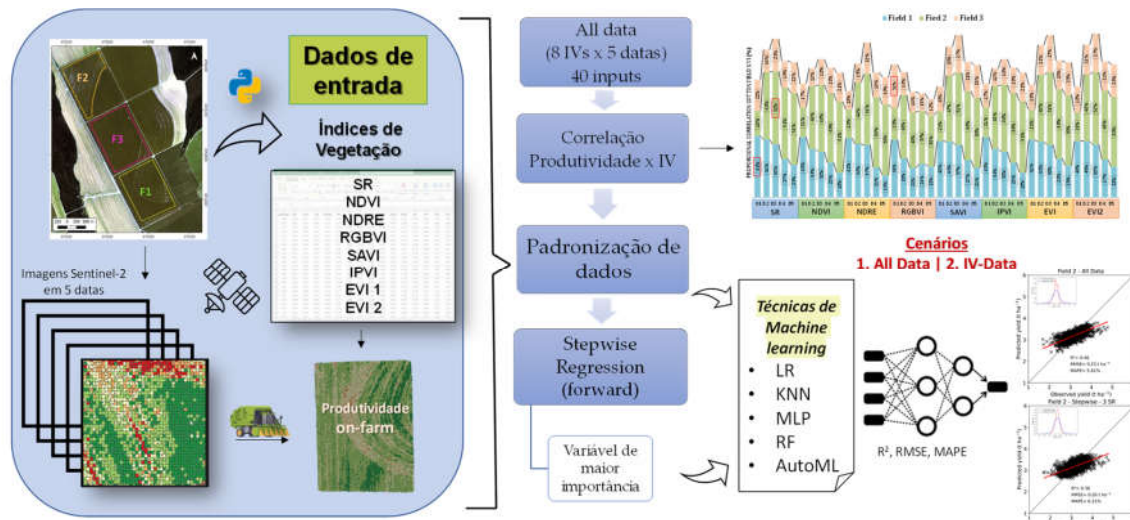


Figura 1. Fluxograma das etapas percorridas na previsão da produtividade do algodão utilizando algoritmos de machine learning e índices de vegetação, apresentadas no Capítulo 2.

No terceiro capítulo deste trabalho apresenta-se um estudo inédito realizado em duas etapas (Brasil e EUA), utilizando-se de tecnologia existente nas colhedoras de algodão (RFID), porém pouco explorada no Brasil, que objetiva vincular as informações de localização de colheita dos módulos de algodão com as informações de classificação de fibra recebidas na algodoeira.

Para isso, métodos de amostragem de algodão para avaliar os parâmetros de qualidade da fibra foram testados (coletas da máquina em operação, coletas do módulo em campo, e coletas na algodoeira), avaliados em relação ao controle estatístico do processo por meio de cartas de controle, e por meio da geoespacialização da área colhida dos módulos, mapas de parâmetros de qualidade da fibra em campos comerciais foram gerados (Figura 2).

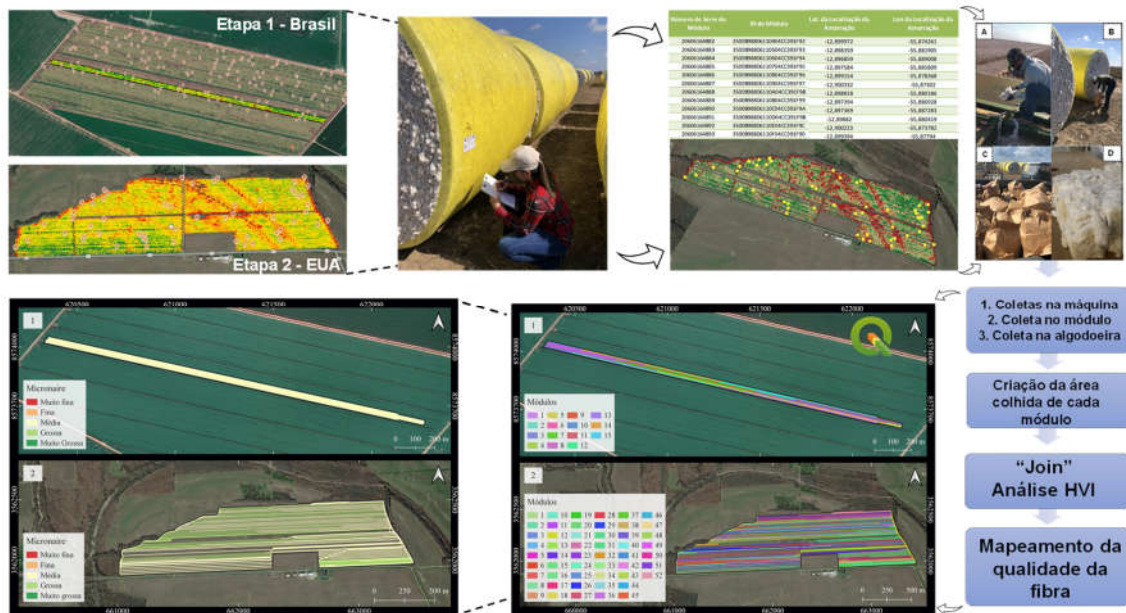


Figura 2. Fluxograma das etapas realizadas para o mapeamento dos parâmetros de qualidade da fibra do algodão, apresentadas no Capítulo 3.

A exploração das ferramentas da agricultura digital abordadas nos capítulos dois e três visam auxiliar produtores e gerenciadores de dados nas tomadas de decisões, sendo assim, no capítulo quatro são apresentadas as considerações finais deste trabalho, os quais demonstram os principais resultados obtidos de forma objetiva e coesa, bem como recomendações e perspectivas de estudos futuros.

Pressupondo que índices de vegetação sejam preditores da produtividade de algodão utilizando aprendizagem de máquina supervisionada e que os parâmetros de qualidade da fibra podem ser rastreados e identificados na resolução de módulo, atrelado ao fato de que a exploração dessas ferramentas auxiliarão o produtor nas tomadas de decisões, este trabalho teve como objetivo geral o desenvolvimento de técnicas de auxílio na predição da produtividade e melhoria da rastreabilidade da qualidade da fibra do algodoeiro em campos comerciais.

REVISÃO DE LITERATURA

Panorama da cultura do algodão

O algodão é uma fonte crucial de fibra natural e desempenha papel importante na economia de diversos países. Os maiores produtores mundiais de algodão são: Índia, China, Estados Unidos da América (EUA), Brasil e Paquistão (USDA, 2021a).

Os Estados Unidos (EUA) e o Brasil são países que se destacam tanto em produção como em exportação de algodão. Os EUA é o segundo produtor mundial e o maior exportador, o Brasil é o quarto maior produtor e segundo exportador mundial de algodão (USDA, 2021a), com objetivo de se tornar o maior exportador mundial até 2030 (ABRAPA, 2022).

A maior parte do algodão produzido e comercializado no mundo trata-se do algodão upland ou herbáceo (*Gossypium hirsutum*), que se caracteriza por ter um cultivo anual e fibras com comprimento de até 32 mm (Fryxell, 1971).

No Brasil, a produção brasileira chegou ao recorde de 3 milhões de toneladas de pluma na safra de 2019/2020. Devido à pandemia e desafios climáticos na safra de 2020/2021, a produção foi reduzida em 22,5% totalizando 2,33 milhões de toneladas (Abrapa, 2022). Na safra atual (2021/2022) houve um incremento na produção para 2,6 milhões de toneladas de pluma em relação à safra de 2020/2021.

Esse crescimento é o resultado do esforço dos produtores, técnicos, pesquisadores e governos por meio de associações de produtores, das instituições públicas e empresas privadas na geração e transferência de novas tecnologias visando aperfeiçoar o sistema produtivo. Este aspecto tem feito com que o Cerrado brasileiro detenha as mais altas produtividades na cultura do algodoeiro no Brasil e no mundo, em áreas não irrigadas (Araujo, 2017).

No Brasil, os Estados de Mato Grosso e Bahia se destacam no cenário nacional e mundial, pois juntos, correspondem a mais de 88% da produção (CONAB, 2021).

Enquanto no Brasil fala-se em recordes e incrementos, nos EUA, o ano de 2022 não foi favorável, pois devido ao estresse hídrico que assola o país, houve abandono em cerca de 43% das áreas de algodão, sendo o mais alto da série histórica registrada pelo Departamento de Agricultura dos Estados Unidos (USDA). A produção caiu de

3,37 milhões de toneladas de pluma estimados em julho/2022 para 2,74 milhões de toneladas. O novo volume representou um recuo anual de 28% e é a menor produção de algodão dos EUA desde 2009/10 (Siqueira, 2022).

Apesar dos EUA serem o segundo maior produtor e o maior exportador com relevante diferença do Brasil, no quesito produtividade o Brasil fica bem à frente dos EUA (Figura 3). Na safra 1990/91, a produtividade média do Brasil era de 363 kg de pluma por hectare, contra 711 kg nos EUA, ou seja, naquela época no Brasil a produtividade era bem menor. Dez anos depois (em 2000/01), o rendimento do Brasil já estava em 1,101 kg por hectare, enquanto o americano estacionou em 709 kg. E 20 anos depois, estamos no ciclo 2021/22, e a safra brasileira rendeu média de 1,633 mil kg por hectare, bem à frente dos 918 kg dos EUA (Siqueira, 2022).

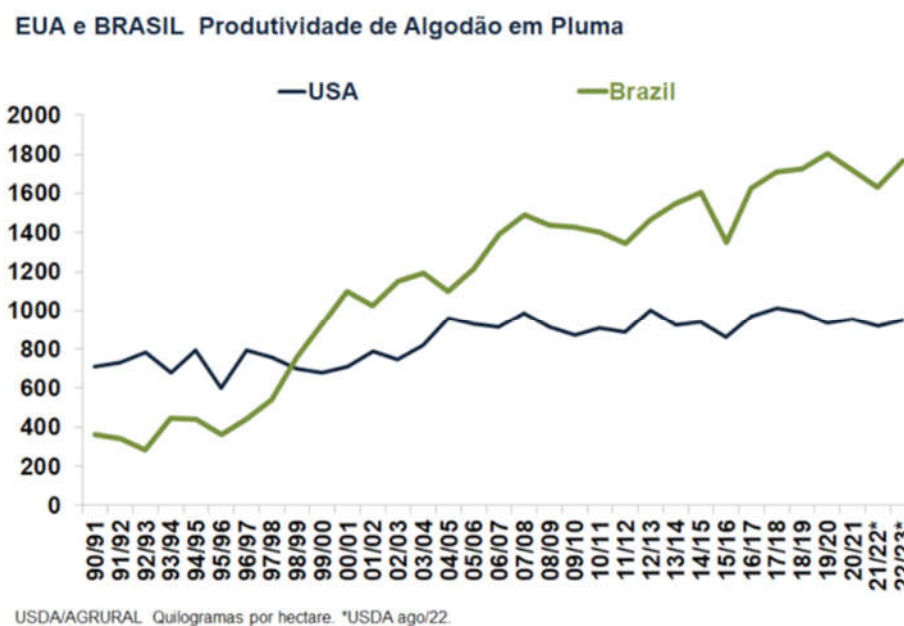


Figura 3. Gráfico comparativo de produtividade média de algodão em pluma entre Brasil e EUA desde a safra 1990/91 a 2022/23 em quilogramas por hectare. Fonte: AgRural (2022).

Em relação ao ciclo do algodoeiro, ele é dividido em quatro fases: a vegetativa (V), formação de botões florais (B), abertura de flores (F) e abertura de capulhos (C) (Rosolem, 2020).

Dentre os fatores ambientais, a temperatura é um dos parâmetros que mais influenciam o desenvolvimento do algodoeiro. Exemplo claro disso é a utilização do

monitoramento da fenologia da cultura por meio da somatória dos graus-dia ou unidades de calor, pois temperaturas médias mais elevadas resultam no adiantamento do ciclo da cultura (Reddy et al., 1996). Esta determinação, referindo-se a unidades de calor (UC) é obtida pelos valores máximos e mínimas diárias de temperatura e subtraída da temperatura base (Oosterhuis, 1999)

A temperatura ótima para o desenvolvimento do algodão está na faixa entre 20 °C e 30 °C, no entanto, é comum o algodoeiro ser cultivado em regiões com temperaturas mínimas abaixo de 15 °C (ex. EUA - período de emergência) e máximas acima de 40 °C (ex. Índia) (Snider & Kawakami, 2014).

O USDA estima a produção mundial da safra 2021/22 em 26,34 milhões de toneladas (+8,3% frente à temporada anterior), impulsionada pelos Estados Unidos (+20,6 %), Brasil (+22 %) e Paquistão (+28,9 %). Mesmo com a redução de 8,5 % na produção chinesa (em relação a 2021), o país asiático ainda deve continuar como segundo maior produtor mundial de algodão. A liderança continua sendo da Índia, que pode ter ligeira queda de 0,4% quando comparado ao da safra passada. Já o consumo global pode aumentar 2,8 %, para 27,05 milhões de toneladas, ficando 2,7 % acima da oferta mundial. A comercialização deve totalizar 10,14 milhões de toneladas, com recuos de 5% nas importações e de 4,4% nos embarques frente aos da safra anterior (Agroinsight, 2022).

No Brasil, a semeadura do algodão ocorre de outubro a fevereiro, a depender da região e a colheita passa a ocorrer entre os meses de março e setembro de cada ano. Em geral, há concentração da colheita de junho a julho, período de pico de safra em Mato Grosso, Goiás e Bahia, principais produtores (Ferreira Filho et al., 2009). Nos EUA a semeadura ocorre de abril a junho e a colheita de setembro a novembro (IMAmt, 2010)

Colheita do algodão

A colheita é uma etapa de alto custo e de extrema importância no processo produtivo do algodão, uma vez que, sendo mal conduzida, poderá acarretar prejuízos qualiquantitativos no produto final (Morelli-Ferreira et al., 2015).

A colheita pode ser realizada manualmente ou mecanicamente. Qualquer que seja o processo de colheita, é indispensável realizá-las em boas condições, a fim de apresentar uma matéria prima – o algodão em caroço – que permita produção durante o beneficiamento de uma fibra de boa qualidade – tipo e qualidade intrínseca – ao menor custo possível (Belot e Vilela, 2006)

Para a colheita do algodão, existem dois tipos de máquinas: a colhedora de fusos (*picker*), é bastante seletiva, os fusos puxam a fibra do capulho aberto e quando a colhedora está bem ajustada, a quantidade de matéria estranha levada para o cesto (sujeira grossa) é razoável, a contaminação vegetal é média, com casquilhas, caules e folha. Já a colhedora *stripper*, dotada de sistema de roldanas, retira os capulhos inteiros, as casquilhas e os invólucros, originando fibras com mais impurezas (Embrapa, 2003).

Mais de 90% do parque de colhedoras do tipo “picker” são da marca John Deere; o restante é da marca Case IH (Mion e Belot, 2018). A colhedora com sistema de fusos (*picker*) é a mais utilizada na cotonicultura (Figura 4) por colher algodão em caroço de boa qualidade, dependendo da variedade, da distribuição espacial de plantas, do uso adequado de desfolhantes, das regulagens das máquinas e velocidade de colheita (Buainain e Batalha, 2007).

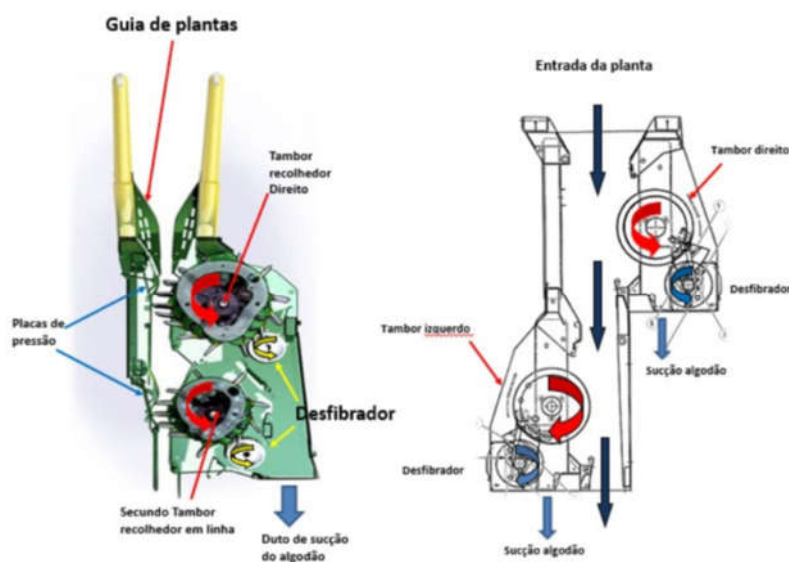


Figura 4. Unidades de colheita com fusos de uma colhedora de algodão John Deere (à esquerda) e Case IH (à direita). Fonte: John Deere Case International citado por Mion e Belot (2019).

As colhedoras do tipo picker tem como principal elemento os fusos em rotação, que extraem de forma seletiva o algodão em caroço dos capulhos abertos da planta do algodão, sem colher as casquilhas, em seguida ele é desprendido dos fusos com desfibradores de borracha (doffer) e levado para o cesto armazenador da máquina por correntes de ar (Belot et al. 2010).

As últimas colhedoras que foram comercializadas para a cotonicultura são equipadas com módulos enfardadores junto às colhedoras, tornando o processo mais eficiente, uma vez que elimina a operação de prensa e transporte feita antigamente pelo bass boy (Kazama, 2015). Para as colhedoras da marca John Deere, os fardos são cilíndricos e possuem em média 2,5 toneladas, sendo embaladas em um filme de polietileno (Figura 5). Este processo permite que a máquina desempenhe uma colheita contínua, sem paradas para descarregamento do produto colhido (Silva, 2015).



Figura 5. Colhedora JD 7760, que possui câmara de enfardamento de módulos interno possibilitando-a colher, prensar, enfardar e depositar o fardo no campo sem parar a máquina. Foto: Francielle Morelli Ferreira.

A formação dos fardos cilíndricos de colheita é realizada pelo módulo acumulador, um sistema que é equipado com sensores de nível de algodão, que projetam um feixe infravermelho, neste sistema, quando o acumulador atinge o topo, o feixe é interrompido, e inicia o processo de formação do módulo. O algodão é então

direcionado para o construtor do módulo, que gira a uma taxa consistente e formam um módulo de algodão que possui tamanho pré-definido e por fim, o processo de enfardar no filme de polietileno é iniciado (John Deere, 2015 citado por Kazama, 2015).

A qualidade da fibra do algodão

O principal produto do algodoeiro é a fibra, que, com 94% de celulose em sua composição, representa 35% a 45% da produção total e corresponde à fibra natural mais consumida no mundo, abastecendo 50% do mercado mundial de fibras têxteis (Santos et al., 2008).

A fibra de algodão é classificada com base em suas características intrínsecas e extrínsecas. Entre as intrínsecas pode-se destacar: comprimento, finura ou diâmetro, espessura da parede secundária ou maturidade, micronaire, resistência e alongamento. As extrínsecas também são muito importantes, pois refletem indiretamente a qualidade da fibra. São elas: cor, aspecto visual, presença e teor de neps e contaminações com matérias estranhas (Chanselme e Marquié, 2006)

Segundo Santana et al. (2008), o termo “classificação do algodão” refere-se à aplicação de procedimentos padronizados e desenvolvidos pelo Departamento de Agricultura dos Estados Unidos da América do Norte – USDA, para avaliar as qualidades físicas do algodão.

O sistema HVI (High Volume Instrument) é a combinação de aparelhos de medição usados para acessar quantitativamente as características físicas das fibras do algodão, que podem fornecer informações mais rapidamente e em maior quantidade, com tanta precisão quanto às classificações e ensaios laboratoriais (Sestren e Lima, 2011).

Numerosas características estão associadas à qualidade das fibras de algodão; as principais são: comprimento, uniformidade, índice micronaire, maturidade, finura, resistência à ruptura, refletância, grau de amarelo, impurezas e contaminantes das fibras do algodão (Bachelier e Gourlot, 2018).

De acordo com a Instrução Normativa n° 24, de 14 de julho dezembro de 2016, do Ministério da Agricultura Pecuária e Abastecimento (MAPA), definem-se os termos

empregados para a fibra de algodão em pluma e para os subprodutos da sua industrialização. O Art. 1º estabelece o Regulamento Técnico do Algodão em Pluma, definindo o seu padrão oficial de classificação e o Art. 2º define os seguintes conceitos (BRASIL, 2016):

- Comprimento de fibra (UHML): o comprimento médio da metade superior (Upper Half Mean Length - UHML) ou comprimento médio dos 50% das fibras mais longas, expresso em polegadas ou em milímetros;
- Grau de amarelecimento (+b): é o valor correspondente ao amarelecimento das fibras com a ajuda de um filtro amarelo e que corresponde à escala do eixo das abscissas do diagrama de cores do Colorímetro de Nickerson/Hunter.
- Grau de cor (C.G.): o resultado do cruzamento dos valores do grau de refletância (%Rd) e do grau de Amarelamento (+b) determinado no diagrama de cores de Nickerson/Hunter.
- Grau de folha ou grau de lixo: o código (L.G.) obtido pela comparação dos corpos de prova com os Padrões Físicos Universais;
- Grau de refletância (%Rd): o valor da luminosidade e da cor branca refletida pelas fibras de algodão, expresso em percentual, que corresponde à escala do eixo das ordenadas do diagrama de cores do Colorímetro (Colorimeter) de Nickerson/Hunter, expresso em porcentagem;
- Índice de maturidade (Mat): o grau de espessura das camadas de celulose que constituem a parede secundária das fibras que formam os corpos de prova
Índice de uniformidade do comprimento da fibra (%UI): a relação entre o comprimento médio dos 100% das fibras (Mean length -ML) e o comprimento médio dos 50% das fibras mais longas (Upper Half Mean Length - UHML), expresso em porcentagem;
- Micronaire da fibra (Mic): o índice determinado pelo complexo finura/maturidade da fibra;
- Índice de fibras curtas (%SFI): o percentual de fibras menores que 0,50 polegadas ou 12,7 milímetros presentes nos corpos de prova;
- Resistência específica ou tenacidade à rotura da fibra (Str - gf/tex): a força, em gramas, requerida para romper um feixe de fibras (barbas de fibras) de um tex.

O grau da cor para os algodões equivalentes ao algodão americano Upland é representado por um código numérico com três dígitos, sendo os dois primeiros relativos a um tipo de cor dos seus Padrões Físicos Universais e o terceiro dígito

representando um dos quadrantes localizados na área do diagrama de cores (Brasil, 2016). Por exemplo, um algodão 31-4 é um algodão de cor 31, e de grau 4 em folha (Morais et al., 2021).

A busca em geral da indústria têxtil é por plumas maduras com maiores percentuais de refletância, menores índices de amarelecimento e impurezas (Santana et al., 2008).

Tendo em vista que o aumento da quantidade produzida de algodão deve ser acompanhado de um aumento da qualidade reportada, uma vez que há penalidades para os comerciantes que negociem fardos de algodão de baixa qualidade ou não classificados, a Associação Brasileira dos Produtores de Algodão (ABRAPA) criou um programa de controle de qualidade dos dados reportados pelos instrumentos em laboratórios parceiros. Um total de 11 laboratórios e 71 instrumentos fazem parte do programa da ABRAPA, para validar os resultados obtidos por todos os laboratórios de classificação instrumental (Abrapa, 2020b).

A variabilidade nas propriedades da fibra pode reduzir a eficiência da fiação e a qualidade do fio. Reduzir a variabilidade do algodão de fardo a fardo permitirá que o algodão compita com fibras sintéticas mais uniformes e previsíveis (Tesema e Drieling, 2019).

A produção de algodão com propriedades mais uniformes (Tabela 1) permitirá que o algodão seja processado com sistemas de alta velocidade, como a fiação a jato de ar, que é duas vezes mais rápida que os atuais sistemas de fiação de algodão (Oxenham, 1992).

As fiações tentam comprar fardos de algodão que sejam adequados para gerar fios que atendam aos contratos acordados. Infelizmente, nem sempre é possível conseguir todos os fardos com a qualidade desejada. Nesse caso, as fiações vão precisar comprar fardos com qualidade inferior ou superior ao necessário, e misturar os fardos conforme a necessidade (Morais et al., 2021).

Tabela 1. Classificação dos principais parâmetros de qualidade da fibra avaliados para comercialização em fibras produzidas no Brasil e nos EUA. Fonte: Brasil (2016) e USDA (2018).

Parâmetro	Classificação	Brasil (Brasil, 2016)	EUA (USDA, 2018)
Micronaire	Muito fina	≤ 3,0	≤ 3,0
	Fina	3,1 - 3,6	3,1 - 3,6
	Média	3,7 - 4,7	3,7 - 4,7
	Grossa	4,8 - 5,4	4,8 - 5,4
	Muito Grossa	≥ 5,5	≥ 5,5
Resistência (g/tex)	Fraca	≤ 24,9	≤ 23,4
	Intermediária	25,0 - 26,9	23,5 - 25,4
	Média	27,0 - 28,9	25,5 - 28,4
	Resistente	29,0 - 31,9	28,5 - 30,4
	Muito Resistente	≥ 32,0	≥ 30,5
Uniformidade de comprimento (%)	Muito baixa	≤ 77,0	≤ 76,5
	Baixa	77,0 - 79,0	76,5 - 79,4
	Regular	80,0 - 82,4	79,5 - 82,4
	Alta	82,5 - 85,5	82,5 - 85,4
	Muito alta	≥ 85,5	≥ 85,4
Índice de fibras curtas (%)	Muito baixa	≤ 6,0	-
	Baixa	6,0 a 7,9	-
	Regular	8,0 a 10,0	-
	Alta	10,1 a 13,0	-
	Muito Alta	≥ 13,0	-

Assim as fiações podem aplicar um desconto quando comprarem fardos com a qualidade inferior às negociadas e poderão aplicar prêmio para fardos com a qualidade superior (Morais et al., 2021) (Tabela 2).

Segundo as Normas Específicas de Algodão (Brasil, 2022) e The Classification of Cotton (Cotton Incorporated, 2018), um valor de micronaire entre 3,5 e 4,9 é considerado um valor adequado, a resistência é adequada para fiação se estiver entre 26,00 a 28,90, o comprimento tem que ser 26,6 mm ou maior, e a uniformidade de comprimento atender um mínimo de 80%, e possuir um máximo de 10% de índices de fibras curtas. Para o mercado interno, seja no Brasil ou EUA, tem que atender a classificação final de CG de 41-4 e para exportação e atender o mercado internacional deve apresentar classificação mínima de 31-4. Todas essas classificações e parâmetros serão evidenciados e discutidos no terceiro capítulo desta tese.

Tabela 2. Tabela com os intervalos padrões para comercialização das fibras de agricultores brasileiros (Brasil, 2022) e americanos (USDA, 2021b). Classificação fora desse intervalo sofrem deságio ou ágio (Brasil)/ descontos ou prêmios (EUA).

Parâmetros	Brasil (Brasil, 2022)	EUA (USDA, 2021b)
Micronaire	3,5 – 4,9	3,5 – 4,9
Resistência (g/tex)	27,0 – 29,0	26,0 – 28,9
Comprimento (mm)	27,4 – 28,0	26,16 ou maior
Uniformidade de comprimento (%)	-	80,0 – 81,9
Índice de fibras curtas (%)	≤ 10,0	-
Gru de cor e folhas (CG)	31-4 (mercado externo) 41-4 (mercado interno)	31-4 (mercado externo) 41-4 (mercado interno)

Rastreabilidade da qualidade da fibra

O termo internet das coisas (internet of things, IoT) foi primeiramente utilizado por Kevin Ashton em 1999 e originalmente se referia à leitura de dispositivos RFID (radio-frequency identification) para o rastreamento de produtos. Em 2015, o grupo de trabalho de IoT do Institute of Electrical and Electronics Engineers (IEEE) produziu uma definição do termo IoT como sendo “um domínio de aplicação que integra diferentes campos tecnológicos e sociais” (Minerva et al., 2015).

Ainda no sistema de colheita convencional (colhedoras de cesto, uso de prensa e bass boy) pesquisadores em propriedade agrícolas no Kansas, California e Texas, USA, criaram um protótipo para obter informações georreferenciadas por meio de sistema Wireless e GPS, obtendo as coordenadas da área colhida de cada fardão, classificando a qualidade da fibra na forma de mapa em nível de módulo de algodão colhido (~10 ton) e sua localização no talhão, e observaram diferenças significativas na maioria das propriedades de fibra de algodão entre os módulos mapeados e tendências foram identificadas (Ge et al., 2008 e Ge et al., 2012); entretanto citam que novos estudos devem ser realizados com aplicações de mapas de qualidade de fibra em conjunto com mapas de rendimento dentre outros dados, podendo ainda produzir mapas de lucro (previstos ou reais) de um campo.

Em 2011, a John Deere introduziu no mercado a colhedora JD 7760 que possui câmara de enfardamento de módulos interno possibilitando-a colher, prensar, enfardar

e depositar o fardo no campo sem parar a máquina, o que revolucionou o armazenamento, o manuseio do algodão após a colheita e a capacidade operacional (Silva e Sofiatti, 2017). Ela também possui sistema RFID (Radio Frequency Identification), que por meio de etiquetas presentes no envoltório plástico (TamaWrap™), permitem a identificação no campo dos módulos redondos de algodão através de um sistema único de etiquetas (Franco, 2020).

Cada módulo criado pela colhedora contém quatro etiquetas RFID e uma etiqueta externa adicionada que identifica exclusivamente o módulo com dados como: número de série do módulo, ID do módulo, variedade, coordenadas geográficas de amarração e descida dos módulos, diâmetro, umidade, área colhida, peso do módulo, e produtividade, por meio do Rastreador de Colheita ou Harvest Identification – HID, Cotton Pro (John Deere, 2022).



Figura 6. Colhedora CP690 representando o equipamento associado ao sistema de rastreador de colheita (HID cotton system). 1) Leitor RFID presente no acumulador do coletor. 2) Medidor de umidade embutido na lateral do acumulador. 3) Sistema de pesagem hidráulico integrado à plataforma traseira. Fonte: Fuhrer (2022).

Os dados HID são gerados a partir da colhedora de algodão com os sensores apropriados e assinatura de licença do produtor. O sistema HID pode ser equipado com os modelos CP e CS 7760, e nos modelos CP e CS 690/770. Esse sistema

completo inclui sistema de pesagem de módulo, leitor RFID dentro do acumulador e um medidor de umidade (Fuhrer, 2022) (Figura 6).

A cadeia de fornecimento do algodão atualmente tem algum nível de integração, pois os dados sobre as medidas de qualidade da fibra de um fardo são repassados ao descaroador e produtor, bem como disponibilizados ao cliente têxtil. No entanto, ainda falta uma integração completa dos dados coletados na fazenda com os dados coletados no gin. Consequentemente, apesar de mais de um século de pesquisa sobre processamento de algodão e qualidade da fibra, poucos estudos examinaram as interações sobre os diferentes componentes da cadeia de suprimentos (Hardin et al., 2022).

Nos EUA, nos últimos cinco anos, a Cotton Incorporated financiou uma demonstração com o laboratório de descaroador do USDA-ARS em Lubbock para permitir a opção de usar apenas a etiqueta RFID para rastrear o algodão do campo durante todo o processo de descaroador sem adicionar outras etiquetas ou tinta spray. Além disso, as informações de propriedade são inseridas apenas uma vez. “As etiquetas RFID e a automação de dados estão começando a permitir que os produtores deixem de depender de bilhetes de papel para rastrear o algodão do campo até o descaroador” (Barnes, 2020).

No Brasil, a ID-COTTON surgiu como solução completa para rastreabilidade de algodão, da colheita até o beneficiamento com o uso da tecnologia RFID. As soluções da ID-Cotton para rastreabilidade do algodão visam combater falhas nos processos de marcação e identificação dos fardos colhidos. Além disso, eliminam rótulos manuais, planilhas manuscritas, “pintura em spray sobre tela e outros métodos antigos de identificação de produtor, fazenda, variedade, campo, borda e contaminantes (IoP Journal, 2020).

Na busca por artigos e projetos abordando a rastreabilidade, mapeamento e a variabilidade da qualidade da fibra do algodão, apenas pesquisas iniciais foram encontradas na USDA (Texas), na North Carolina State University e na University of Georgia.

Em relação ao Brasil, não se encontrou pesquisa científica em desenvolvimento com esse tema e abordagem, fazendo-se importante o presente estudo para explorar mais da tecnologia disponível nas colhedoras de algodão e possibilitar a

rastreabilidade da qualidade da fibra do algodão no campo por meio de mapas da variabilidade espacial dos parâmetros poderão ser gerados.

Com o avanço das tecnologias embarcadas nas colhedoras de algodão (HID Cotton Pro) e pouco exploradas, os produtores podem usar os dados gerados para fornecer aos clientes as informações que eles esperam sobre o algodão que usam, como qualidade da fibra, valor do fardo e outras métricas de sustentabilidade. Esta tecnologia simplifica a transferência de informações do módulo do produtor para a algodoeira, entretanto, no Brasil, o caminho inverso é descontinuado e pouco explorado, necessitando de pesquisa neste tema.

Agricultura de Precisão e Agricultura Digital

A Agricultura de Precisão (AP) se caracteriza por fazer uso intensivo de tecnologias, incluindo geotecnologias como o Sensoriamento Remoto (SR), que se resume em adquirir informação de um objeto sem que se tenha contato físico com o alvo. Em dias atuais, as imagens dos sensores remotos são obtidas em vários tipos de plataforma, tais como, aeronaves, máquinas agrícolas autopropelidas, satélites, veículos aéreos não tripulados (VANTs), dentre outros (Shiratsuchi et al., 2014).

A origem do termo “agricultura de precisão” está fundamentada no fato de que as lavouras não são uniformes no espaço e nem no tempo. Assim, foi necessário o desenvolvimento de estratégias para gerenciar os problemas advindos da desuniformidade das lavouras com vários níveis de complexidade (Molin et al., 2015).

Baseado nisso é que o conceito de AP se adequa, pois considerando a variabilidade espacial do meio agrícola, a AP pode reduzir os custos de produção e aumentar a produtividade, fazendo-se o uso, muitas vezes, de automação e alta tecnologia voltadas à produção agrícola (McKinion et al., 2001; Brandão et al., 2008)

Como já apresentado, a AP é uma forma de gestão de todo o processo agrícola e leva em consideração a variabilidade existente nas lavouras. Logo, é necessário avaliar, quantificar e mapear a variabilidade a fim de geri-la eficientemente. Um ponto de partida para prática de AP é o mapeamento da produtividade da lavoura, caracterizado por apresentar a informação mais completa para se visualizar a variabilidade nos cultivos (Molin et al., 2015).

A agricultura, incluindo a Agricultura de Precisão, está iniciando sua transformação digital. A adoção da automação, de sensores de alta tecnologia, da computação em nuvem, de algoritmos nas tomadas de decisão e da internet das coisas está criando a agricultura digital, na qual dados coletados são usados para aumentar a eficiência dos recursos utilizados pela agricultura. Esses dados, muitas das vezes disponibilizados em tempo real, são utilizados tanto para monitoramento como para a tomada de decisões na agricultura (Borém et al., 2021)

Fazem parte da era digital: inteligência artificial (IA), machine learning, blockchain, big data, computação em nuvem, telemetria, realidade aumentada, deep learning, sensoriamento remoto, mineração de dados, robotização, automação e Sistemas de Informações Geográficas (SIG), entre outras tecnologias (Borém, 2022).

A agricultura digital vem fazendo uso intensivo de tecnologias, e na colheita do algodão, as novas máquinas que fornecem rastreabilidade e documentação da colheita em campo estão em ampla adoção. Segundo Franco (2020) os módulos redondos originados pelas colhedoras JD7760, CP690 e JD770 são a maioria do algodão colhido e processado por descaroçadores nas regiões produtoras de algodão dos EUA e do Brasil, com cerca de 80%.

O avanço desta produção é oriundo de investimentos em tecnologia, destacando a Agricultura de Precisão (AP), em que a ênfase principal está na coleta, gerenciamento e utilização de uma grande quantidade de dados para tomada de decisões (Pham e Stack, 2018).

Uma tecnologia que leva ao aumento da digitalização na agricultura é a Internet das Coisas (IoT), que envolve a adição de capacidade de computação e rede a sensores e máquinas. Esse aumento da digitalização pode melhorar a produtividade, a sustentabilidade e a competitividade com fibras sintéticas (Hardin IV et al., 2022).

A produção e o processamento de algodão podem se beneficiar especificamente das tecnologias IoT. Várias etapas de processamento ocorrem entre a colheita e a produção de produtos têxteis acabados. Muitos parâmetros de qualidade da fibra, como comprimento, finura, resistência e cor da matéria-prima, desempenham um papel na eficiência do processamento e no uso final ideal dessa matéria-prima. Essas medidas de qualidade, assim como a produtividade, são influenciadas pela complexa interação da genética, fatores ambientais e manejo (Hardin IV et al., 2022).

Os dispositivos e as tecnologias de AP evoluem constantemente, bem como a gama de suas aplicações na rotina das fazendas. A AP já é uma realidade no dia a dia dos produtores de algodão, mas o potencial de ampliação de seu uso é enorme, e chegará um momento em que a maioria dos processos de planejamento e decisão será integrada por meio de ferramentas inteligentes, alimentadas por dados especializados de monitoramento dos ambientes de produção e do clima, além de registros sobre o desenvolvimento e o desempenho das lavouras (Brandão et al., 2020).

No geral, diversas técnicas de sensoriamento remoto foram validadas e têm grande aplicabilidade na área agrícola dentro do conceito de AP, porém o custo e a falta de mão de obra qualificada ainda limitam o uso por parte dos usuários (Shiratsuchi et al., 2014).

Aprendizado de máquina na agricultura

O termo machine learning (ML) vem do inglês e pode ser entendido como a capacidade dos computadores de aprenderem com base em dados (Valente et al., 2021). Como resultado, eles são promissores para analisar dados agrícolas em grande volume (big data), agindo de forma mais rápida, eficiente e precisa.

Os algoritmos de ML podem ser divididos em algoritmos de classificação e regressão. A classificação gera resultados categóricos (ou discretos), enquanto a regressão gera previsões de valores contínuos (Valente et al., 2021)

O aprendizado de máquina, que é um ramo da Inteligência Artificial (IA) com foco no aprendizado, é uma abordagem prática que pode fornecer uma melhor previsão de rendimento com base em vários recursos. O ML pode determinar padrões e correlações e descobrir conhecimento de conjuntos de dados. Os modelos precisam ser treinados usando conjuntos de dados, em que os resultados são representados com base na experiência passada (Klompenburg et al., 2020).

O modelo preditivo é construído usando vários recursos e, como tal, os parâmetros dos modelos são determinados usando dados históricos durante a fase de treinamento. Para a fase de teste, parte dos dados históricos que não foram usados

para treinamento são usados para fins de avaliação de desempenho (Klompenburg et al. 2020).

A IA, uma imitação da inteligência humana é utilizada na agricultura para melhorar a qualidade e a precisão no processamento de dados para tomada de decisões, permitindo aos agricultores analisar uma série de variáveis em tempo real, como condições climáticas, temperatura, uso da água ou condições do solo, para dar suporte às suas decisões. Essa tecnologia facilitará praticamente todas as atividades do produtor (Borém et al., 2022).

O aprendizado de máquina é um método que trabalha com análise de dados e busca automatizar a construção de modelos analíticos (Shekoofa et al. 2014; Li et al. 2016). As técnicas de aprendizado de máquina são promissoras para análises de big data mais rápidas, eficientes e precisas (Rehman et al. 2019). Os algoritmos de aprendizado de máquina são o novo paradigma em uso na área agrícola para previsão de produtividade de culturas (Elavarasan et al. 2018).

O uso do ML aliado ao sensoriamento remoto tem sido aplicado na agricultura em diversos temas como na previsão de produção das culturas do milho (Soares et al., 2015; Singh 2008); previsão precisa de produtividade de cana-de-açúcar (Everingham et al. 2016), na previsão de produtividade de soja usando um satélite com a integração de aprendizado de máquina e dados climáticos (Schwalbert et al. 2020), entre outras aplicações.

Estudos avaliando algoritmos com a cultura do algodão ainda são incipientes, mas começam a surgir, com métodos baseados em *deep learning* para identificar características morfológicas em imagens de algodão, como detecção de flores recém-abertas e capulhos para estimar a produtividade (Fue et al., 2018; Xu et al. 2018; Yeom et al., 2018)

No Brasil, um estudo com técnicas da IA aplicadas ao algodão foi realizado por Tedesco-Oliveira et al. (2020), que propuseram abordagem inovadora estimando produtividade do algodão de forma rápida utilizando-se câmera de smartphone na identificação e contagem de capulhos em campos de produção comercial, em que foi possível a contagem de capulhos de algodão usando imagens adquiridas em diferentes horários ao longo do dia, com erros médios de 8,84%.

Em revisão sistemática da literatura (RSL) sobre o uso de aprendizado de máquina na previsão de produtividade de culturas, Klompenburg et al. (2020) afirmam que os algoritmos de ML mais utilizados nas pesquisas até o ano de 2019 foram o Artificial Neural Networks, Linear Regression, Random Forest, Support Vector Machine e Gradient Boosting Tree. Na busca dos algoritmos de deep learning (aprendizado profundo), os três mais utilizados foram a CNN (Convolutional Neural Networks), LSTM (Long-shor Term Memory) e a DNN (Deep Neural Networks).

Sensoriamento remoto na agricultura

Uma das ferramentas da agricultura digital que tem contribuído com grandes avanços para o setor agrícola é o sensoriamento remoto (SR). De maneira geral, a conceituação desse termo indica que o SR é a aquisição de dados sobre um objeto sem tocá-lo diretamente, captando o registro das informações na região do ultravioleta, visível, infravermelho e micro-ondas do espectro-eletromagnético, sem contato, por meio de instrumentos tais como câmeras, escâneres, lasers, dispositivos lineares e/ou matriciais localizados em plataformas como aeronaves, satélites, e a análise da informação adquirida por meio visual ou processamento digital de imagens” (Jensen, 2007).

Em SR, os sensores utilizados podem ser divididos em duas categorias: passivos ou ativos. Sensores passivos registram a energia eletromagnética refletida ou emitida pelo alvo, como a radiação solar refletida ou radiação termal emitida, como exemplo sensores multiespectrais a bordo de diversos satélite (Shiratsuchi et al., 2014), como exemplo o satélite Sentinel utilizado nesta pesquisa. Já os sensores ativos proporcionam fonte própria de energia eletromagnética, como os radares, sonares, sensores ativos de dossel (Crop Circle e Greenseeker) e LiDAR.

Várias aplicações podem ser abordadas através de SR, dentre elas, se destacam a estimativa da biomassa, o monitoramento de estresse hídrico e do vigor nas plantas, a avaliação do estágio fenológico, até estimativa de produtividade da cultura (Brandão, 2009), que é um dos objetivos desta tese.

Todas essas aplicações são possíveis porque alterações nas condições fisiológicas das plantas, impactam diretamente na forma como a radiação

eletromagnética interage com o alvo. Ao incidir sobre o alvo, a radiação pode ser absorvida, transmitida (atravessa) ou refletida, o que define isso são as características do objeto, no caso das plantas, majoritariamente, a estrutura foliar e processos fisiológicos (Almeida, 2022).

Dentro da faixa do espectro-eletromagnético mais trabalhada no SR (Fig. 7), que abrange as regiões do visível (azul (B), verde (G) e vermelho (R)) e infravermelho próximo (NIR), a vegetação, tem menor refletância nos comprimentos de onda do vermelho e azul, devido à alta absorção por parte dos pigmentos fotossintetizantes (Ponzoni et al., 2012).

Alterações na taxa fotossintética, por motivos como senescência, causam o aumento da refletância nestes comprimentos de onda. Já na região do NIR, intimamente relacionada à estrutura celular, a planta em condição sadia tem alta refletância (Molin et al., 2015).

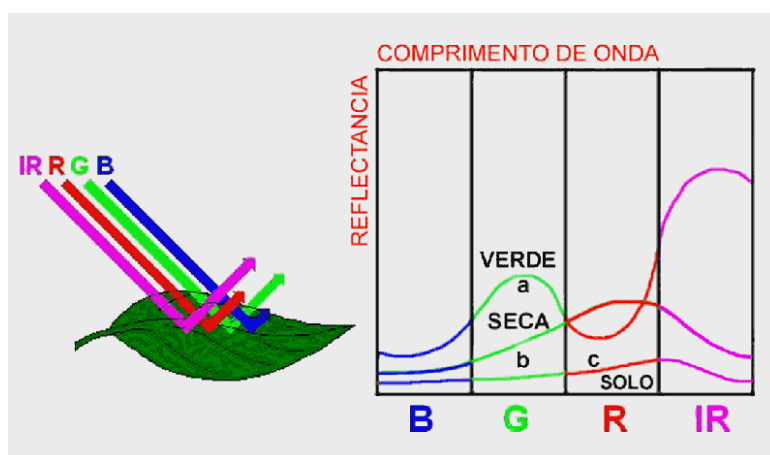


Figura 7. Assinatura espectral da vegetação sadia, seca e do solo. Fonte: Steffen (s/d)

Na agricultura, as propriedades da refletância em cada faixa do espectro eletromagnético podem ser melhor avaliadas por meio de combinações matemáticas (cálculos) de diferentes bandas espectrais (Atzberger, 2013). Essas transformações são medidas da atividade da vegetação utilizam em especial as faixas do visível e do infravermelho próximo e são denominados de índices de vegetação (IVs).

Dentro de SR existem diferentes níveis de coleta de dados, que são eles: orbital, aéreo e proximal (Figura 8). A definição de qual utilizar, deve ser baseada nas

características da área que se deseja monitorar, sendo fundamental considerar também as resoluções fornecidas por cada um dos sensores (espacial, temporal, espectral e radiométrica) (Almeida, 2022).

A resolução espacial diz respeito à dimensão do terreno que é representada pelo menor elemento em uma imagem, o pixel. A resolução espectral diz respeito ao número e ao tamanho dos intervalos de comprimento de onda (bandas) do espectro eletromagnético mensurado. A resolução temporal diz respeito ao intervalo de tempo necessário para a obtenção de imagens de um mesmo local. A resolução radiométrica diz respeito à forma e a precisão de como a radiação eletromagnética é medida pelo sensor remoto (Shiratsuchi et al., 2014).

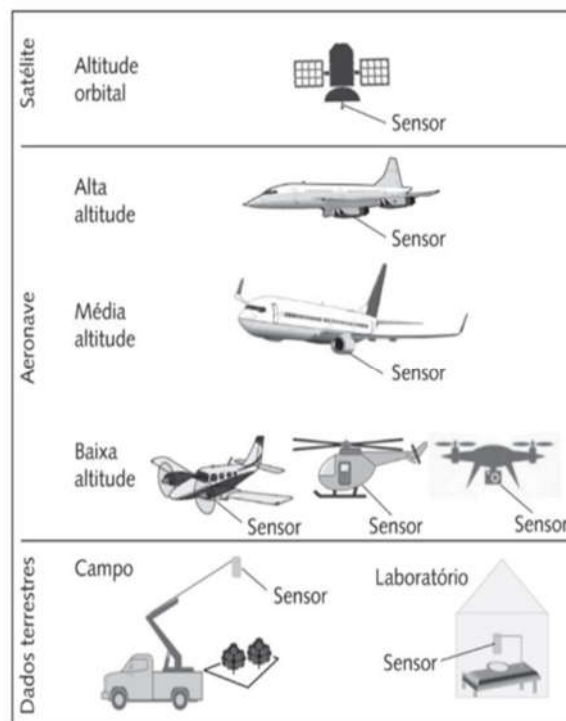


Figura 8. Plataformas de sensoriamento remoto. Fonte: Formaggio e Sanches (2017).

O principal produto do SR utilizado na agricultura atualmente são os índices de vegetação (IV). Para Formaggio e Sanches (2017), os IVs são formulações matemáticas desenvolvidas a partir de dados espectrais obtidos por sensores remotos. Por se tratar da combinação de bandas, os IVs possibilitam minimizar efeitos indesejáveis causados por solo, atmosfera, e variações de geometria Sol-alvo-sensor.

Entre os índices mais conhecidos estão o índice de vegetação da diferença normalizada (NDVI - Rouse et al., 1973) e o Simple Ratio (SR, Jordan, 1969), ambos obtidos a partir das bandas do vermelho e infravermelho próximo. A grande aplicação do NDVI na agricultura se dá pela excelente resposta as variações de vigor da vegetação.

O SAVI (Soil-Adjusted Vegetation Index) (Huete, 1988), foi criado visando amenizar o efeito de solo e se baseia no princípio de que a curva de vegetação tende a aproximar-se da curva de solo para baixas densidades de vegetação, passando por uma mistura de respostas espectrais para densidades médias e quase nenhuma influência do solo para densidades altas de vegetação.

Outros índices bastante utilizados são os que se propõem a reduzir os efeitos atmosféricos com a utilização da faixa do azul do espectro eletromagnético, como os EVI por exemplo, também utilizado nesta pesquisa.

Compreender o crescimento e o desenvolvimento do algodoeiro possibilita a implementação de estratégias de manejo para alto rendimento pois o crescimento dessa espécie é muito dinâmico, variando em função do ambiente e do manejo (Oosterhuis, 1999), e esse monitoramento pode ser realizado por meio dos índices de vegetação.

Uma grande limitação do SR através de sensores passivos orbitais são as nuvens, atmosfera e revisita do imageamento em tempo hábil para a tomada de decisão no manejo agrícola. A nebulosidade se torna crítica principalmente para culturas de verão porque o estágio fenológico da cultura que deve ser imageada geralmente coincide com a época de grande pluviosidade e nebulosidade (Shiratsuchi et al., 2014).

Controle estatístico do processo

A avaliação da qualidade de operações agrícolas é possível por meio do uso de controle estatístico de processos (CEP), caracterizadas como ferramentas que busca melhoria contínua da qualidade do processo através da redução da variabilidade. Porém, não há processo ausente de variabilidade, logo a melhor solução é manter o processo dentro dos padrões de estabilidade (Montgomery, 2013).

O uso de CEP permite que sejam identificadas variações na execução de processos possibilitando classificá-las como advindas de causas normais, acaso e/ou inerentes ao processo, ou de causas especiais, externas, sempre tendo como base indicadores pré-selecionados, limites específicos, ou por indicadores de padrões calculados (limites de controle). Busca-se com o uso do CEP evitar a ocorrência de falhas no processo e evitar gastos desnecessários com ações corretivas nos processos (Milan e Fernandes, 2002; Silva et al., 2013).

Associado à melhor qualidade nos resultados de classificação dos parâmetros de qualidade da fibra do algodão, o CEP torna-se fundamental para o monitoramento da qualidade nas etapas desse processo agrícola, em que as cartas de controle expressam os resultados em gráfico de forma sequencial, permitindo verificar o comportamento médio, estabilidade e variabilidade ao longo da operação (Silva et al., 2008).

Os módulos colhidos geram diversas informações do campo, e na algodoeira, esse módulo gera cinco fardos que são classificados quanto aos parâmetros de qualidade da fibra e estas classificações irão servir de base para venda no mercado interno e externo desta fibra. E será que essa classificação sofre variação?

Segundo Voltarelli et al. (2013) e Tavares et al. (2018), o intuito das cartas de controle, de modo geral, é detectar as possíveis variações externas ao processo; forçar o gerenciamento da operação com a criação de um plano de melhorias; inferir na capacidade e nos limites estabelecidos para o processo. O processo instável ou fora dos limites de controle pode ser analisado ou melhorado utilizando as premissas dos fatores 6 M's (máquina, mão-de-obra, medida, método, matéria-prima e meio).

Referências

Abrapa. Associação Brasileira dos Produtores de Algodão (2022) **A conquista da Ásia.** Disponível em: <<https://www.abrapa.com.br/Paginas/Not%C3%ADcias%20Abrapa.aspx?noticia=1171#:~:text=O%20Brasil%2C%20que%20se%20tornou,4%20milh%C3%B5es%20de%20toneladas%20exportadas>>. Acesso em: 01 out. 2022

Agroinsight (2022) Conjuntura atual do mercado de algodão. Disponível em: <<https://agroinsight.com.br/conjuntura-atual-do-mercado-do-algodao/>>. Acesso: 13 out 2022.

AgRural (2022). Seca nos EUA Deve Ampliar Espaço do Brasil no Mercado do Algodão. Disponível em: <<https://www.czapp.com/pt/analyst-insights/seca-nos-eua-deve-ampliar-espaco-do-brasil-no-mercado-do-algodao/>>. Acesso em 01 nov. 2022.

Aparecido LEO, Meneses KC, Souza Rolim, GR, Carvalho MJN, Pereira WBS., Silva PA, Santos TS, Moraes JRSCM (2020) Algorithms for forecasting cotton yield based on climatic parameters in Brazil. **Arch. Agron. Soil Sci.** 984-1001. Doi 10.1080/03650340.2020.1864821.

Atzberger C (2013) Advances in remote sensing of agriculture: context description, existing operational monitoring systems, and primary information needs. **Remote Sensing**, v. 5, n. 2, p. 949-981 DOI: <http://dx.doi.org/10.3390/rs5020949>

Araujo A (2017). Importância econômica do algodão. Sistemas de Produção – Embrapa Algodão. 2 ed. Disponível em: <https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistema_de_producao_1ga1ceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1&p_r_p_-76293187_sistemaProducaoId=7718&p_r_p_-996514994_topicId=7986>. Acesso em 05 ago. 2022.

Almeida SLH (2022) **Sensoriamento remoto aplicado ao modelo SAFER na estimativa de parâmetros biofísicos de cultivos**. 104f. Tese (Doutorado em Produção Vegetal) – Universidade Estadual Paulista, Jaboticabal.

Bélot JL, Vilela PMCA (2006) Colheita de Algodão. In: Fundo de apoio à cultura do algodão (Ed.). Algodão: pesquisas e resultados para o campo, Cuiabá: FACUAL p.304-325.

Belot JL, Farias FJC, Vilela PMCA (2010) Cultivares de algodoeiro herbáceo para sistema de cultivo adensado. O Sistema de Cultivo do Algodoeiro Adensado em Mato Grosso. Cuiabá. Editora Defanti. 390 p.

Bertolini M, Mezzogori D, Neroni M, Zammori F (2021) Machine Learning for industrial applications: A comprehensive literature review. **Expert Syst. Appl.** 175, 1-29. Doi: 10.1016/j.eswa.2021.114820

Borém A, Queiroz DM, Valente DSM, Pinto FAC (2020). Agricultura Digital. 2 ed. São Paulo: Oficina de textos. Disponível em: <<https://ebooks.ofitexto.com.br/epubreader/agricultura-digital-2-ed>>.

Brandão ZN (2009) **Estimativa da produtividade e estado nutricional da cultura do algodão irrigado via técnicas de sensoriamento remoto**. 152 f. Tese (Doutorado em Recursos Naturais) - Universidade Federal de Campina Grande, Campina Grande.

Brandão ZN, Bezerra MVC, Freire EC, Silva BB (2008) Agricultura de precisão para gerenciamento do algodão. In: Azevêdo DMP, Beltrão, NEM. **O agronegócio do algodão no Brasil**. Brasília: Embrapa Informação Tecnológica, 2008. 1309 p. cap. 20. v. 2.

Brandão ZN, Resende AV, Grego CR, Speranza EA (2020) Agricultura de precisão: tecnologias para o algodoeiro. In: AMPA & IMAMT. Manual de boas práticas de manejo do algodoeiro em Mato Grosso. Cuiabá. 2019/20. 4 ed. p.78-97.

Brasil. Ministério da Agricultura, Pecuária e Abastecimento (2016). Título 41 - Normas Específicas de Algodão – Safras 2021/2022 e 2022. Comunicado CONAB/MOC N.º 008 de 14/04/2022. Disponível em: <https://www.conab.gov.br/images/arquivos/moc/41_NORMAS_ESPECIFICAS_DE_ALGODAO_SAFRA_2021_2022_E_2022.pdf>.

Brasil. Ministério da Agricultura, Pecuária e Abastecimento (2016). Instrução Normativa nº 24, de 14 de julho de 2016. Regulamento Técnico do algodão em pluma. Disponível em: <http://www.inmetro.gov.br/barreirastecnicas/pontofocal/textos/regulamentos/BRA_683.PDF>.

Buainain AM, Batalha MO (2007) **Cadeia Produtiva do Algodão**. Ministério da Agricultura, Pecuária e Abastecimento, Secretaria de Política Agrícola, Instituto Interamericano de Cooperação para a Agricultura. Série Agronegócios, v. 4, 108 p. Brasília. Jan. 2007.

Chanselme JL, Marquié C (2006) A qualidade da fibra de algodão em relação a seus mercados. In: Fundo de apoio à cultura do algodão (Ed.). **Algodão: pesquisas e resultados para o campo**, Cuiabá: FACUAL p.327-358.

Elavarasan D, Vincent DR, Sharma V, Zomaya AY, Srinivasan K. 2018. Forecasting yield by integrating agrarian factors and machine learning models: A survey. **Comput Electron Agric** 155:257–282. Doi: 10.1016/j.compag.2018.10.024.

Embrapa Empresa Brasileira De Pesquisa Agropecuária (2003) Cultura do algodão no cerrado: Colheita.

Everingham Y, Sexton J, Skocaj D, Inman-Bamber G. 2016. Accurate prediction of sugarcane yield using a random forest algorithm. **Agronomy for Sustainable Development**. 36(2):27. Doi:10.1007/s13593-016-0364-z

Filho JBSF, Alves LRA, Villar PMD (2009) Estudo da competitividade da produção de algodão entre Brasil e Estados Unidos - safra 2003/04. **Rev Eco Sociol Rural**, 47 (1). DOI: <<https://doi.org/10.1590/S0103-20032009000100003>>.

Formaggio AR, Sanches IDA (2017) Sensoriamento remoto em agricultura. Oficina de Textos.

Franco BC (2020) Tama e o algodão. Disponível em: <<https://soudealgodao.com.br/blog/apoiadores/tama/>>. Acesso em: 20 out. 2022.

Fue KG, Porter WM, Rains GC (2018) Deep Learning based Real-time GPU-accelerated Tracking and Counting of Cotton Bolls under Field Conditions using a Moving Camera, In 2018 Detroit, Michigan, American Society of Agricultural and Biological Engineers. Doi: 10.13031/aim.201800831.

Fuhrer L (2022) **Mapping in-field cotton fiber quality utilizing John Deere's Harvest Identification System (HID)**. 73 f. Thesis (Master of Science). University of Georgia, Tifton.

Hardin IV RG, Barnes EM, Delhom CD, Wanjura, JD, Ward, JK (2022). Internet of things: Cotton harvesting and processing. *Computers and Electronics in Agriculture*, 202, 107294. Doi: <https://doi.org/10.1016/j.compag.2022.107294>

Huete AR (1988) A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, New York, v. 25, n. 3, p. 295-309. Doi: 10.1016/0034-4257(88)90106-X

IMAmt – Instituto Mato-grossense do Algodão (2010) Cenário do Algodão publico. Disponível em: <https://imamt.org.br/wp-content/uploads/2019/03/Cen%C3%A1rio_do_Algod%C3%A3o_publico_Cen%C3%A1rio_do_Algod%C3%A3o_publico.pdf>. Acesso em 01 nov. 2022.

Jensen JR (2007). Remote Sensing of the Environment: An Earth Resource Perspective. 2nd. ed. Upper Saddle River, NJ: Prentice Hall, 592p

John Deere. Colhedora de Algodão CP690. Disponível em: <<https://www.deere.com.br/pt/magazines/publication.html?id=4f26a16d#5>>. Acesso em: 01 nov. 2022.

Jordan CF (1969) Derivation of leaf área index from quality of light on the Forest floor. Ecology, v. 50, p.663-666, 1969

Kazama EH (2015) **Perdas na colheita mecanizada de algodão em variabilidade espacial e em função de velocidades de colheita**. 73f. Dissertação (Mestrado em Agronomia – Ciência do Solo) – Universidade Estadual Paulista, Jaboticabal.

Klompenburg TV, Kassahun A, Catal C (2020) Crop yield prediction using machine learning: A systematic literature review. **Comput. Electron. Agric.** 177, 105709. DOI: <https://doi.org/10.1016/j.compag.2020.105709>

Li L, Baker TE, White SR, Burke K. (2016) Pure density functional for strong correlation and the thermodynamic limit from machine learning. Physical Review B. 94(24):245129. others. Doi:10.1103/PhysRevB.94.245129.

Mckinion JM, Jenkins JN, Akins D, Turner SB, Willers JL, Jallas E, Whisler FD (2001) Analysis of a precision agriculture approach to cotton production. Computers and Electronics in Agriculture, v. 32, n. 3, p. 213- 218. DOI: [http://dx.doi.org/10.1016/S0168-1699\(01\)00166-1](http://dx.doi.org/10.1016/S0168-1699(01)00166-1)

Milan M, Fernandes RAT (2002) Quality of tillage operations by statistical process control. Scientia Agrícola, Piracicaba, v.59, n.2, p.261-266.

Minerva R, Biru A, Rotondi D (2015) Towards a Definition of the Internet of Things (IoT). IEEE Internet Initiative, v. 1, n. 1, p. 1-86.

Mion RL, Belot JL (2018) Colheita, armazenamento, transporte e qualidade da fibra. In: AMPA & IMAMT. Manual de qualidade da fibra da AMPA. Ed. Jean-Louis Belot. Cuiabá. 2018. p.238-268.

Molin JP, Amaral LR, Colaço AF (2015) Agricultura de precisão. Oficina de Textos: São Paulo, p. 119-153.

Montgomery DC (2013) Statistical Quality Control. 7th Edition. New York: John Wiley & Sons, Inc.

Morais JPS, Farias FJC, Belot J, Martins RSA, Mizoguchi ET (2021). Interpretação das características avaliadas no SITC para qualidade de fibra de algodão – uma abordagem prática. Campina Grande: Embrapa Algodão. Disponível em: <<https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1134241/1/PUBLICA-R-AINFO-DOC285.pdf>>. Acesso em 05 ago. 2022.

Morelli-Ferreira F, Fiorese DA, Silva, ARB (2013). Sistemas de colheita picker e stripper: características e influências da colheita mecanizada de algodão adensado no estado de Mato Grosso. *Enciclopedia Biosfera*, 9(17). Disponível em: <<https://conhecer.org.br/ojs/index.php/biosfera/article/view/3213>>. Acesso em 01 set 2022.

Oosterhuis DM (1999) Growth and development of a cotton plant. In: Cia, E.; Freire, E.C.; Santos, W.J. **Cultura do algodoeiro**. Piracicaba: POTAFOS, 286p.

Oxenham W (1992) Influence of fibre properties in air jet spinning **Ind. J. Fibre Textile Res.**, 17, pp. 194-200.

Pham X, Stack M (2018) How data analytics is transforming agriculture. **Business Horizons** 61(1), 125–133. Doi: 10.1016/j.bushor.2017.09.011

Ponzoni FJ, Shimabukuro YE, Kuplich TM (2012) Sensoriamento Remoto no Estudo da Vegetação. Oficina de textos, São Paulo.

Reddy VR, Hodges HF, Mccarty WH, Mckinnon JM (1996) Weather and cotton growth: Present and Future. Mississippi Agr. & Forestry Exp. Sta., Mississippi State University, Starkville, MS.

Rehman TU, Mahmud MS, Chang YK, Jin J, Shin J (2019) Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. **Computers and Electronics in Agriculture**. 156:585–605. Doi:10.1016/j.compag.2018.12.006.

Rosolem CA (2020) Crescimento do algodoeiro. p. 105. In: BÉLOT, J.; VILELA, P. **Manual de boas práticas de manejo do algodoeiro em Mato Grosso**. Cuiabá, 461 p. 4 ed. Disponível em: <http://www.casadoalgodao.com.br/images/publicacoes/manualdeboaspraticas2020-4ed-vf-web.pdf>.

Rouse JW, Haas RH, Schell JÁ, Deering DW (1973) Monitoring vegetation systems in the great plains with ERTS. In: Earth Resources Technology Satellite-1 Symposium, 3. Washington. **Proceedings...** Washington: NASA, Goddard Space Flight Center, 1, 309-317.

Santana JCF et al. Características da fibra e do fio do algodão: Análise e interpretação dos resultados. In: BELTRÃO, N. E. M. e AZEVEDO, D. M. P (Ed.). **O agronegócio do algodão no Brasil**. 2 ed. Brasília. Embrapa Informação Tecnológica. 2008. p. 1101-1120.

Santos RF, Koury J, Santos JW (2008) O Agronegócio do algodão crise e recuperação no mercado brasileiro da matéria-prima agrícola. In: Beltrão NEM, Azevedo DMP (2008) **O agronegócio do algodão no Brasil**. 2. ed. Brasília, DF: Embrapa Informação Tecnológica, v.1, p.31-60.

Severino LS, Silva Filho J. L. E.; Santos, J. B.; Alencar, A. R. **Plantio de Algodão Adensado no Oeste Baiano: Safra 2002-2003**. Campina Grande: Embrapa Algodão. Comunicado Técnico 209, 2004. 4 p.

Siqueira D (2022) Seca nos EUA Deve Ampliar Espaço do Brasil no Mercado do Algodão. Disponível em: <<https://www.czapp.com/pt/analyst-insights/seca-nos-eua-deve-ampliar-espaco-do-brasil-no-mercado-do-algodao/>>. Acesso em 01 nov. 2022.

Schwalbert RA, Amado T, Corassa G, Pott LP, Prasad PVV, Ciampitti IA. 2020. Satellite-based soybean yield forecast: integrating machine learning and weather data for improving crop yield prediction in southern Brazil. **Agricultural and Forest Meteorology**. 284:107886. Doi: <https://doi.org/10.1016/j.agrformet.2019.107886>

Sestren JA, Lima JJ (2011) Características e classificação da fibra do algodão. In: FREIRE, E. C. **Algodão no cerrado do Brasil**. Associação Brasileira dos Produtores de Algodão – ABRAPA. Aparecida de Goiânia: Mundial Gráfica. 2ªed. 1082 p.

Shekoofa A, Emam Y, Shekoufa N, Ebrahimi M, Ebrahimie E, Anisimova M (2014) Determining the most important physiological and agronomic traits contributing to maize grain yield through machine learning algorithms: a new avenue in intelligent agriculture. **PLoS One**. 9(5): 97288. Doi:10.1371/journal.pone.0097288

Shiratsuchi LS, Brandão ZN, Vicente LE, Victoria DC, Ducati JR, Oliveira RP, Vilela MF (2014) Sensoriamento Remoto: conceitos básicos e aplicações na Agricultura de Precisão. In: Bernardi ACC, Naime JM, Resende AV, Bassoi LH, Inamasu RY (2014) **Agricultura de precisão: resultados de um novo olhar**. Brasília: EMBRAPA, 2014. Cap. 4. p.58-73.

Silva ORRF, Sofiatti V, Mion RL (2015) Colheita mecanizada do algodão. In: Freire EC (ed). **Algodão no Cerrado do Brasil**. Brasília: Positiva. Cap 16, 583-611, 2015.

Silva ORRF, Sofiatti V (2017) Colheita e Beneficiamento In: Sistemas de Produção 2 ed. Embrapa Algodão. ISSN 1678-8710. Disponível em: <https://www.spo.cnptia.embrapa.br/conteudo?p_p_lifecycle=0&p_p_id=conteudoporlet_WAR_sistemasdeproducao16_1ga1ceportlet&p_p_col_count=1&p_p_col_id=column-1&p_p_state=normal&p_r_p_-76293187_sistemaProducaoId=7718&p_r_p_-996514994_topicId=7993&p_p_mode=view>.

Silva RP, Cassia MT, Voltarelli MA, Compagnon AM, Furlani CEA (2013) Qualidade da colheita mecanizada de feijão (*Phaseolus vulgaris*) em dois sistemas de preparo do solo. **Revista Ciência Agronômica**, v.44, p.61-69.

Silva RP, Corrêa CF, Cortez JW, Furlani CEA (2008) Statistical control applied to the mechanized harvesting process of sugarcane. **Engenharia Agrícola**, v. 28, p. 292-304.

Singh RK (2008). Artificial neural network methodology for modelling and forecasting maize crop yield. **Agric Econ Res Rev**. 21(347–2016–16813):5–10

Snider J, Kawakami EM (2014) Efeito da temperatura no desenvolvimento do algodoeiro. In.: Echer. F. R. **O algodoeiro e os estresses abiótico: Temperatura, luz, água e nutrientes**. Cuiabá (MT), 123 p.

Tavares TO, Borba MAP, Oliveira BR, Silva RPS, Voltarelli MA, Ormond ATS (2018) Effect of Soil Management Practices on the Sweeping Operation during Coffee Harvest. **Agronomy Journal**. 110 (5): 1689. Doi: 10.2134 / agronj2017.10.0598

Tedesco-Oliveira D, Silva RP, Maldonado-Jr W, Zerbato C (2020) Convolutional neural networks in predicting cotton yield from images of commercial fields. **Comput. Electron. Agric**. 171, 105307. DOI: 10.1016/j.compag.2020.105307

Tesema GB, Drieling A (2019) Statistical analysis and instrumental characterization of commercial Ethiopian cotton varieties Retrieved from **J Textile Apparel Technol Manage**, 11 (1), pp. 1-16.

USDA - United States Department of Agriculture (2018) Cotton Classification: Understanding the Data. Disponível em: <<https://www.ams.usda.gov/sites/default/files/media/CottonDBUnderstandingtheData.pdf>>. Acesso em: 04 ago. 2022.

USDA - United States Department of Agriculture (2021a) National Agricultural Statistics Service. Crop Production – December. ISSN: 1936-3737. Disponível em: <https://www.nass.usda.gov/Publications/Todays_Reports/reports/crop1221.pdf>. Acesso em: 10 mar 2022.

USDA - United States Department of Agriculture (2021b) Premiums and Discounts for Grade, Staple Length, and Leaf Content of 2021 - Crop American Upland Cotton 1. 5 p. Disponível em: <<https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdafiles/Price-Support/pdf/2021/2021%20Cotton%20%20Premiums%20and%20Discounts.pdf>>.

Valente DSM, Queiroz DM, Pereira GW (2021) Machine Learning In.: Borém A, Queiroz DM, Valente DSM, Pinto FAC (2020). **Agricultura Digital**. 2 ed. São Paulo: Oficina de textos. Cap. 14. Disponível em: <<https://ebooks.ofitexto.com.br/epubreader/agricultura-digital-2-ed>>.

Voltarelli MA, Silva RP, Rosalen DL, Zerbato C, Cassia MT (2013) Quality of performance of the operation of sugarcane mechanized planting in day and night shifts. *Australian Journal of Crop Science*, v.7, n.9, p.1396-1406.

Xu R, Li C, Paterson AH, Jiang Y, Sun S, Robertson JS (2018). Aerial images and convolutional neural network for cotton bloom detection. **Front. Plant Sci.** 8, 1–17. Doi: 10.3389/fpls.2017.02235

Yeom J, Jung J, Chang A, Maeda M, Landivar J (2018) Automated open cotton boll detection for yield estimation using Unmanned Aircraft Vehicle (UAV) Data. *Remote Sens.* 10, 1–20. Doi: 10.3390/rs10121895

CAPÍTULO 2 - Comparação de algoritmos de aprendizado de máquina na predição de produtividade do algodão em campos comerciais utilizando sensoriamento remoto por satélite

RESUMO - É de grande importância estimar a produtividade do algodão com rapidez e precisão antes da operação de colheita. A utilização de técnicas de machine learning para previsão de rendimento com base em sensoriamento remoto está se consolidando nesta tarefa. Pressupondo que índices de vegetação sejam preditores da produtividade de algodão utilizando aprendizagem de máquina supervisionada e visando auxiliar o cotonicultor na antecipação das tomadas de decisões, propusemos pesquisas neste tema, mas com foco em campos comerciais e utilizando imagens de satélites gratuitas (resolução de 10 m). O objetivo deste capítulo foi aplicar ferramentas da agricultura digital e inteligência artificial para identificar o quão cedo, quais índices de vegetação e quais algoritmos de machine learning conseguem prever a produtividade do algodão com boa precisão e acurácia em campos comerciais. Para isso percorremos as seguintes etapas: 1) Observamos a produtividade em 398 ha (três campos) e oito índices de vegetação (IV) foram calculados em cinco datas durante o ciclo do algodão. 2) Cenários foram criados para facilitar as análises e interpretação dos resultados: Cenário 1: All Data (oito índices em 5 datas = 40 inputs) e Cenário 2: variável selecionada por Stepwise (um input). 3) Na busca do melhor algoritmo, ajustes de hiperparâmetros, calibrações e testes usando machine learning foram realizados na predição da produtividade e as performances foram avaliadas. O cenário 1 obteve as melhores métricas em todos os campos de estudo, e os algoritmos Multilayer Perceptron (MLP) e Random Forest (RF) apresentaram os melhores desempenhos com R^2 ajustado de 47% e RMSE de apenas 0.24 t ha^{-1} , entretanto, neste cenário necessita-se de todos os inputs preditores que foram gerados ao longo de todo o ciclo da cultura (aprox. 180 dias), sendo assim, otimizamos a predição e testamos apenas o melhor IV em cada campo, e detectamos que dentre os oito IV, o Simple Ratio (SR), impulsionado pelo algoritmo K-Nearest Neighbor (KNN) prediz com $0,26$ e $0,28 \text{ t ha}^{-1}$ de RMSE e MAPE de 5.20% antecipando com baixo erro em aproximadamente ± 143 dias antes da colheita a produtividade de algodão, e com aspecto importante de exigir menor demanda computacional na geração da predição quando comparado a MLP e RF por exemplo, possibilitando sua utilização como técnica que auxilia produtores e gerenciadores de dados na predição de produtividade, resultando em ganho de tempo no planejamento, seja da comercialização ou na intervenção no manejo e colheita do algodão.

Palavras-chave: agricultura digital, índice de vegetação, KNN, mapa de colheita, random forest, SR

COMPARISON OF MACHINE LEARNING ALGORITHMS IN-FIELD COTTON YIELD PREDICTION USING SATELLITE REMOTE SENSING

ABSTRACT - It is of great importance to estimate cotton yield quickly and accurately before the harvesting operation. The use of machine learning techniques for yield forecasting based on remote sensing is consolidating in this task. Assuming that vegetation indices can be predictors of cotton productivity using supervised machine learning and aiming to help the cotton grower in anticipating decision-making, we proposed research on this topic, focusing on commercial fields and using free medium-resolution satellite images. The objective of this chapter was to apply digital agriculture tools and artificial intelligence to identify how early, which vegetation indices and machine learning algorithms can predict cotton yield with good precision and accuracy in commercial fields. Yield data was collected in 398 ha (three fields), and eight vegetation indices (VI) were calculated on five dates during the growing season. Scenarios were created to facilitate the analysis and results interpretation. Scenario 1: All Data (8 index on 5 dates = 40 inputs) and Scenario 2: best variable selected by Stepwise regression (one input). The algorithms hyperparameter were optimized, calibrations and tests using machine learning were used to predict yield, and performance metrics were evaluated. Scenario 1 achieved the best performance metrics in all fields of study. Multilayer Perceptron (MLP) and Random Forest (RF) algorithms showed the best performances with adjusted R^2 of 47% and RMSE of only 0.24 t ha^{-1} ; however, in this scenario, all inputs that were generated throughout the growing season (approx. 180 days) are needed. Then we optimized the prediction and tested only the best VI in each field, and the Simple Ratio (SR) plus the K-Nearest Neighbor (KNN) algorithm, predicting with 0.26 and 0.28 t ha^{-1} of RMSE and 5.20% MAPE, anticipating the cotton yield with a low error by ± 143 days. An essential aspect of requiring less computational demand in the prediction generation is MLP and RF, for example. This enables its use as a technique that helps farmers predict cotton yield, resulting in time savings for planning, whether in marketing crop management strategies.

Keywords: digital agriculture, vegetation index, yield mapping, KNN, random forest, SR

Introdução

O algodão (*Gossypium hirsutum* L.) é a principal fonte de fibra natural e desempenha papel importante na economia de diversos países. Apesar de nos Estados Unidos da América (EUA) e no Brasil (maiores exportadores mundiais da fibra – USDA, 2020) a colheita do algodão ser totalmente mecanizada com alto nível de tecnologia, ainda se observa utilização de métodos manuais para se estimar a produtividade antes da operação de colheita, principalmente em fazendas que não possuem colhedoras com sensor de rendimento.

Na agricultura convencional a produtividade média é determinada com auxílio de balanças. Apesar de ser um método confiável, caracteriza-se por coletas de amostragem destrutiva no campo e que demanda tempo. O monitoramento não destrutivo e preciso do rendimento é de grande importância para a produção de algodão.

Em função do rápido avanço tecnológico e do complexo cenário big data (grande volume de dados), a exploração da inteligência artificial (IA) na agricultura digital permitem estudos sobre as diversas respostas mensuradas por tecnologias de monitoramento e processamento de dados disponíveis gratuitamente nas plataformas digitais (imagens de satélite, parâmetros climáticos, séries temporais de safras) processadas em softwares livres (QGIS e ambientes de linguagem de programação como Java, R, Python), possibilitando o uso de algoritmos de aprendizado de máquina, de forma gratuita, incentivando a realização de pesquisas e desenvolvimento de softwares comerciais.

As técnicas mais populares utilizadas para analisar imagens incluem índices de vegetação, análise de regressão e correlações, em que, nos últimos anos técnicas de aprendizados de máquina (Singh et al., 2016) estão sendo direcionadas para o meio agrícola devido a intensificação da agricultura digital e geração de grande volume de dados. Porém, para o algodoeiro, que é uma cultura de ciclo longo (até 220 dias) e de hábito de crescimento indeterminado este tipo de abordagem é incipiente, principalmente com experimentação em campos comerciais, tornando este trabalho um diferencial para a literatura.

Algoritmos de aprendizado de máquina (ML) colocam em prática a ideia de que as máquinas podem aprender parcial ou totalmente de forma autônoma diretamente dos dados de entrada (Bertolini et al., 2021). Como resultado, estes são promissores para analisar dados agrícolas em grande volume (big data), agindo de forma mais rápida, eficiente e precisa. Alguns exemplos de algoritmos de ML são K-Nearest Neighbor (KNN), Linear Regression (LR), Artificial Neural Networks – Multilayer Perceptron (MLP) e Random Forest (RF) (Badnakhe et al., 2018).

Estudos avaliando algoritmos com a cultura do algodão começam a surgir. Pesquisadores como Fue et al. (2018); Xu et al. (2018) e Yeom et al. (2018) propuseram métodos baseados em *deep learning* para identificar regiões de interesse na planta em imagens de algodão.

Pesquisadores desenvolveram um sistema que permite rastrear e contar cápsulas de algodão cultivadas em vasos de plantas usando imagens em movimento (Fue et al, 2018). Um método usando uma rede neural convolucional para detectar e contar o número de flores de algodão recém-abertas em imagens aéreas foi desenvolvido (Xu et al., 2018) e criaram um algoritmo para identificar e contar capulhos estimando o rendimento de algodão usando imagens aéreas (Yeom et al., 2018).

A previsão de produtividade foi alcançada em grande e pequena escala usando deep learning para extrair cápsulas de algodão, melhorando a precisão da previsão, e os autores geram mapas de rendimento de alta resolução de acordo com o modelo (Xu et al., 2021).

Li et al. (2022) avaliaram a estimativa de rendimento de campos de algodão de alta densidade usando imagens de UAV de baixa altitude e aprendizado profundo e concluíram que este estudo fornece uma referência metodológica para estimativa de rendimento de algodão na China.

Estudos com aquisição de imagens de algodão por meio de UAV carregando sensor RGB com base em índices de vegetação visível, características de textura e suas combinações, concluíram que os resultados encontrados comprovam certo potencial no monitoramento da produtividade do algodão, o que pode fornecer base teórica e suporte técnico para a avaliação da produção de algodão no campo (Ma et al. 2022).

No Brasil, estudos com técnicas da IA aplicadas ao algodão se iniciaram em 2018 em que Tedesco-Oliveira et al. (2020), que propuseram abordagem inovadora estimando produtividade do algodão de forma rápida utilizando-se câmera de smartphone na identificação e contagem de capulhos em campos de produção comercial, e com os pesquisadores confirmaram a possibilidade de contagem de capsulas usando imagens adquiridas em diferentes horários ao longo do dia, com erros médios de 8,84%.

Baseando-se em parâmetros climáticos no âmbito de regiões do Brasil, Aparecido et al. (2020) estimaram a produtividade de algodão na região centro-oeste do Brasil, utilizando algoritmos de machine learning e enfatizaram que é possível prever com precisão a produtividade do algodão para as principais regiões produtoras do Brasil, concluindo que o melhor algoritmo foi o Extras-tree-regressor (TREE) e o de menor desempenho foi a Multiple Linear Regression. Segundo os autores, o algoritmo TREE conseguiu prever com uma antecipação de aproximadamente 80 dias da colheita.

Portanto, é de grande importância estimar a produtividade do algodão com rapidez e precisão antes da operação de colheita. A utilização de técnicas de machine learning para previsão de rendimento com base em sensoriamento remoto está se consolidando nesta tarefa.

Pressupondo que índices de vegetação possam ser preditores da produtividade de algodão utilizando aprendizagem de máquina supervisionada e visando auxiliar o cotonicultor na antecipação das tomadas de decisões, propusemos pesquisas neste tema, mas com foco em campos comerciais e utilizando imagens de satélites gratuitas de média resolução.

O objetivo deste capítulo foi aplicar ferramentas da agricultura digital e inteligência artificial para identificar o quão cedo, quais índices de vegetação e quais algoritmos de machine learning conseguem prever a produtividade do algodão com boa precisão e acurácia em campos comerciais.

Material e Métodos

Campos de estudo

A pesquisa foi conduzida na safra de 2019 na mesorregião norte do Estado de Mato Grosso, maior produtor de algodão do Brasil (CONAB, 2020). Os locais de estudo (Fig. 1) foram três campos de algodão comercial em fazenda próxima a cidade de Santa Rita do Trivelato ($13^{\circ}59'17''$ S, $55^{\circ}23'15''$ W).

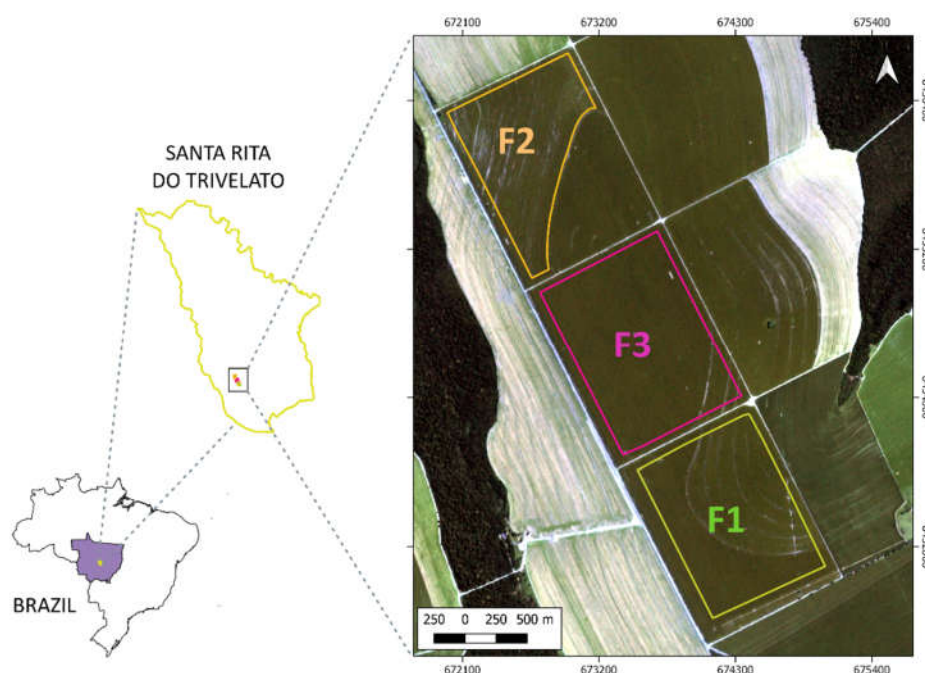


Figura 1. Localização dos campos de estudo na unidade de produção em Santa Rita do Trivelato - MT. Foto: Planet Team (2017).

Os campos foram denominados como: Campo 1 - F1 (137 ha), Campo 2 – F2 (102 ha) e Campo 3 – F3 (159 ha) totalizando 398 ha de estudo. Os campos apresentam topografia predominantemente plana e os solos da região de estudo estão classificados como Latossolo Vermelho-Amarelo Distrófico, seguindo os critérios estabelecidos pelo Sistema Brasileiro de Classificação dos Solos (Santos et al., 2018), apresentando textura média (30 a 35% de argila).

O clima na região é o tropical quente e úmido, com inverno seco (Aw) (Köppen e Geiger, 1928), com temperatura média anual próxima de 25 °C e pluviosidade de aproximadamente 1800 mm ao ano.

A cultura foi monitorada por técnicos da propriedade para controle de pragas e doenças, assim como todas as operações necessárias para o bom estabelecimento da cultura. A cultivar utilizada foi IMA5801B2RF, com espaçamento entre linhas de 0,76 m, e semeadura em janeiro/2019 e a colheita em julho/2019.

Durante o ciclo do algodão avaliado nos campos de estudo, a temperatura média foi de 22.5 °C com precipitação acumulada em 654 mm, estando próximo aos 700 mm exigidos pela cultura. Os dados meteorológicos influenciam diretamente nas condições de desenvolvimento do algodoeiro e são apresentados na Figura 2.

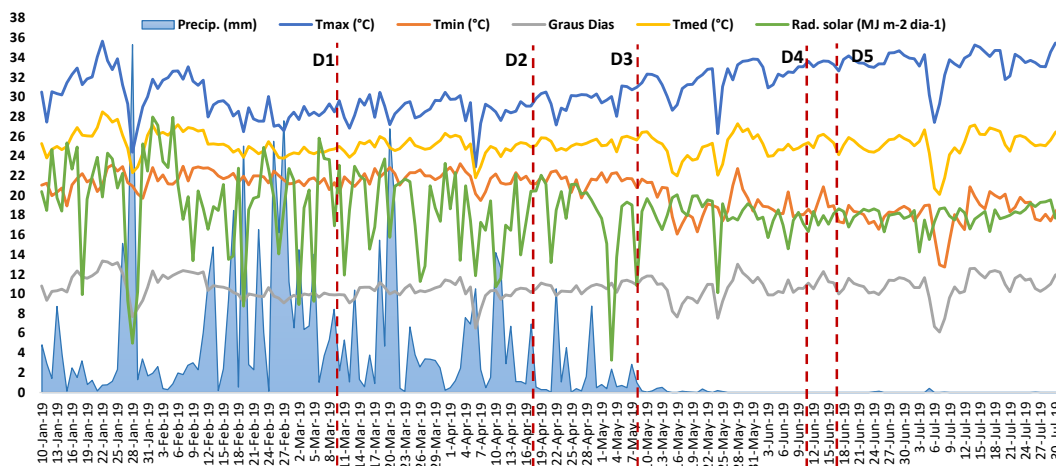


Figura 2. Precipitação (mm), temperatura máxima (°C), mínima (°C), graus dias, temperatura média (°C) e radiação solar (MJ m² dia⁻¹) para os campos de estudo em Santa Rita do Trivelato – MT, Brasil. D1, D2, D3, D4 e D5 equivalem às datas de aquisição das imagens

Os dados climáticos foram obtidos em uma escala diária na National Aeronautics and Space Administration/Prediction of World Ampla plataforma de recursos de energia - NASA/POWER (Stackhouse, 2010) e se fez importante para interpretação dos resultados de índice de vegetação nas datas de estudo.

Aquisição e processamento dos dados das imagens de satélite

No decorrer das safras de 2019 foram adquiridas imagens (Tabela 1) do Sentinel-2A pelo portal Earth Explorer da USGS (United States Geological Survey) ao longo dos estádios fenológicos do algodão para monitoramento e extração dos índices de vegetação.

Os locais do estudo foram subconjuntos da imagem total, normalizados para a refletância, corrigidos para os efeitos atmosféricos e, por meio dos valores espectrais das bandas Green (560 nm), Blue (490 nm), Red (665 nm), RedEdge (705 nm) e NIR (842 nm), os índices de vegetação foram calculados sendo os preditores de entrada para a predição de produtividade do algodão.

Tabela 1. Características das imagens de satélite Sentinel-2^a (resolução de 10 m) nas datas das áreas de estudo. Santa Rita do Trivelato – MT, 2019.

Data de aquisição	Estádio Fenológico	Ângulo médio do azimute do sol	Ângulo médio do zênite do sol	UR (%)
D1 – 08/03/19	B (botão floral)	27.85	73.30	86.95
D2 – 17/04/19	F (florescimento)	34.10	45.16	84.66
D3 – 07/05/19	C (abertura de capulho)	38.40	37.17	79.53
D4 – 11/06/19	C (abertura de capulho)	43.87	32.44	51.16
D5 – 16/07/19	C (abertura de capulho)	44.22	32.53	47.38

Fenologia e soma térmica do algodão

Compreender os efeitos do estresse térmico sobre as culturas é extremamente importante, devido às grandes variações de temperaturas, que podem ocorrer de um dia para o outro ou até entre períodos de um mesmo dia. A soma térmica dos graus dias seguiram equação proposta por Arnold (1959), utilizando temperatura basal de 15° C para a cultura do algodão segundo Rosolem (2020).

$$ST = \sum GD = \frac{(TM - Tm)}{2} - Tb \quad (1)$$

Em que: ST é a soma térmica; GD é graus-dias; TM é temperatura máxima diária, e Tm é temperatura mínima diária; e Tb é a temperatura base (15° C).

Dentre os fatores ambientais, a temperatura é um dos parâmetros que mais influenciam o desenvolvimento do algodoeiro. Exemplo claro disso é a utilização do monitoramento da fenologia da cultura (Fig. 3) por meio da somatória dos graus-dia que foi realizada neste estudo, pois temperaturas médias mais elevadas resultam no adiantamento do ciclo da cultura (Reddy et al., 1996).

A temperatura ótima para o desenvolvimento do algodão está na faixa entre 20°C e 30°C, no entanto, é comum o algodoeiro ser cultivado em regiões com temperaturas abaixo de 15°C (ex. EUA - período de emergência) e acima de 40°C (ex. Índia) (Snider e Kawakami, 2014).

O ciclo do algodoeiro é dividido em quatro fases: a vegetativa (V), formação de botões florais (B), abertura de flores (F) e abertura de capulhos (C) (Rosolem, 2020). Na Figura 3 podemos observar detalhadamente as fases fenológicas do algodão descritas por Aparecido et al. (2020).

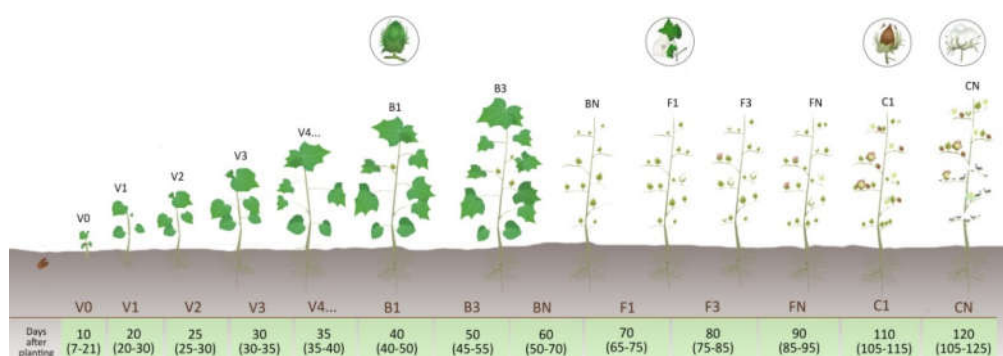


Figura 3. Média fenológica da cultura do algodão. Legenda: V0 = início da emergência da planta; V1 = V0 para a nervura principal da segunda folha; V2 = V1 à nervura principal da terceira folha; V3 = V2 à nervura principal da quarta folha; V4 = V3 à nervura principal da quinta folha; B1 = primeiro botão floral visível; B3 = primeiro botão floral visível no terceiro ramo; F1 = abertura da primeira flor no primeiro ramo do fruto; F3 = abertura da primeira flor no terceiro ramo do fruto e C1 = primeira maçã do primeiro ramo aberto (capulho de algodão). Fonte: Aparecido et al. (2020).

Índices de vegetação – preditores de entrada

Foi realizada malha amostral de 10.000 pontos em cada campo de estudo para a extração dos valores espectrais das bandas das imagens, juntamente com o valor de produtividade no mesmo ponto. Após a extração das informações nos 10.000

pontos às informações geolocalizadas das bandas 2,3,4,5 e 8 (Blue, Green, Red, RedEdge1 e NIR), calculou-se em linguagem Python os índices de vegetação (IV) para cada área em cada data das imagens obtidas, conforme esquema da Figura 4.

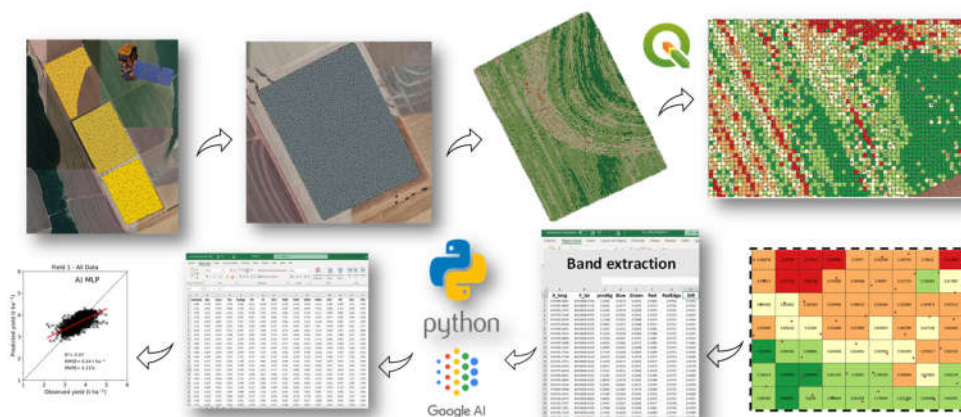


Figura 4. Esquema da criação de pontos e interpolação dos dados de produtividade e extração dos valores dos pixels vetorizados dos valores espectrais das bandas 2 (Blue), 3 (Green), 4 (Red), 5 (RedEdge), 8 (NIR) das imagens do Sentinel 2A, totalizando 10.000 feições por campo de estudo.

Os índices de vegetação utilizados neste estudo com suas respectivas equações (2 a 9) estão listados na Tabela 2.

Tabela 2. Índices de vegetação calculados no estudo como preditores de entrada da produtividade do algodão.

Índice de Vegetação	Equação	Referências
Simple Ratio	$SR = \frac{NIR}{R}$	(2) Birth and McVey (1968)
Normalized Difference Vegetation Index	$NDVI = \frac{NIR - R}{NIR + R}$	(3) Rouse et al. (1973)
Normalized Difference Red Edge Index	$NDRE = \frac{NIR - RE1}{NIR + RE1}$	(4) Barnes et al. 2000
Red Green Blue Vegetation Index	$RGBVI = \frac{G^2 - (B \cdot R)}{G^2 + (B \cdot R)}$	(5) Bendig et al. 2015
Soil Adjusted Vegetation Index	$SAVI = \frac{NIR - R}{(NIR + R + L)} \cdot (1 + L)$	(6) Huete, 1988
Infrared Percentage Vegetation Index	$IPVI = \frac{NIR}{(NIR + R)}$	(7) Crippen, 1990
Enhanced Vegetation Index 1	$EVI 1 = 2.5 * \frac{(NIR - R)}{(NIR + 6 * RED - 7.5 * B + 1)}$	(8) Huete et al. (1997)
Enhanced Vegetation Index 2	$EVI 2 = 2.5 * \frac{(NIR - R)}{NIR + (2.4 * R) + 1}$	(9) Huete et al. (1997)

Em que: NIR = refletância no infravermelho próximo, R = refletância no Vermelho, RE1 = refletância no Red Edge1, G = refletância no Verde, B é a refletância no Azul e L é um coeficiente empírico de ajuste para solos (0,5).

Produtividade observada – variável y que o modelo tentará prever

A colheita mecanizada do algodão foi realizada em julho de 2019 usando colhedora de fusos modelo CP690, que possui tecnologia de módulos enfardadores internos possibilitando-a colher, prensar, enfardar e depositar os módulos no campo sem parar a máquina. A colhedora possui sensor de rendimento e os dados foram coletados da máquina e exportados usando o software Ag Leader SMS Basic. Os dados foram exportados sem pré-processamento, sendo necessário eliminar valores incoerentes em cada campo.

Para essa filtragem utilizamos metodologia proposta por Menegatti e Molin (2004). Por fim, analisamos os dados filtrados e coerentes por meio de estatística descritiva. Devido a grande quantidade de dados, para a interpolação e geração dos mapas, fizemos novamente outra filtragem seguindo metodologia sugerida por Tukey (1977), para remoção de valores discrepantes, que, com base nos quartis, calcula-se o limite superior e inferior da produtividade de algodão em caroço ($t\ ha^{-1}$).

Um buffer de -50 m foi realizado no polígono de cada campo para assegurar que os valores de refletância espectral em cada pixel das imagens orbitais sejam oriundos apenas da refletância do dossel das plantas do algodoeiro, sem influência das áreas circunvizinhas, como estradas ou talhões adjacentes.

Os arquivos de dados vetoriais possuíam de 108.000 a 162.700 feições/pontos em cada campo de estudo, sendo assim, esses dados foram rasterizados interpolando-os pela Ponderação do Inverso das Distâncias (IDW), por meio do software QGIS versão 3.12.3 (QGIS, 2009), com grade de pixel de 10 x 10 m para coincidir com o tamanho do pixel das imagens do Sentinel-2A.

Os valores de produtividade interpolada constituem metas que o modelo está tentando prever durante o treinamento dos algoritmos avaliados conforme esquema demonstrado na Figura 3. Assim, os valores de produtividade foram extraídos dos pixels juntamente com os dados espectrais de cada banda usando o Point Sampling Tool (plugin do software QGIS), para que os as entradas e alvos possuam geolocalização correspondentes.

A avaliação da normalidade dos dados de produtividade foi realizada por meio do teste de Ryan-Joiner.

Correlação, seleção e padronização das variáveis preditoras

Utilizou-se a Correlação de Pearson entre as variáveis, a fim de compreender a relação entre os índices de vegetação e a produtividade de algodão.

Utilizou-se Stepwise Regression (Forward) para selecionar a variável de maior importância para a variável resposta (produtividade observada), e foi conduzida para resolver um problema de multicolinearidade no modelo de regressão. Sendo assim, este método fornece uma triagem inicial das variáveis candidatas quando se tem um grande grupo de variáveis, selecionando assim a(as) variável(eis) que possui(em) o maior R^2 .

A padronização de um conjunto de dados é um requisito comum para muitos estimadores de aprendizado de máquina. Sendo assim, padronizou-se o banco de dados transformando-os em desvio padrão que variam de 1 a -1 (Pedregosa et al., 2011).

Algoritmos de *machine learning* para predição de produtividade do algodão

Cenários foram criados para facilitar a análise e interpretação dos resultados. *Cenário 1*: All Data – todos os dados (40 inputs) e *Cenário 2*: a variável mais importante selecionada por Stepwise Regression (1 input), conforme Tabela 3.

Utilizamos diferentes algoritmos para realizar a previsão da produtividade do algodão. Em todos os casos, a produtividade foi a variável dependente. Os resultados dos oito índices de vegetação (IV) nas cinco datas de avaliação, do estágio de botão floral à abertura de capulho, foram as variáveis independentes do cenário All Data. A melhor variável deste conjunto, foi a variável independente do Cenário 2.

Os algoritmos usados para prever a produtividade do algodão foram: 1) Linear Regression; 2) Multilayer Perceptron (MLP); 3) Random Forest Regressor (RF); 4) K-Nearest Neighbor Regressor (KNN) e 5) AutoML: Aprendizado de máquina automático.

Tabela 3. Variáveis preditoras utilizadas nos algoritmos de machine learning para predição de produtividade do algodão. Cenário 1 All Data contempla todos os índices de vegetação em todas as datas avaliadas. Cenário 2: contempla uma única variável de maior importância selecionada por Stepwise Regression.

	CAMPO 1	CAMPO 2	CAMPO 3
Cenário 1 All Data	SR-D1, SR-D2, SR-D3, SR-D4, SR-D5, NDVI-D1, NDVI-D2, NDVI-D3, NDVI-D4, NDVI-D5, NDRE-D1, NDRE-D2, NDRE-D3, NDRE-D4, NDRE-D5, RGBVI-D1, RGBVI-D2, RGBVI-D3, RGBVI-D4, RGBVI-D5, SAVI-D1, SAVI-D2, SAVI-D3, SAVI-D4, SAVI-D5, IPVI-D1, IPVI-D2, IPVI-D3, IPVI-D4, IPVI-D5, EVI-D1, EVI-D2, EVI-D3, EVI-D4, EVI-D5, EVI2-D1, EVI2-D2, EVI2-D3, EVI2-D4, EVI2-D5	SR-D1, SR-D2, SR-D3, SR-D4, SR-D5, NDVI-D1, NDVI-D2, NDVI-D3, NDVI-D4, NDVI-D5, NDRE-D1, NDRE-D2, NDRE-D3, NDRE-D4, NDRE-D5, RGBVI-D1, RGBVI-D2, RGBVI-D3, RGBVI-D4, RGBVI-D5, SAVI-D1, SAVI-D2, SAVI-D3, SAVI-D4, SAVI-D5, IPVI-D1, IPVI-D2, IPVI-D3, IPVI-D4, IPVI-D5, EVI-D1, EVI-D2, EVI-D3, EVI-D4, EVI-D5, EVI2-D1, EVI2-D2, EVI2-D3, EVI2-D4, EVI2-D5	SR-D1, SR-D2, SR-D3, SR-D4, SR-D5, NDVI-D1, NDVI-D2, NDVI-D3, NDVI-D4, NDVI-D5, NDRE-D1, NDRE-D2, NDRE-D3, NDRE-D4, NDRE-D5, RGBVI-D1, RGBVI-D2, RGBVI-D3, RGBVI-D4, RGBVI-D5, SAVI-D1, SAVI-D2, SAVI-D3, SAVI-D4, SAVI-D5, IPVI-D1, IPVI-D2, IPVI-D3, IPVI-D4, IPVI-D5, EVI-D1, EVI-D2, EVI-D3, EVI-D4, EVI-D5, EVI2-D1, EVI2-D2, EVI2-D3, EVI2-D4, EVI2-D5
Cenário 2 Stepwise	SR-D1	SR-D3	RGBVI-D1

A LR é a abordagem mais simples e rápida que foi utilizada no estudo. LR ajusta um modelo linear com coeficientes para minimizar a soma residual dos quadrados entre os alvos observados no conjunto de dados e os alvos previstos pela aproximação linear (Pedregosa et al. 2011).

KNN é uma técnica simples, de fácil implementação e muito flexível, este método encontra um grupo de k amostras (dados de treinamento) mais próximo de amostras desconhecidas (dados de teste). Nas amostras k, as amostras desconhecidas são determinadas pela média da variável de resposta (Aparecido et al. 2020).

A RF é uma técnica de ML muito utilizada para previsão de safras (Everingham et al. 2016). As Florestas aleatórias são uma combinação de preditores de árvore, de modo que cada árvore depende dos valores de um vetor aleatório amostrado de forma independente e com a mesma distribuição para todas as árvores na floresta (Breiman, 2001).

A MLP foi a rede neural artificial utilizada, em que se define um mínimo de três camadas: as camadas de entrada, ocultas e de saída. A função de ativação utilizada foi a linear retificada (ReLU). O treinamento da MLP foi a backpropagation

(retropropagação dos erros). A propagação de retono é uma forma de aprendizado supervisionado em que a taxa de erro é enviada de volta pela rede para alterar os pesos para melhorar a previsão e diminuir o erro (Kaul et al. 2005).

Para o ML automático, utilizou-se o AutoML da plataforma H2O, que foi projetado para ter o mínimo de parâmetros possível, de modo que o usuário precise apenas apontar o conjunto de dados, identificar a coluna de resposta e se optar, definir um limite no número total de modelos treinados. De maneira rápida, o AutoML realiza uma pesquisa de hiperparâmetros em uma variedade de algoritmos H2O para fornecer o melhor modelo (LeDell & Poirier, 2020).

Para fornecer base de comparação, o mesmo conjunto de preditores foi mantido para todos os algoritmos avaliados, separando-os aleatoriamente em 70% dos dados para treinamento e 30% para teste.

Todos os hiperparâmetros utilizados estão apresentados no Apêndice A. Para os algoritmos MLP, RF e KNN, otimizamos os modelos com hiperparâmetros utilizando GridSearchCV (Pedregosa et al., 2011).

Avaliação do desempenho dos algoritmos

Os dados de produtividade de campo observada e predita em todos os modelos foram comparados usando índices estatísticos em relação a acurácia e a precisão.

A acurácia indica a proximidade de uma estimativa do valor observado e foi avaliada usando a raiz do erro quadrático médio (RMSE) dado na mesma unidade da variável predita, ou seja, em $t\ ha^{-1}$, e o erro médio percentual absoluto (MAPE), eq. 10 e 11.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Yobs_i - Yest_i)^2}{n}} \quad (10)$$

$$MAPE = \frac{\sum_{i=1}^n \left(\left| \frac{Yest_i - Yobs_i}{Yobs_i} \right| * 100 \right)}{n} \quad (11)$$

Em que: n é o número de dados, $Yest_i$ é o valor da produtividade estimada pelo algoritmo e $Yobs_i$ é o valor da produtividade real observada.

A precisão é a capacidade de um modelo de repetir uma estimativa e foi avaliada usando o coeficiente de determinação R^2 ajustado (CORNELL; BERGER, 1987), eq. 12.

$$R^2_{ajustado} = 1 - \frac{(1 - R^2) \times (n - 1)}{n - k - 1} \quad (12)$$

Resultados e Discussão

Produtividade observada x variáveis preditoras

A produtividade média de algodão em caroço observada nos campos de estudo variou de 3.20 a 3.65 t ha⁻¹ (Tabela 4). A estatística descritiva é apresentada na Tabela 4 e observamos que os dados de produtividade se comportaram conforme distribuição normal (valores próximos a 1) pelo teste de Ryan-Joiner nos três campos de estudo.

A maior produtividade média foi observada no campo 1 (3.65 t ha⁻¹) com o menor coeficiente de variação de 9.31 %, podendo ser observado no mapa de produtividade (Fig. 5). Essa baixa variação por tratar de dados comerciais a nível de fazenda. Apesar da baixa variação na média de produtividade entre os campos avaliados, esses resultados se encontram abaixo da média do estado de Mato Grosso que foi de 4.29 t ha⁻¹ (CONAB, 2020).

Tabela 4. Estatística descritiva dos dados de produtividade de algodão em caroço observada nos campos de estudo na fazenda em Santa Rita do Trivelato, Mato Grosso, Brasil, 2019.

Descrição	Campo 1	Campo 2	Campo 3
Ryan-Joiner	0.99 ^N	0.99 ^N	0.98 ^N
Média (t ha ⁻¹)	3.65	3.20	3.48
Mediana(t ha ⁻¹)	3.65	3.21	3.47
Desvio Padrão	0.34	0.32	0.45
CV (%)	9.31	10.12	12.92
Mínimo	2.16	1.54	1.46
Máximo	5.36	5.10	5.31
Quartil1	3.45	3.01	3.27
Quartil3	3.87	5.01	5.31
Distância Interquartílica	0.42	0.39	0.48

^N: distribuição normal dos dados; CV: Coeficiente de variação.

A fazenda já adota algumas técnicas de agricultura de precisão entre elas o mapeamento da produtividade para geração de mapas interpolados. Os valores de produtividade observada foram extraídos dos pixels dos mapas interpolados e se constituíram de metas que o modelo tentou prever durante o treinamento dos algoritmos. Como visto na análise descritiva, também observamos nos mapas que as menores produtividade foram observadas no campo 2, e a maior variabilidade espacial da produtividade ocorreu no campo 3. O campo com menor variabilidade dos valores de produtividade foi o campo 1.

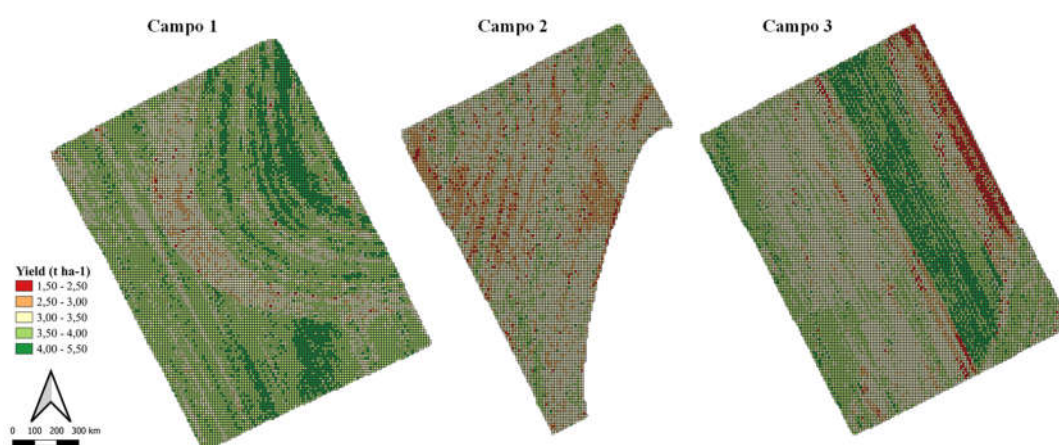


Figura 5. Representação dos mapas de produtividade observada nos campos de estudo na safra de 2019 em fazenda próxima a cidade de Santa Rita do Trivelato – MT, Brasil.

Oito índices de vegetação (SR, NDVI, NDRE, RGBVI, SAVI, IPVI, EVI e EVI2) foram calculados em cinco datas correspondente a 08/03, 17/04, 07/05, 11/06 e 16/06 no ano de 2019 que serão denominadas no texto como D1, D2, D3, D4 e D5.

Na Tabela 5 observamos o desenvolvimento do algodão para cada campo de estudo, e nota-se que a emergência do algodão ocorreu com 9, 6 e 17 dias após a semeadura, respectivamente para os campos 1, 2 e 3, o que resultou em somatória de graus dias acumuladas nas plantas em quantidades diferentes, resultando na variação de soma térmica e estágio fenológico entre os campos, o que refletirá nas variáveis mensuradas neste estudo.

Observa-se que os estádios fenológicos referentes a cada data de aquisição das imagens variaram entre a formação de botões florais (B1, B6 e B9), abertura de flores/florescimento (F5, F7, F8 e F10) e abertura de capulhos (C8, C9, C10, C12, C14), para os campos 1, 2 e 3, conforme Tabela 5.

Tabela 5. Informações sobre as datas de semeadura, emergência, estágio fenológico, colheita, ciclo e população final em função das datas de aquisição das imagens para os campos de estudo na fazenda em Santa Rita do Trivelato – MT.

	Campo 1			Campo 2			Campo 3		
Data da Imagem	Semeadura: 12/01/19 Emergência: 21/01/19 Colheita: 23/07/19 Ciclo: 183 dias Pop. Final 55,476.00 Área: 137 ha			Semeadura: 11/01/19 Emergência: 17/01/19 Colheita: 31/07/19 Ciclo: 195 dias Pop. final: 53,390.33 Área: 102 ha			Semeadura: 14/01/19 Emergência: 31/01/19 Colheita: 27/07/19 Ciclo: 177 dias Pop. final: 57,390.33 Área: 159 ha		
Sentinel	DAS(DAE)	EF	ST	DAE(DAE)	EF	ST	DAE(DAE)	EF	ST
08/03/2019	55(46)	B6	606	56(50)	B9	615	53(35)	B1	569
17/04/2019	95(86)	F8	1015	96(90)	F7	1024	93(76)	F5	978
07/05/2019	115(106)	C1	1223	116(110)	C3	1238	113(96)	F10	1191
11/06/2019	150(141)	C10	1596	151(145)	C12	1605	148(131)	C8	1558
16/06/2019	155(146)	C12	1652	156(150)	C14	1682	153(136)	C9	1636

Pop. Final: população final; DAS: dia após semeadura; DAE: dias após emergência; EF: estágio fenológico; ST: soma térmica. B: botão floral; F: florescimento; C: abertura de capulho

Essa variação ocasionou respostas espectrais diferentes em cada campo de estudo, conforme apresentado na Figura 6, sendo refletidas nos índices de vegetação em cada campo, como mostra a correlação de Pearson proporcional entre a produtividade e IVs durante as cinco datas analisadas.

Todas os índices de vegetação foram significativos positivamente ao nível de significância de 1% e as correlações variaram de 16 a 44% (campo 1), 18% a 52% (campo 2), 12 a 31% (campo 3), sendo as maiores no campo 2 e as menores no campo 3.

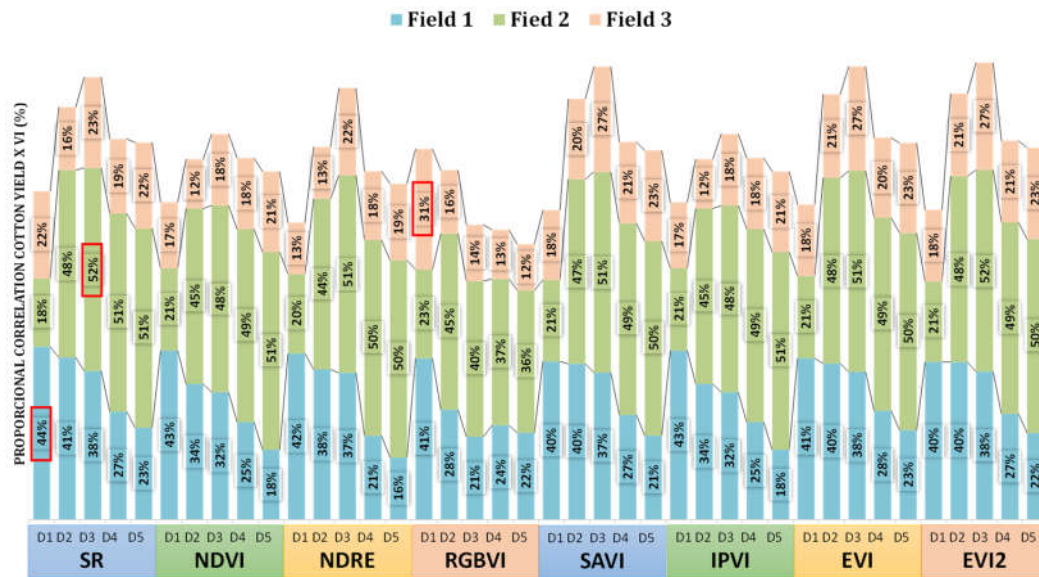


Figura 6. Correlação de Pearson Proporcional entre a produtividade observada e índice de vegetação na sequência das cinco datas de aquisição das imagens para os campos 1, 2 e 3.

Segundo classificação de Cohen (1988), que classifica como forte (0,50 e 1), moderada (0,30 a 0,49) e fraca entre (0,10 e 0,29) a relação entre duas variáveis, sendo assim, nota-se que apenas o campo 2 apresentou forte correlação com valores acima de 50%, e apenas nas datas D2, D3, D4 e D5.

Por meio da correlação nota-se dois aspectos importantes: 1) Todos os índices apresentaram comportamento semelhante entre si durante as todas as datas para o mesmo campo. 2) No campo 1 as maiores correlações foram observadas na D1, porém, no campo 2 a D1 apresentou as menores correlações. 3) No campo 3 as correlações apresentaram correlação fraca, e os maiores valores permaneceram na data 3. Ou seja, os valores se correlacionam com a produtividade, mas variam em cada campo para um mesmo índice de vegetação avaliado.

Seleção e padronização das variáveis preditoras

O método de seleção por Stepwise é direto e simples, pois inicia-se sem nenhuma variável candidata no modelo até a completa análise das variáveis entre si. Seleciona-se a(as) variável(eis) que possui(em) o maior R-quadrado ou, à cada etapa,

seleciona-se a combinação de variáveis candidatas que mais aumenta o R ao quadrado.

No campo 1, a variável selecionada foi o SR na D1. No campo 2 a variável de maior importância foi novamente o SR, porém na D3. E no campo 3 a variável selecionada por stepwise foi o RGBVI na D1. Nota-se que esses índices de vegetação SR e RGBVI conseguiram prever nas datas 1, 2 e 3 a produtividade do algodão com boa acurácia (baixo RMSE).

Algoritmos x Predição de produtividade do algodão

Os algoritmos demonstraram acurácia e precisão diferentes na previsão de produtividade de algodão nos cenários e campos avaliados (Tabela 6) e os melhores desempenhos estão destacados.

Cenário 1) All Data: Para o campo 1 a melhor performance foi a MLP, apresentando boa precisão e baixo erro (R^2 : 0.47; RMSE: 0.24 t ha⁻¹) na predição de produtividade do algodão em caroço a nível da fazenda (MLP) resultando no menor MAPE de todos os cenários e campos avaliados, de apenas 5.11%.

Para os campos 2 e 3, a melhor performance foi obtida pela RF, com o menor RMSE obtido encontrado em todos os cenários e campos, de apenas 0.23 t ha⁻¹.

O campo 3 apresentou as piores métricas, com RMSE de 0.1 a 0.2 t ha⁻¹ a mais que os outros campos. O MAPE também foi maior atingindo 2.82% a mais quando comparado ao campo 1 e 2, variando de 7.54 a 9.86 %.

Tabela 6. Avaliação de desempenho por RMSE e MAPE (entre parênteses) para os algoritmos de aprendizado de máquina avaliados na predição de produtividade do algodão na safra 2019 nos campos 1, 2 e 3.

Campo	Cenário	Input (entrada)	RMSE (MAPE) - t ha ⁻¹ (%)				
			RLM	MLP	RF	KNN	AutoML
1	1. All Data	40	0.29 (6.06)	0.24(5.11)	0.25(5.29)	0.25(5.37)	0.28(6.09) ^{GLM}
	2. Stepwise	SR-D1	0.31(6.66)	0.30(6.45)	0.30(6.42)	0.28(6.21)	0.31(6.74) ^{GLM}
2	1. All Data	40	0.27(6.29)	0.25(5.98)	0.23(5.61)	0.24(5.70)	0.26(6.16) ^{GLM}
	2. Stepwise	SR-D3	0.28(6.72)	0.27(6.45)	0.27(6.40)	0.26(6.17)	0.28(6.51) ^{DL}
3	1. All Data	40	0.40(8.88)	0.36(7.86)	0.34(7.54)	0.36(7.78)	0.38(8.15) ^{DL}
	2. Stepwise	RGBVI-D1	0.43(9.86)	0.42(9.60)	0.42(9.59)	0.40(9.17)	0.42(9.63) ^{DL}

RLM: Regressão Linear Múltipla, **MLP:** Multilayer Perceptron, **RF:** Random Forest Regressor, **KNN:** K-Nearest Neighbor Regressor; **AutoML:** Aprendizado de máquina automático; **DL:** Deep Learning-1-AutoML; **GLM:** Generalized Linear Model-1-AutoML; **1 SR:** Índice SR na data 1 (08/03/2019) com 606 de soma térmica, **3 SR:** Índice SR na data 3 (07/05/2019) com 1238 de soma térmica; **1 RGBVI:** Índice RGBVI na data 1 (08/03/2019) com 569 de soma térmica.

Cenário 2: Variável selecionada por stepwise. Neste cenário, as melhores variáveis dentro o conjunto de dados com 40 inputs foram o índice SR na data 1 e 3, para o campo 1 e 2 respectivamente. Para o campo 3 a variável de maior importância em resposta à produtividade foi o RGBVI na D1.

Nota-se que independente das métricas, dentre um conjunto de variáveis preditoras que corresponderam a cinco datas antecedendo a colheita, observa-se que, as melhores foram na data 1 e 3, ou seja, antecedendo em 137, 85 e 141 dias a colheita dos campos 1, 2 e 3 respectivamente.

O índice SR, apesar de ser considerado de cálculo simples e de sofrer influências da atmosfera, apresentou as melhores correlações com a produtividade observada nos campos 1 e 2 deste estudo na maioria das datas. O campo 3 se comportou de maneira distinta, possivelmente devido ao atraso na emergência em relação aos outros campos (8 e 11 dias de atraso em relação ao campo 1 e 2 respectivamente), o que pode ter ocasionado estresse abiótico do algodão deste campo, resultando em menores valores de soma térmica que refletiu em maior variabilidade, desenvolvimento retroativo e baixas correlações dos índices com a produtividade observada nele.

Uma estrutura de aprendizado de máquina foi desenvolvida para estimativa de rendimento de algodão usando dados de sensoriamento remoto multi-temporais por Ashaputre et al. (2020) e eles compararam o desempenho do modelo de rede neural artificial (ANN) com regressão de vetor de suporte (SVR) e regressão de floresta aleatória (RFR). Os resultados da comparação revelaram que o modelo ANN supera SVR e RFR.

Pesquisadores encontraram R^2 variando de 0,13 a 0,44 e RMSE de 0,16 a 0,53 $t\ ha^{-1}$ corroborando com a algumas métricas do presente estudo utilizando dados públicos (dados climáticos e de solo) para prever a produtividade de algodão em campos comerciais na Austrália com os algoritmos RF e GBM (Gradient boosting machine) em três datas de imagens do satélite Landsat com o índice GNDVI (Leo et al., 2020).

O desempenho dos algoritmos na etapa de teste que apresentaram o melhor desempenho em cada cenário (All Data e Stepwise) são apresentados na Figura 7.

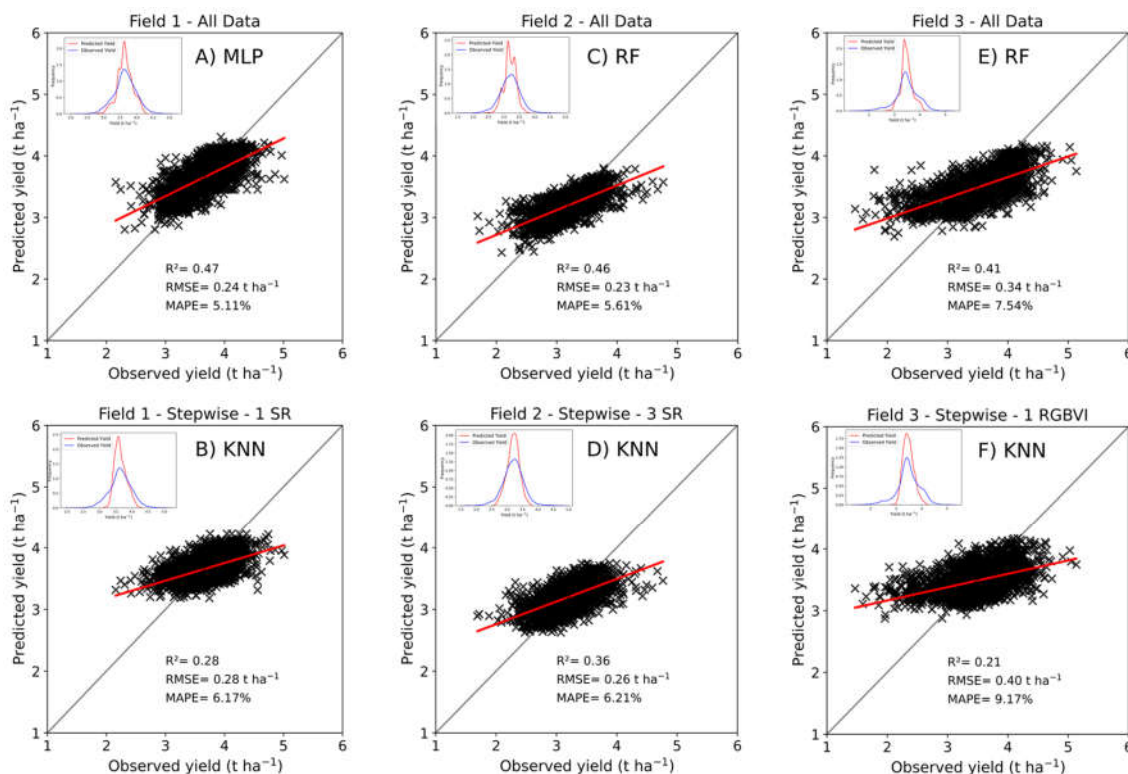


Figura 7. Relação dos valores observados e preditos para produtividade do algodão (t ha⁻¹) na etapa de teste para os melhores desempenhos apresentados como CAMPO 1: A) MLP no cenário 1: All Data e B) KNN para o cenário 2: SR-D1; CAMPO 2: C) RF para o cenário 1: All Data e B) KNN para o cenário 2: SR-D1; CAMPO 2: C) RF para o cenário 1: All Data e D) KNN para o cenário 2: SR-D3; CAMPO 3: E) RF para o cenário 1: All Data e D) KNN para o cenário 2: RGBVI-D1.

Dentre os algoritmos com ajustes de hiperparâmetros, que são análises de ajustes morosas e que requerem grande demanda computacional, comparando-o com a MLP e a RF (gridsearch mais moroso), o KNN foi o que apresentou processo mais satisfatório, ficando apenas depois do AutoML e da LR.

Destacamos um resultado importante no Cenário 2 que, para treinamento e teste usando apenas um único input, o algoritmo KNN Regressor apresentou o melhor desempenho em todos os campos de estudo, apresentando as melhores performances aos 46 DAE (D1), bem como 110 DAE (D3) que equivale a predição antecedendo a colheita em 143 dias na D1, usando um único índice em uma única data nos campos de estudo.

Os melhores desempenhos novamente ficaram similares entre o campo 1 e 2, com o melhor algoritmo sendo o KNN quando se usa 1 input, prevendo com 0,26 a 0,28 t ha⁻¹ de RMSE, valores, ou seja, com diferença de apenas 0.04 e 0.01 t ha⁻¹ de RMSE. Essa diferença é mínima por se tratar de área comercial com talhões de agricultura de larga escala no Mato Grosso, destacando que no cenário 2 reduzimos 39 variáveis de entrada e conseguimos prever com mínima diferença no RMSE.

O KNN de modo geral apresentou ótima performance neste estudo. Segundo Aparecido et al. (2020), que também avaliou esse algoritmo na predição de algodão a região centro-oeste do Brasil, o ponto positivo é que o KNN é um algoritmo bastante simples, pois baseia-se no cálculo da distância, ou seja, o algoritmo considera que novos pontos de dados são semelhantes aos dados do conjunto de treinamento para fazer sua classificação ou regressão. Na regressão, o KNN determina um valor de um determinado atributo da amostra de dados vizinhos do conjunto de treinamento.

Considerações Finais

Nos campos 1 e 2 do presente estudo a produtividade observada apresentou baixa variabilidade e os índices de vegetação em todas as datas de aquisição das imagens apresentaram correlação significativa e positiva com destaque para o índice SR que foi o de maior correlação e o índice selecionado por Stepwise Regression nestes dois campos como a variável de maior importância para a predição da produtividade de algodão.

Devido à diferença nas datas de emergência, o campo 3 apresentou resposta espectral e comportamento distinto em relação aos índices de vegetação dentre os campos avaliados. Esse fato influenciou na produtividade observada e na correlação das variáveis, resultando menor desempenho dos algoritmos quando comparado aos campos 1 e 2. Estudos adicionais serão realizados com melhores modelos para outros campos e locais de estudo para confirmar os resultados.

Além do melhor desempenho apresentado pelo algoritmo KNN, outro aspecto importante observado no estudo foi que esse algoritmo exigiu menor demanda computacional nos ajustes dos hiperparâmetros e calibração dos algoritmos, quando comparado a MLP e RF por exemplo.

Conclusão

O cenário 1 (40 inputs) apresentou as melhores métricas em todos os campos de estudo e os algoritmos Multilayer Perceptron e Random Forest apresentaram os melhores desempenhos, ressaltando que neste cenário necessitamos de todos os inputs preditores que foram gerados ao longo de todo o ciclo da cultura (aprox. 180 dias).

De modo geral o KNN apresentou a melhor performance para atingir o objetivo neste estudo. A diferença apresentada usando 1 input dentre 40 inputs na predição foi mínima por se tratar de área comercial com talhões de agricultura de larga escala no Mato Grosso.

Referências

Aparecido LEO, Meneses KC, Souza Rolim GR, Carvalho MJN, Pereira WBS, Silva PA, Santos TS, Moraes JRSCM (2020) Algorithms for forecasting cotton yield based on climatic parameters in Brazil. **Arch. Agron. Soil Sci.** Doi: 10.1080/03650340.2020.1864821

Arnold CY (1959) The determination and significance of the base temperature in a linear heat unit system. **Proceedings... The American Society for Horticultural Science**, Alexandria, v. 74, n.1 p. 430-445.

Ashapure A, Jung J, Chang A, Oh S, Yeom J, Maeda M, Maeda A, Dube N, Landivar J, Hague S, Smith W (2020) Developing a machine learning based cotton yield estimation framework using multi-temporal UAS data. **ISPRS J. Photogramm. Remote Sens.** 169, 180-194. Doi: 10.1016/j.isprsjprs.2020.09.015

Badnakhe MR, Durbha SS, Jagarlapudi A, Gade RM (2018) Evaluation of Citrus Gummosis disease dynamics and predictions with weather and inversion-based leaf optical model. **Comput Electron Agric** 155:130–141

Barnes EM et al. (2000) Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In: INTERNATIONAL CONFERENCE ON PRECISION AGRICULTURE, Bloomington. Madison: ASA: CSSA: SSSA, 2000.

Bendig J, Yu K, Aasen H, Bolten A (2015) Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. **Int J Appl Earth Obs Geoinf**, v. 39, p. 79–87, 2015. Doi: 10.1016/j.jag.2015.02.012.

Birth GS, McVey GR (1968) Measuring the color of growing turf with a reflectance spectrophotometer. **Agron. J.** 60 (6), 640–643. Doi: 10.2134/agronj1968.00021962006000060016x.

Breiman L (2001) Random Forests. *Machine Learning* 45, 5-32. Doi: 10.1023/A:1010933404324.

Cohen J (1988) *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Conab, Companhia Nacional de Abastecimento (2020). *Acomp. safra bras. grãos*, v. 7 - Safra 2019/20 - n. 5 - Quinto levantamento, Brasília, p. 1-25 fevereiro. Disponível em: <<https://www.conab.gov.br/info-agro/safras/graos>>.

Crippen RE (1990) Calculating the Vegetation Index Faster. *Remote Sensing of Environment*, V. 34, P. 71-73, 1990. Doi: 10.1016/0034-4257(90)90085-Z

Cornell JA, Berger RD (1987) Factors that Influence the Value of the Coefficient of Determination in Simple Linear and Nonlinear Regression Models. **Phytopathology**, v. 77, n. 1, p. 63, 1987.

Everingham Y, Sexton J, Skocaj D, Inman-Bamber G (2016) Accurate prediction of sugarcane yield using a random forest algorithm. **Agronomy for Sustainable Development**. 36(2):27. Doi:10.1007/s13593-016-0364-z.

Feng A, Zhou J, Vories ED, Sudduth KA, Zhang M (2020) Yield estimation in cotton using UAV-based multi-sensor imagery. **Biosyst. Eng.** 193, 101-114. DOI:10.1016/j.biosystemseng.2020.02.014

Fue KG, Porter WM, Rains GC (2018) Deep Learning based Real-time GPU-accelerated Tracking and Counting of Cotton Bolls under Field Conditions using a Moving Camera, In: 2018 Detroit, Michigan. American Society of Agricultural and Biological Engineers. Doi: 10.13031/aim.201800831.

Huete AR (1988) A soil-adjusted vegetation index (SAVI). **Remote Sensing of Environment**, New York, v. 25, n. 3, p. 295-309. Doi: 10.1016/0034-4257(88)90106-X

Huete AR, Liu HQ, Batchily K, Leeuwen van W (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. **Remote Sensing of Environment**, 59, 440–451.

Kaul M, Hill RL, Walthall C (2005) Artificial neural networks for corn and soybean yield prediction, **Agricultural Systems**, v. 85, n. 1, p. 1-18, 2005, ISSN 0308-521X, Doi: 10.1016/j.agsy.2004.07.009.

Kottek MJ, Grieser C, Beck B. Rudolf, and F. Rubel (2006) World Map of KöppenGeiger Climate Classification updated. **Meteorol. Z.**, 15, 259-263.

LeDell E, Poirier S (2020) *H2O AutoML: Aprendizado de máquina automático escalonável*. 7° ICML Workshop on Automated Machine Learning (AutoML) **In: ICML Workshop on Automated Machine Learning**, 7. Disponível em: <https://www.automl.org/wp-content/uploads/2020/07/AutoML_2020_paper_61.pdf>.

Leo S, Migliorati MA, Grace PR (2020) Predicting within-field cotton yields using publicly available datasets and machine learning. **Agron J.** 113:1150-1163. Doi: <https://doi.org/10.1002/agj2.20543>

Li F, Bai J, Zhang M, Zhang R (2022) Yield estimation of high-density cotton fields using low-altitude UAV imaging and deep learning. **Plant Methods**, 18, 1-11. DOI: <https://doi.org/10.1186/s13007-022-00881-3>.

Menegatti L, Molin JP (2004) Remoção de erros em mapas de produtividade via filtragem de dados brutos. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.8, n.1, p.126-134.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V. (2011). Scikit-learn: machine learning in Python. **J Mach Learn Res.** 12: 2825–2830. others.

Planet Team (2017) Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>

QGIS, Equipe de Desenvolvimento (2009) Sistema de Informação Geográfica QGIS. Fundação Geoespacial de Código Aberto. <http://qgis.org>

Reddy, V. R.; Hodges, H. F.; Mccarty, W. H.; Mckinnon, J. M. Weather and cotton growth: Present and Future. Mississippi Agr. & Forestry Exp. Sta., Mississippi State University, Starkville, MS.1996.

Rosolem, C. A. Crescimento do algodoeiro. p. 105. In: BÉLOT, J.; VILELA, P. **Manual de boas práticas de manejo do algodoeiro em Mato Grosso**. Cuiabá, 461 p. 4 ed. 2020. Disponível em: <http://www.casadoalgodao.com.br/images/publicacoes/manualdeboaspraticas2020-4ed-vf-web.pdf>.

Rouse Jr, J.W., Haas, R., Schell, J., Deering, D., 1973. Monitoring vegetation systems in the Great Plains with ERTS. In: NASA Special Publication 351. pp. 309–317.

Santos HG, JacominE PKT, Anjo LHC, Oliveira VA, Lumbreras JF, Coelho MR, Almeida JA, Araújo Filho, JC; Oliveira, JB, Cunha, TJF. **Sistema brasileiro de classificação de solos** (2018) Brasília, DF: Embrapa. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/199517/1/SiBCS-2018-ISBN-9788570358004.pdf>>. Acesso em: 10 maio. 2021.

Singh A, Ganapathysubramanian B, Singh AK, Sarkar S (2016) Machine learning for high-throughput stress phenotyping in plants. **Trends Plant Sci.** 21 (2), 110–124. Doi: 10.1016/j.tplants.2015.10.015

Snider J, Kawakami EM (2014) Efeito da temperatura no desenvolvimento do algodoeiro. In: Echer. F. R. **O algodoeiro e os estresses abiótico: Temperatura, luz, água e nutrientes**, 123 p.

Stackhouse P (2010) **Prediction of worldwide energy resource**. Hampton: NASA Langley Research Center. Disponível em: <<https://power.larc.nasa.gov/>>. Acesso em 01 jul. 2022.

Tedesco-Oliveira D, Silva RP, Maldonado-Jr W, Zerbato C (2020) Convolutional neural networks in predicting cotton yield from images of commercial fields. **Comput Electron Agric** 171 (2020) 105307. Doi: 10.1016/j.compag.2020.105307

Tukey JW (1997) **Exploratory data analysis**. 1 ed. Reading, Massachusetts, v.1, n.3.

USDA – United State Department of Agriculture (2020) **World Agricultural Production**. Circular Series WAP 2-20 February 2020. Disponível em: <<https://apps.fas.usda.gov/psdonline/circulars/production.pdf> >. Acesso em: 22 fev. 2020

Xu R, Li C, Paterson, AH, Jiang Y, Sun S, Robertson JS (2018). Aerial images and convolutional neural network for cotton bloom detection. **Front. Plant Sci.** 8, 1–17. Doi: 10.3389/fpls.2017.02235

Xu W, Chen P, Zhan Y, Chen S, Zhang L, Lan Y (2021) Cotton yield estimation model based on machine learning using time series UAV remote sensing data. **Int J Appl Earth Obs Geoinf**. Doi: <https://doi.org/10.1016/j.jag.2021.102511>

Yeom J, Jung J, Chang A, Maeda M, Landivar J (2018) Automated open cotton bol detection for yield estimation using Unmanned Aircraft Vehicle (UAV) Data. **Remote Sens.** 10, 1–20. Doi: 10.3390/rs10121895

CAPÍTULO 3 – É possível mapear os parâmetros de qualidade da fibra de algodão em campos comerciais?

RESUMO - A agricultura digital faz uso intensivo de tecnologias, e na colheita do algodão, novas máquinas que fornecem rastreabilidade e documentação da colheita em campo estão em ampla adoção. A hipótese deste estudo é a de que os parâmetros de qualidade da fibra podem ser rastreados e identificados por meio de tecnologias presentes nas colhedoras atuais, e auxiliar o produtor na tomada de decisão. Assim, o objetivo deste capítulo foi comparar métodos de amostragem de algodão para avaliar os parâmetros de qualidade da fibra para mapear e classificar estes parâmetros em campos comerciais no Brasil e nos Estados Unidos (EUA). O estudo foi realizado em duas etapas: Etapa 1 - em agosto de 2021, em uma área de 6,22 ha, no município de Sorriso - MT, Brasil; e Etapa 2 - em outubro de 2021, em uma área de 50 hectares situada em Newellton - LA, EUA. Em ambas as áreas foi realizada a colheita mecanizada, por uma colhedora John Deere, modelo CP690, ano 2018, com sistema de identificação por radiofrequência. Os três métodos de coleta de amostras de algodão para classificação da fibra foram: 1) Coleta manual na colhedora; 2) Coleta manual dos módulos no campo; 3) Coleta manual do fardo, após o beneficiamento. Os dados de produtividade obtidos pela colhedora foram filtrados para obtenção da média em cada módulo, por meio da área de polígonos. Foram coletadas amostras de algodão em caroço em 15 módulos na Etapa 1 e de 52 na 2. Todas as amostras foram identificadas pelo número do módulo, descaroçadas e os parâmetros (micronaire, resistência, comprimento, índice de uniformidade, índice de fibra curta, grau de cor e de folha) foram determinados utilizando o aparelho HVI (High Volume Instrument). Em seguida foram criados os polígonos referentes a área de coleta de cada módulo para a geração de mapas temáticos nas etapas 1 e 2, utilizado o software QGIS. A produtividade pode ser mapeada e classificada à nível de módulo. O método de coleta do fardo (algodoeira) apresentou a classificação de fibra com menor variabilidade. O processo de classificação ficou dentro dos limites superior e inferior de controle, apresentando variabilidade apenas entre os métodos de coleta, com exceção do parâmetro de classificação grau de cor e folha, que apresentaram diversos pontos fora de controle indicando que o processo não foi estável, coincidindo com a produção de fibras de padrão com qualidade inferior, fora da faixa de comercialização para o mercado internacional (31-4 ou inferior). É possível mapear os parâmetros de qualidade da fibra nas Etapas 1 e 2 desde que se vincule os dados de identificação do módulo com a etiqueta criada na algodoeira, para os fardos beneficiados. A conclusão deste trabalho é que esta técnica pode ser utilizada como ferramenta para os produtores visualizarem a variabilidade espacial da qualidade da fibra em seus campos comerciais, por meio de mapas, e tomarem decisões mais assertivas quando ao manejo da lavoura.

Palavras-chave: mapas de HVI, rastreador de colheita, módulo de algodão, rastreabilidade, RFID

IS IT POSSIBLE TO MAP COTTON FIBER QUALITY PARAMETERS IN COMMERCIAL FIELDS?

ABSTRACT – Digital agriculture makes intensive use of technologies, and in cotton harvesting, new machines that provide traceability and harvest documentation in-field are widely adopted. This study hypothesizes that the fiber quality parameters may be tracked and identified using technologies present in current harvesters and help the producer in decision-making. Thus, the objective of this chapter was to compare cotton sampling methods to evaluate fiber quality parameters to map and classify these parameters in commercial fields in Brazil and the United States (USA). The study was carried out in two stages: Stage 1 - in August 2021, in an area of 6.22 ha located in Sorriso - MT, Brazil; and Stage 2 - in October 2021, in an area of 50 ha located in Newellton - LA, USA. In both regions, mechanized harvesting was carried out using a John Deere harvester, model CP690, year 2018, with a radio frequency identification system (RFID). The three cotton sample collection methods for fiber classification were: 1) Manual collection on the harvester; 2) Manual collection of modules in the field; 3) Manual collection of the bale after processing. The yield data obtained by the harvester were filtered to get the average in each module through the area of polygons. Cotton samples were collected from 15 modules in Stage 1 and 52 in Stage 2. All samples were identified by module number, cotton was ginned, and the parameters (micronaire, strength, length, uniformity index, short fiber index, color, and leaf grade) were determined using the HVI classification (High Volume Instrument). Then, the polygons referring to the harvest area of each module were created for the generation of thematic maps in Stages 1 and 2 using the QGIS software. Yield may be mapped and classified at the module level. The bale collection method (gin) presented the fiber classification with less variability. The classification process was within the upper and lower control limits, showing variability only between the collection methods, except the classification parameter color and leaf degree, which gave several points out of control, indicating that the process was not stable, coinciding with the production of lower quality standard fibers, outside the commercialization range for the international market (31-4 or less). It is possible to map the fiber quality parameters in Stages 1 and 2 as long as the module identification data is linked with the label created in the gin for the processed bales. This work concludes that this technique can be used as a tool for cotton producers to visualize the spatial variability of fiber quality in their commercial fields through maps and make more assertive decisions regarding crop management.

Keywords: HVI maps, harvest identification cotton, cotton module, traceability, RFID

Introdução

O algodão origina a principal commodity de fibra natural do mercado cujo preço é ajustado conforme parâmetros de classificação da fibra, sendo a colheita mecanizada uma operação que está intimamente ligada à qualidade e lucratividade final do produto.

Após a colheita, a próxima etapa é o beneficiamento do algodão em caroço e classificação que determina a qualidade da fibra. É bem documentado que a variabilidade da fibra de algodão existe “entre e dentro” de cultivares e entre ambientes (qualidade do solo, estresses bióticos e abióticos dentre outros fatores) tanto no Brasil (Belot et al., 2017; Bachelier e Gourlot, 2018) como nos EUA (Johnson et al., 2002; Kothari et al., 2015; Ayele, et al., 2018; Porter, et al., 2017).

Numerosas características estão associadas à qualidade das fibras de algodão; as principais são: índice micronaire, comprimento, uniformidade, maturidade, finura, resistência à ruptura, reflectância, grau de amarelo, impurezas e contaminantes das fibras do algodão (Bachelier e Gourlot, 2018).

Sabendo da existência da variabilidade “intraplanta”, ainda assim, a variabilidade espacial dos parâmetros de qualidade da fibra não são avaliados e mapeados no âmbito de fazenda (campos comerciais) para melhoria nas tomadas de decisões, devido a não utilização de todas as tecnologias disponíveis na colhedora e no centro de operações e devido à falta de continuidade das informações geradas no campo com as geradas pela algodoeira.

Em 2011, a John Deere introduziu no mercado a colhedora JD 7760 que possui câmara de enfardamento de módulos interno possibilitando-a colher, prensar, enfardar e depositar o fardo no campo sem parar a máquina, o que revolucionou o armazenamento, o manuseio do algodão após a colheita e a capacidade operacional (Silva e Sofiatti, 2017). Ela também possui sistema RFID (Radio Frequency Identification), que por meio de etiquetas presentes no envoltório plástico (TamaWrap™), permitem a identificação no campo dos módulos redondos de algodão através de um sistema único de etiquetas (Fig. 1).

Cada módulo criado pela colhedora contém quatro etiquetas RFID e uma etiqueta externa adicionada que identifica exclusivamente o módulo com dados como:

número de série do módulo, ID do módulo, variedade, coordenadas geográficas de amarração e descida dos módulos, diâmetro, umidade, área colhida, peso do módulo, e produtividade, por meio do Rastreador de Colheita ou Harvest Identification – HID, Cotton Pro (John Deere, 2019).



Figura 1. O algodão é envolto em um módulo cilíndrico (A) com uma etiqueta RFID no envoltório plástico (B). Fotos: Francielle Morelli Ferreira.

A agricultura digital vem fazendo uso intensivo de tecnologias, e na colheita do algodão, as novas máquinas que fornecem rastreabilidade e documentação da colheita em campo estão em ampla adoção. O aumento da digitalização pode melhorar a produtividade, a sustentabilidade e a competitividade com fibras sintéticas (Hardin IV et al., 2022).

Os módulos colhidos geram diversas informações do campo, e na algodoeira, esse módulo gera cinco fardos que são classificados quanto aos parâmetros de qualidade da fibra e estas classificações irão servir de base para venda no mercado interno e externo desta fibra. E será que essa classificação sofre variação?

A avaliação da qualidade em processos ou operações agrícolas é possível por meio do uso de controle estatístico de processos (CEP). Associado à melhor qualidade nos resultados de classificação dos parâmetros de qualidade da fibra do algodão, o CEP torna-se fundamental para o monitoramento da qualidade nas etapas desse processo agrícola, em que as cartas de controle expressam os resultados graficamente de forma sequencial, permitindo verificar o comportamento médio, estabilidade e variabilidade ao longo da operação (SILVA *et al.*, 2008).

Por fim, os módulos redondos originados pelas colhedoras JD7760, CP690 são a maioria do algodão colhido e processado por descaroçadores nas regiões produtoras de algodão dos EUA e do Brasil, com cerca de 80% (Franco, 2020).

Entretanto, mesmo com esse nível de adoção, a rastreabilidade por meio do RFID e o completo gerenciamento dos módulos em campo atrelados à classificação da fibra não vem sendo realizado, uma vez que, para interrelacionar a qualidade da fibra ao local de colheita desta, as fazendas precisam interligar as informações de cada módulo com as etiquetas que a algodoeira cria no beneficiamento de cada módulo.

Nos últimos cinco anos, a Cotton Incorporated financiou nos EUA uma demonstração junto ao laboratório de descaroçamento do USDA-ARS em Lubbock, no Texas para permitir a opção de usar apenas a etiqueta RFID para rastrear o algodão do campo durante todo o processo de descaroçamento sem adicionar outras etiquetas ou tinta spray. Além disso, as informações de propriedade são inseridas apenas uma vez (Barnes, 2020).

No Brasil, a ID-COTTON surgiu como solução completa para rastreabilidade de algodão, da colheita até o beneficiamento com o uso da tecnologia RFID (IoP Journal, 2020), entretanto, não realizam o mapeamento da qualidade da fibra, apenas a rastreabilidade.

As etiquetas RFID e a automação de dados estão começando a permitir que os produtores deixem de depender de bilhetes de papel para rastrear o algodão do campo até o descaroçamento (Barnes, 2020).

Na busca por artigos e projetos abordando a rastreabilidade da fibra, foram encontradas pesquisas recentes do USDA em que um novo sistema de código aberto para gerenciamento de módulos de algodão da colhedora até o processo de descaroçamento, utilizando as etiquetas RFID integradas em módulos redondos, foi desenvolvido (Wanjura et al., 2017, Wanjura et al., 2018, Wanjura et al., 2020). O sistema tem a capacidade de importar e gerenciar dados específicos do módulo coletados das colhedoras de construção de módulos redondos da John Deere.

Em relação ao mapeamento da qualidade da fibra, trabalhos iniciais antes do RFID estar presentes nas colhedoras surgiram por Ge et al. (2008) e Ge et al. (2012) no Texas A&M University, e utilizando o RFID para o mapeamento estudos surgiram

na University of Georgia por Fuhrer et al. (2020) e Fuhrer (2022), e na North Carolina State University (Ward, 2020 e Ward et al. 2022).

Em relação ao Brasil, não se encontrou pesquisa científica em desenvolvimento com esse tema e abordagem, fazendo-se importante o presente estudo para se explorar mais da tecnologia RFID disponível nas colhedoras de algodão da John Deere (JD7760, CS/CP690 e CS/CP770) possibilitando a rastreabilidade, bem como o mapeamento completo da qualidade da fibra do algodão produzidas aqui.

Nesse contexto e na hipótese de que os parâmetros de qualidade da fibra possam ser rastreados e identificados na resolução de módulo por meio do RFID e do Rastreador de Colheita (HID Cotton) presentes nas colhedoras mais recentes, e de que a exploração dessas ferramentas auxiliarão o produtor nas tomadas de decisões, o objetivo geral deste capítulo foi avaliar métodos de coleta de algodão para determinar a qualidade da fibra a fim de responder se é possível mapear os parâmetros de classificação em campos comerciais no Brasil e nos Estados Unidos.

Os objetivos específicos deste capítulo foram:

- Coletar os números de identificação dos módulos (Module SN) para relacionar à etiqueta da algodoeira possibilitando seu acompanhamento no processo de descaroçamento e classificação das fibras.
- Manusear as informações do rastreador de colheita (HID) para geoespacializar a área colhida dos módulos.
- Avaliar a classificação HVI em diferentes métodos de coleta do algodão.
- Avaliar a qualidade no processo de classificação das fibras por meio de cartas de controle
- Gerar mapas com os parâmetros de qualidade da fibra em campos comerciais em Sorriso - MT, Brasil e em Newellton - LA, EUA.

Material e Métodos

Etapas e áreas de estudo

O estudo foi realizado em duas etapas, uma no Brasil em agosto de 2021 (etapa 1) e outra nos Estados Unidos (EUA) em outubro de 2021 (etapa 2).

Etapa 1 – Brasil

A etapa 1 foi conduzida no ano agrícola de 2020/2021, no município de Sorriso, Mato Grosso, Brasil. O campo comercial está localizado na Fazenda Santa Maria, nas coordenadas geográficas 12°53'54,43"S, 55°52'54,90"W (Figura 2), e a área de estudo tem 6,22 ha, com altitude média de 410 m. O solo da área foi classificado como Latossolo Vermelho-Amarelo Distrófico, seguindo os critérios estabelecidos pelo Sistema Brasileiro de Classificação dos Solos (Santos et al., 2018).

A cultivar utilizada foi a FM 985GLTP (tecnologia GlyTol LibertyLink Twinlink Plus®). A semeadura ocorreu em 28/01/2021 com população de 80.000 sementes/ha, espaçamento entre linhas de 0,90 m e o ciclo do algodão foi de 194 dias.

Etapa 2 – Estados Unidos (EUA)

A etapa 2 foi conduzida no ano agrícola de 2021, na cidade de Newellton em Louisiana, EUA. O campo comercial está localizado em Somerset Plantation, nas coordenadas geográficas 32°10'56"N, 91°16'46"W (Fig. 2). A área total avaliada foi de 50 hectares, com altitude média de 23 m, topografia com declive suave (até 5%), e o solo da área está classificado como Sharkey (Soil Survey Staff, 2014), apresentando textura argilosa (60-75 % de argila).

A cultivar utilizada foi a CROPLAN 9608 B3X, que é uma variedade de maturidade média e com tecnologia Bollgard® 3 XtendFlex®. A semeadura ocorreu em 16/05/2021 com população de 92.000 sementes/ha, espaçamento entre linhas de 0,91 m (36 inches), e o ciclo do algodão foi de 155 dias.

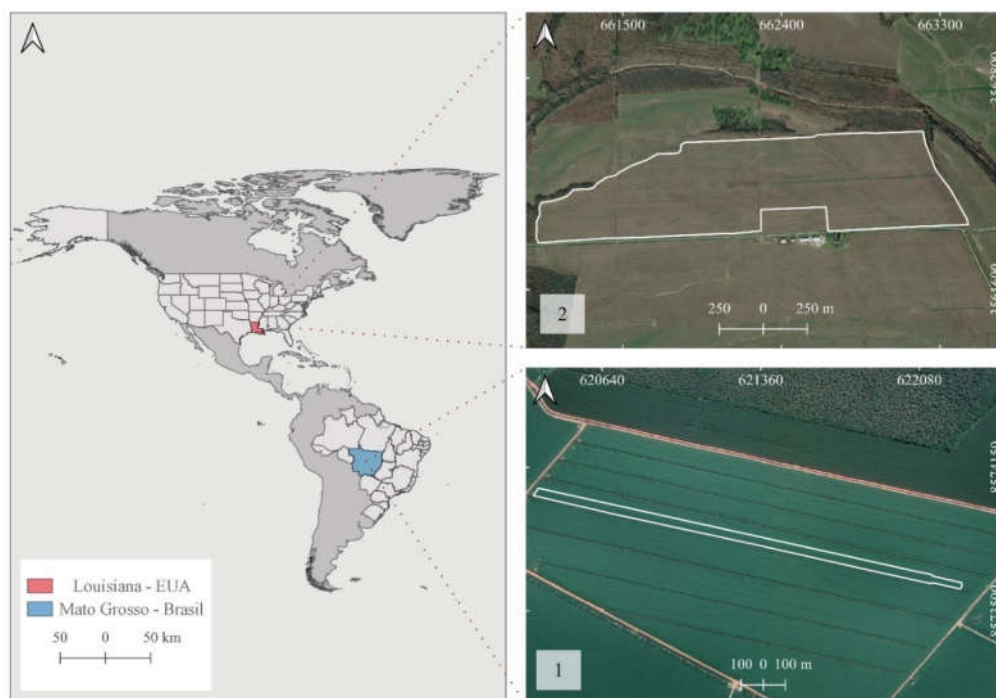


Figura 2. Localização das áreas de estudo na etapa 1 em Sorriso, Mato Grosso, Brasil, e na etapa 2 em Newellton, Louisiana.

A cultura foi monitorada por técnicos das propriedades agrícolas para controle de pragas e doenças, assim como todas as operações necessárias para o bom estabelecimento da cultura.

O clima na região da etapa 1 é o tropical quente e úmido, com inverno seco (Aw), já da etapa 2 é o subtropical úmido com verão quente (Kottek et al., 2006).

Durante o ciclo do algodão avaliado nos campos de estudo, a temperatura média foi de 24,46 °C (etapa 1) e 22,43 °C (etapa 2), com precipitação acumulada em 782 mm (etapa 1) e 1.193,54 mm (etapa 2), atendendo aos valores mínimos exigidos pela cultura em ambas as etapas.

Os dados meteorológicos influenciam diretamente nas condições de desenvolvimento do algodoeiro refletindo na classificação dos parâmetros de qualidade da fibra e são apresentados na Figura 3.

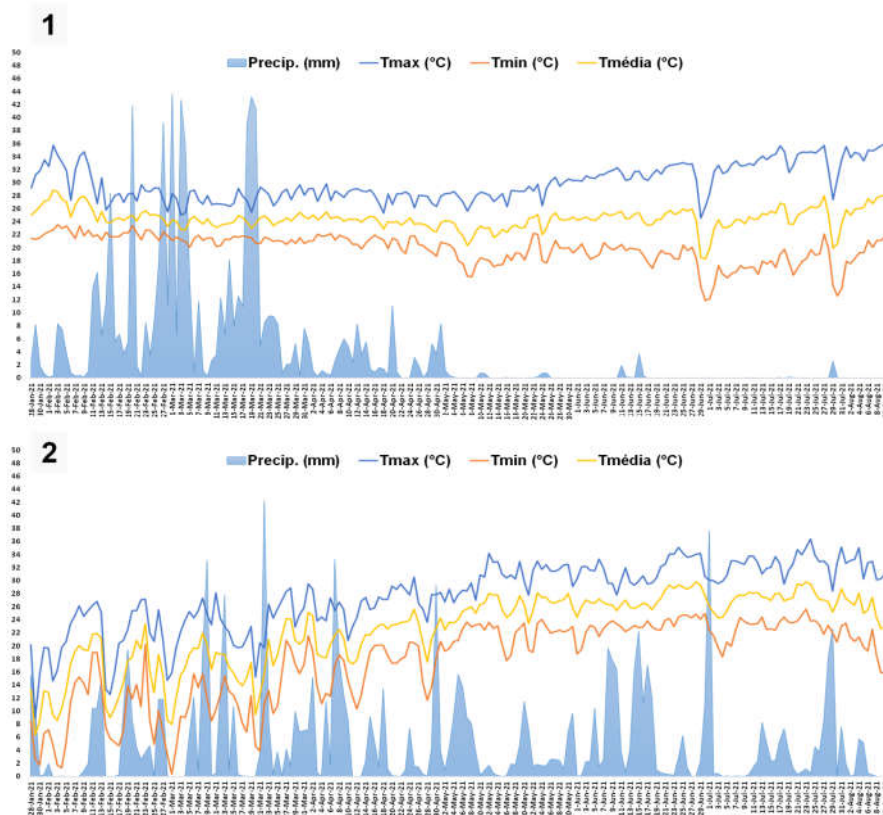


Figura 3. Precipitação (mm), temperatura máxima (°C), mínima (°C) e temperatura média (°C) desde as datas de semeadura até a data de colheita para os campos de estudo nas etapas do Brasil em Sorriso, MT (1) e nos Estados Unidos (2) em Newellton, LA.

Os dados climáticos foram obtidos em uma escala diária na National Aeronautics and Space Administration/Prediction of World Ampla plataforma de recursos de energia - NASA/POWER (Stackhouse, 2010) e se fez importante para interpretação dos resultados de qualidade da fibra em cada etapa de estudo.

Colheita e Produtividade do algodão

A colheita mecanizada foi realizada em 10/08/2021 (Etapa 1 – Brasil) e em 18/10/2021 (Etapa 2 – EUA), ambas utilizando colhedora de algodão, com sistema de fusos (sistema picker), marca John Deere, modelo CP690, ano 2018, com plataforma de 6 tambores/linhas de colheita totalizando 5,4 m de largura.

Os dados de produtividade das áreas de estudo foram exportados das colhedoras e devidamente filtrados utilizando o software MapFilter 2.0 (Maldaner et al., 2019).

A partir dos pontos de produtividade filtrados nos respectivos campos, e por meio da ferramenta “join attributes by location (summary)” no software QGIS obteve-se a média de produtividade em cada módulo por meio da área de cada polígono.

Métodos de coleta de algodão em caroço

Ao todo 67 módulos foram avaliados e as amostras de algodão em caroço foram coletadas manualmente, acondicionadas em saco de papel e identificadas quanto ao número de identificação dos módulos (MODULE SN) duas etapas da pesquisa.

Na etapa 1, os métodos utilizados estão descritos abaixo:

- Método 1 (M1) - Coleta manual na colhedora (durante a colheita): coletou-se uma amostra composta oriunda de três a quatro subamostras (em função do número de vezes que o acumulador enchia para cada módulo).

- Método 2 (M2) - Coleta manual do módulo no campo (após a colheita): coletou-se uma amostra composta por módulo oriunda de 10 subamostras coletadas em várias direções do módulo em campo).

- Método 3 (M3) - Coleta manual do fardo (após o beneficiamento): coletou-se uma amostra composta por módulo oriundas de cinco subamostras (uma de cada fardo gerado do módulo) que originou em uma amostra composta por módulo.

Por fim, na etapa 1, obteve-se quinze médias de quinze módulos como resultados de qualidade da fibra em cada método de coleta (Fig. 4), somando 45 amostras compostas avaliadas.



Figura 4. Método 1 - Coleta manual na colhedora (A); Método 2 - Coleta manual do módulo em campo (B e C); Método 3 – Coleta na algodoeira, após o beneficiamento (D).

Na etapa 2 (EUA) utilizou-se apenas os métodos M2 e M3 (Fig. 4B) e coletou-se uma amostra composta oriundas de dez subamostras retiradas de 52 (cinquenta e dois) módulos. No método 3 houve as coletas da algodoeira, porém os resultados das análises possuíam nova etiqueta de identificação (PBI) nos fardos gerados sem relação com o número do módulo, conseqüentemente sem informação das coordenadas de colheita dos módulos, impossibilitando-nos de associar a informação de qualidade de fibra com a localização dos módulos. Sendo assim, o M3 foi excluído dos resultados.

Beneficiamento e classificação da fibra do algodão

Todas as amostras de algodão em caroço coletadas (~350 g) foram descaroçadas para estarem aptas à classificação dos parâmetros de qualidade da fibra.

Na etapa 1, para as amostras dos métodos 1 e 2 utilizou-se descaroçadora de 10 serras apresentada na Figura 5A. O método 3 de coleta foi descaroçado com procedimento padrão da algodoeira (máquinas industriais).

Após o descaroçamento, todas as amostras foram enviadas ao Laboratório Kuhlmann de Classificação de Fibra (Sorriso, Brasil). As análises foram realizadas em aparelho HVI (High Volume Instrument).

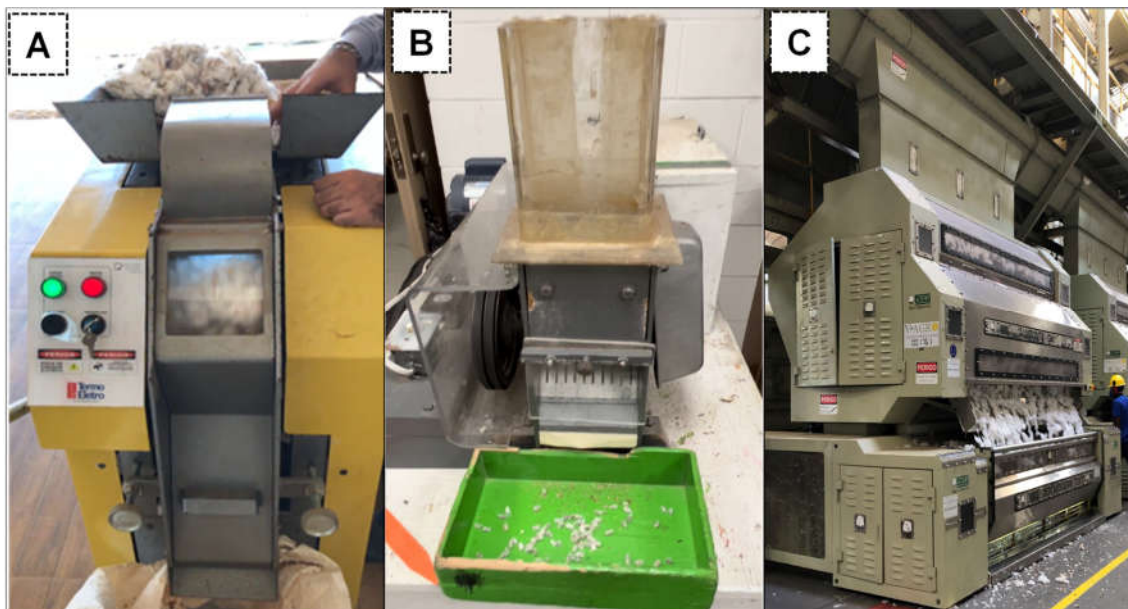


Figura 5. Descaroçadora de 10 serras utilizada nos métodos de coleta 1 e 2 no Brasil (A) e no método 2 nos EUA (B). Descaroçamento de procedimento padrão em algodoeirias utilizada no método 3(C).

Na etapa 2, as amostras de algodão coletadas no método 2 foram descaroçadas com descaroçadora de mesa de 10 serras (Fig. 5B) e as amostras beneficiadas foram enviadas para o Laboratório de Classificação de fibras da LSU AgCenter (Louisiana State University Agricultural Center) em Baton Rouge, USA. As análises da etapa 2 também foram realizadas em aparelho HVI.

Mapeamento da qualidade da fibra

Para entender quais pontos representam a colheita de cada módulo e conseguir geoespacializar a área colhida nas etapas 1 e 2, foi seguido o seguinte passo a passo no software QGIS:

1. Por meio dos dados vetoriais de produtividade, abriu-se a tabela de atributos e com o auxílio da calculadora de campo inseriu-se uma nova coluna com números sequenciais (seq), de acordo com cada ponto coletado pela máquina a partir da coluna data e hora.
2. A partir dos pontos sequenciais e por meio das coordenadas de amarração dos módulos, que coincidem com essa localização, categorizou-se numericamente os pontos referentes ao caminho percorrido pela máquina par a colheita de cada módulo (Figura 6).

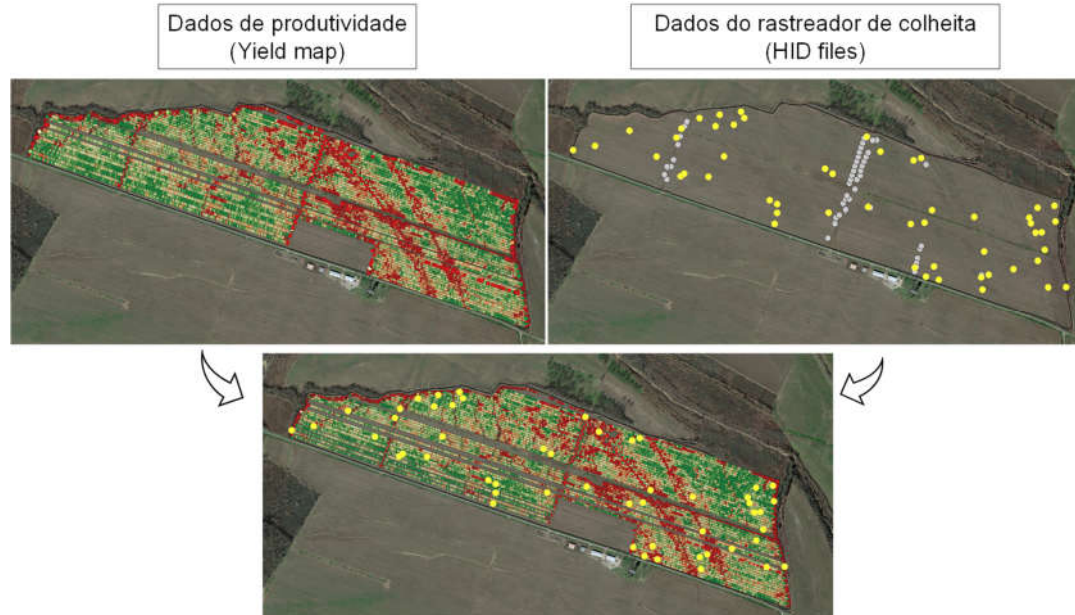


Figura 6. Esquema inicial para geração da área colhida dos módulos em função da nova coluna com números sequenciais (seq), de acordo com cada ponto coletado pela máquina a partir da coluna data e hora.

3. A partir da categorização do conjunto de pontos colhidos nas coordenadas referentes a cada módulo criou-se linhas por meio da ferramenta de conversão de “pontos para linha(s)”.
4. A partir do arquivo de linhas criado, realizou-se um buffer de 2,7 m, que equivale à metade da largura da plataforma da colhedora de algodão (5,4 m) para que o polígono criado por meio desse buffer fique com a largura total da máquina.
5. Com cada polígono criado e com o auxílio da calculadora de campo, calculou-se a área de cada um deles, que também equivale a cada módulo colhido.
6. Por fim, com as áreas de cada módulo geoespacializadas (Figura 7), por meio do comando “join”, uniu-se os resultados de qualidade da fibra média de cada módulo à cada polígono, possibilitando-nos o mapeamento dos parâmetros de qualidade da fibra.

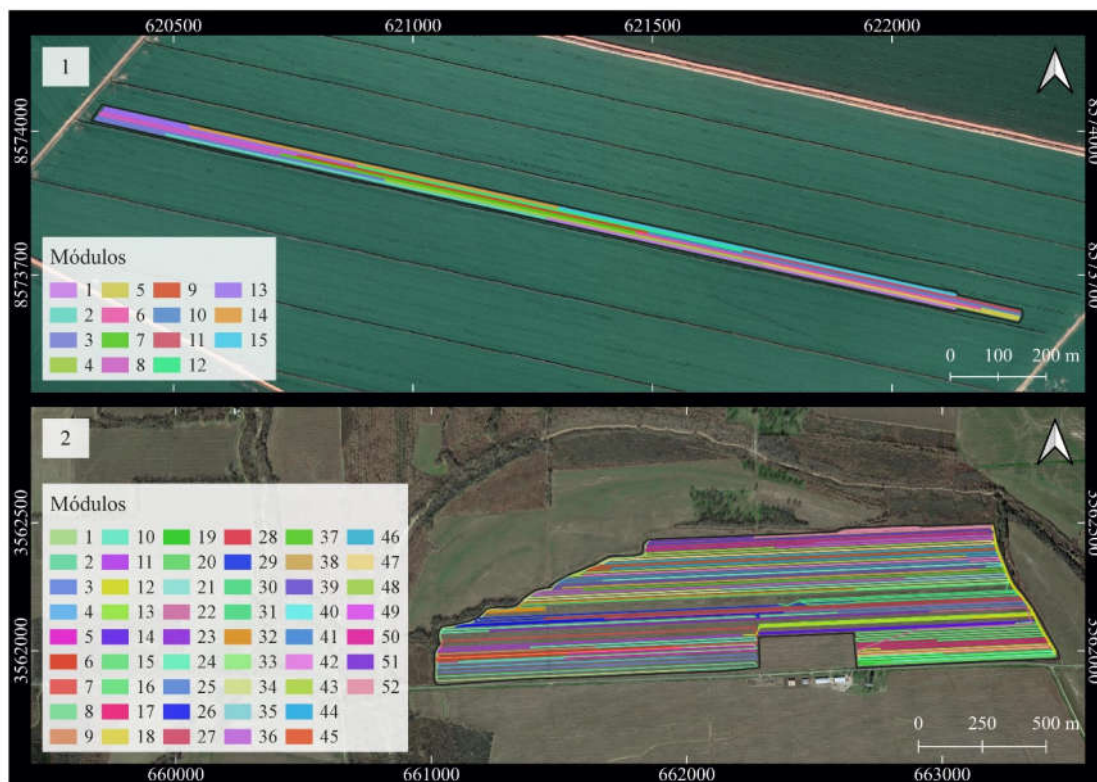


Figura 7. Mapeamento da área colhida de cada módulo nas etapas 1 (Sorriso, Brasil) e 2 (Newellton, EUA) do estudo.

Os parâmetros de classificação da qualidade da fibra mais importantes para comercialização como micronaire, resistência, comprimento, índice de uniformidade, índice de fibra curta, grau de cor e de folha foram associados aos polígonos de cada módulo e mapas temáticos foram gerados cada método de coleta nas etapas de estudo.

Para apresentação e discussão dos resultados das etapas 1 e 2 seguimos a classificação apresentada na Instrução Normativa 24/2016 (Regulamento Técnico da Fibra de Algodão no Brasil - Brasil, 2016) e a “Cotton Classification” (USDA, 2018).

Qualidade do processo de classificação de fibras

Todos os resultados foram analisados por meio da estatística descritiva, utilizando-se a média aritmética, desvio padrão e coeficiente de variação. Em seguida analisou-se os indicadores de qualidade (parâmetros de classificação da fibra) por

meio do controle estatístico de processo (CEP). Os indicadores de qualidade deste estudo foram os parâmetros micronaire, resistência, comprimento, uniformidade de comprimento, índice de fibras curtas e grau de cor e de folha.

Os resultados nas etapas 1 e 2 e seus respectivos métodos de coletas foram analisados por meio das cartas de controle de valores individuais que representam a média dos valores amostrados em cada ponto. Essas cartas apresentam três linhas, sendo que a linha central representa a média geral dos valores individuais, enquanto as outras duas linhas representam os limites superior e inferior de controle (LSC, LIC), respectivamente, calculados com base no desvio-padrão das variáveis (Montgomery, 2013).

Limites superior e inferior de especificação foram inseridos de acordo com os ágios e deságios (prêmios e descontos) segundo Bolsa Brasileira de Mercadorias (2021) para a etapa 1 (Brasil) e Premiums and Discounts for 2021- Crop American Upland Cotton (USDA, 2021b) para a etapa 2 (USA).

Resultados e Discussão

Produtividade do algodão

Observa-se na Figura 8 que a produtividade em nível de módulo pode ser mapeada e os valores podem ser graduados, observando-se que na etapa 1 (campo no Brasil) a produtividade variou de 4,72 a 5,70 ton ha⁻¹ e na etapa 2 (campo nos EUA) variou de 1,07 a 1,28 ton ha⁻¹.

Na etapa 1, a produtividade se apresentou acima da média do estado de Mato Grosso que, na safra 2020/2021, foi de 4,10 ton ha⁻¹ (CONAB, 2021), e segundo a fabricante da cultivar FM 985GLTP utilizada no estudo, a média na região de estudo é de 5,0 t ha⁻¹ (BASF, 2022). Os dados de produtividade da pesquisa demonstraram baixo coeficiente de variação (4,48 %) por ser área comercial (CONAB, 2021).

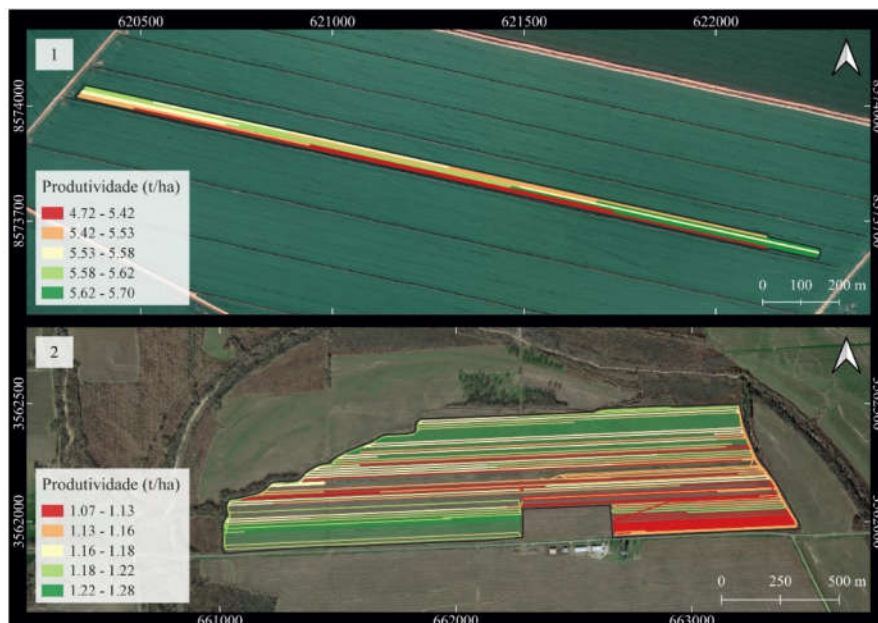


Figura 8. Produtividade média dos módulos colhidos nas etapas 1 (Sorriso, MT, Brasil) e 2 (Newellton, LA, EUA) da pesquisa.

Já o campo 2 apresentou média de $1,12 \text{ ton}^{-1}$, um pouco acima da média do estado da Louisiana e do país (EUA), que foi de $1,10$ e $0,94 \text{ ton ha}^{-1}$ respectivamente (USDA, 2021a).

Indicadores de qualidade da fibra do algodão

Os indicadores de qualidade da classificação HVI da fibra incluem parâmetros como micronaire, comprimento da fibra, uniformidade de comprimento, índices de fibras curtas e grau de cor e folha. Estes indicadores foram analisados por meio de cartas de controle que expressam os resultados graficamente de forma sequencial, permitindo verificar o comportamento médio, estabilidade e variabilidade ao longo da operação. As médias dos dados estão apresentadas na Tabela 1.

Tabela 1. Resultados médios dos parâmetros de qualidade da fibra por meio da análise HVI avaliados na comercialização do algodão em pluma em função dos métodos de coleta nas etapas 1 e 2 da pesquisa.

Métodos de coleta	Micronaire	UHML (mm)	Unf (%)	SFI (%)	Str (g/tex)
Etapa 1					
M1	4.38	28.79	79.34	14.49	28.22
M2	4.18	28.79	77.50	14.38	28.23
M3	4.20	28.70	77.39	10.89	27.80
Etapa 2					
M2	4.79	29.22	83.17	7.88	28.30

UHML: Comprimento Médio da Metade Superior; Unf: Índice de uniformidade; SFI: Índice de Fibra Curta; Str: Resistência;

Analisando a Figura 9A, observa-se que na etapa 1 (Brasil), a maior variabilidade nos resultados de micronaire para os 15 módulos avaliados foi encontrada para o método 2 (coleta do módulo em campo), pois apresentou maior distância entre os limites superior e inferior de controle, enquanto o M1 e M3 apresentaram as menores variabilidades no processo.

Nota-se também que, o M3 apresentou um ponto fora de controle (acima do LSC), indicando a ocorrência de causas especiais, sendo aquelas que ocorrem por alguma falha ou motivo facilmente identificado, como um erro sistemático por exemplo.

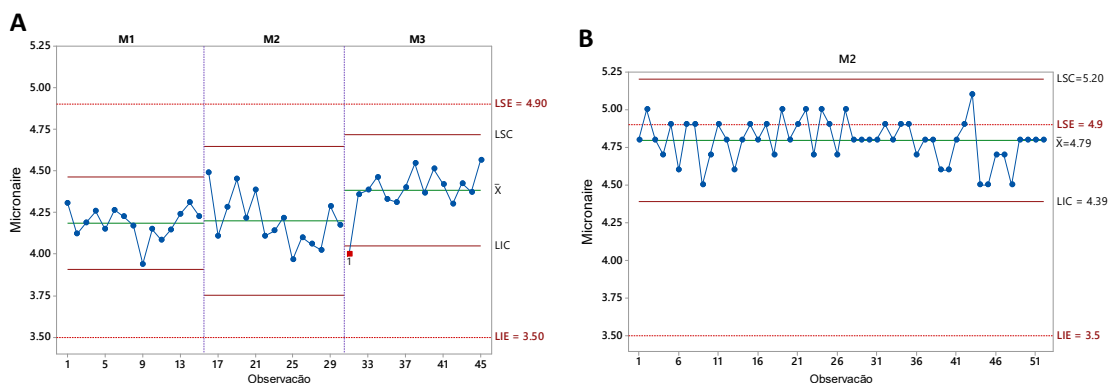


Figura 9. Cartas de controle de valores individuais para Micronaire nas etapas 1 (A) e 2 (B) da pesquisa.

Na etapa 2 (EUA), para os 52 módulos avaliados, a classificação do parâmetro micronaire se mostrou estável dentro dos limites superior e inferior de controle.

Segundo as Normas Específicas de Algodão (CONAB/MOC, 2022) e The Classification of Cotton (Cotton Incorporated, 2018), um valor de micronaire entre 3,5 e 4,9 é considerado um valor adequado. Quando se analisam os limites estabelecidos para ágios e deságios (3,5 a 4,9) observa-se que na etapa 1, nenhum resultado de micronaire ficou fora do padrão estabelecidos para comercialização, entretanto, na etapa 2, seis módulos originaram fibra que recebem desconto por apresentar micronaire acima do padrão exigido na maioria dos contratos.

Valores abaixo de 3,5 são indicativos de uma fibra imatura, enquanto valores acima de 4,9 são indicativos de fibras maduras, porém grossas demais (Gourlot et al., 2020), entretanto o desconto pode não acontecer caso o contrato seja firmado para entrega de valores de micronaire diferente desses limites citados.

A resistência é a força em gramas-força necessária para quebrar um 1 tex (1 grama por quilômetro) de fibras (CONAB/MOC 2022). Com base nessa informação, a resistência é reportada em gf/tex. Esse procedimento padroniza a informação relatada, permitindo a comparação entre diferentes amostras (Taylor, 1986).

Para os resultados de Resistência (Fig. 10), observa-se que na etapa 1, os métodos 1 e 2 ocasionam grande variabilidade no processo, e a menor variabilidade foi constatada no M3. Nota-se também que, o M3 apresentou um ponto fora de controle (acima do LSC), indicando a ocorrência de causas especiais. Na etapa 2, o parâmetro resistência se apresentou dentro dos limites de controle.

Tipicamente, a resistência possui um grande valor de desconto. No sistema de empréstimos americano, é possível que o produtor receba um pequeno ágio se estiver produzindo algodão com uma resistência maior que a média (26,60 – 28,90 gf/tex). No sistema de contratos brasileiro, só haverá descontos mesmo que ele produza fibras mais resistentes, e a faixa de intervalo varia sendo de 27,00 a 29,90 (CONAB/MOC, 2022).

Observa-se que no Brasil apenas um módulo ficou abaixo do mínimo exigido e sofreria deságio, e na etapa dos EUA, nota-se que 19 dos 52 módulos resultarão em fardos que receberão ágios/prêmios por resultarem em fibra acima do padrão.

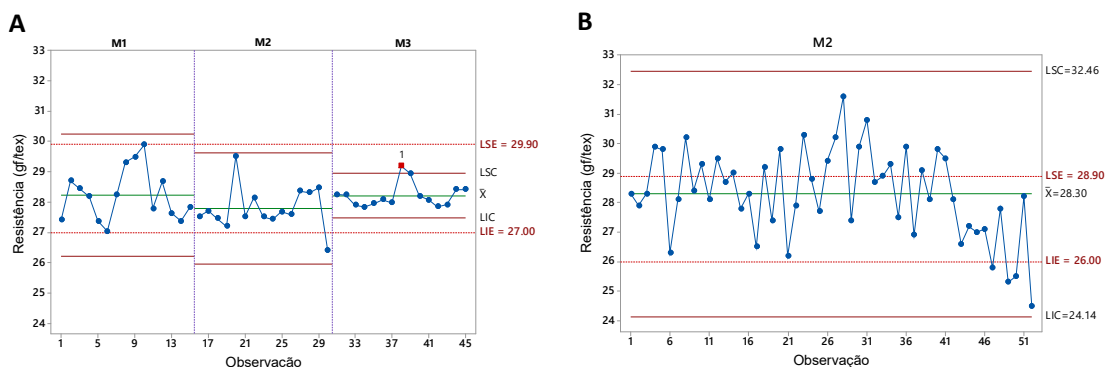


Figura 10. Cartas de controle de valores individuais para Resistência (Str em gf/tex) nas etapas 1 (A) e 2 (B) da pesquisa.

Analisando a Figura 11 observa-se que na etapa 1, a menor variabilidade no processo para comprimento médio da fibra foi encontrada para o método 3 (coleta do fardo), que apresentou menor distância entre os limites superior e inferior de controle, enquanto o M2 e M1 apresentaram as maiores variabilidades no processo. Nota-se também que, o M3 apresentou um ponto fora de controle (acima do LSC), indicando a ocorrência de causas especiais.

Quando se analisam os limites estabelecidos para ágios (acima de 28,00) e deságios (abaixo de 27,40), constata-se que em todos os métodos de coleta os módulos originaram fibras com comprimento acima do padrão, com exceção de apenas um ponto (obs. 26).

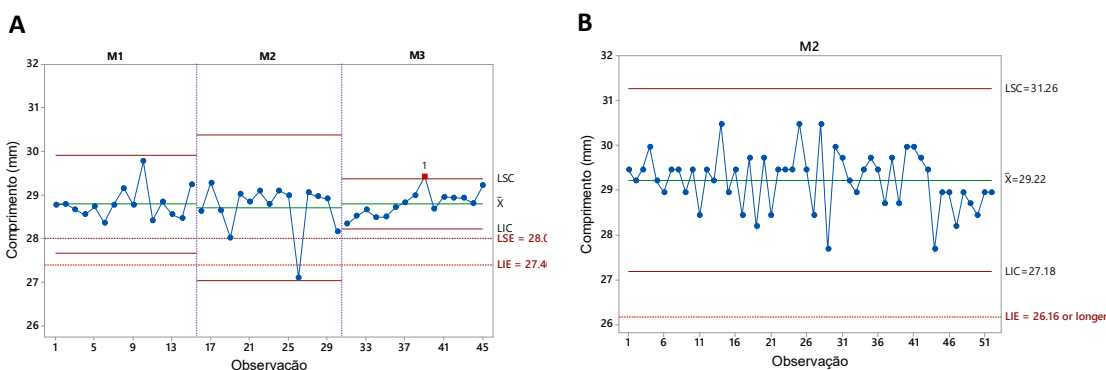


Figura 11. Carta de controle de valores individuais para comprimento médio (mm) nas etapas 1 (A) e 2 (B) da pesquisa.

Esse fato também se observa nas fibras da etapa 2 (EUA) em que todos os módulos resultaram em fibra de comprimento acima do padrão, mas não irão ganhar ágio por isso, apenas não sofrerão descontos.

O comprimento é expresso em polegadas, milímetros (1 polegada = 25,4 milímetros) ou comprimento manual (staple length, representando o numerador da fração de 32 avos de polegada). Segundo Mogahzy et al. (1990) essa informação é importante para os fiadores ajustarem as máquinas que serão usadas na produção de fios.

O comprimento (UHML) e a uniformidade (UI) são parâmetros importante para a qualidade da fiação anel e fiação a jato de ar, sendo comum nos contratos de diversos países produtores que haja prêmios e descontos para fardos de algodão baseados nos valores destes parâmetros (United States, 2020).

No Brasil, segue-se outros critérios, em que, o maior ágio e deságio é praticado sobre o comprimento relativo ao UHML. O UI não é considerado no sistema brasileiro, entretanto nos EUA sim, aplicando-se ágios e deságios para UI.

Na Figura 12 apresentamos o Índice de Uniformidade (%) em cada etapa de estudo. Para uniformidade de comprimento, observa-se que não houve pontos fora de controle nas etapas 1 e 2. Na Figura 12A observa-se que o processo apresentou menor variabilidade entre as amostras no método 3 (algodoeira), possivelmente devido ao tipo de descaroçamento, que foi industrial.

No Brasil, apesar de todas as amostras terem ficado abaixo do mínimo exigido nos EUA, não sofreu desconto devido à uniformidade de comprimento não entrar como parâmetro no sistema brasileiro.

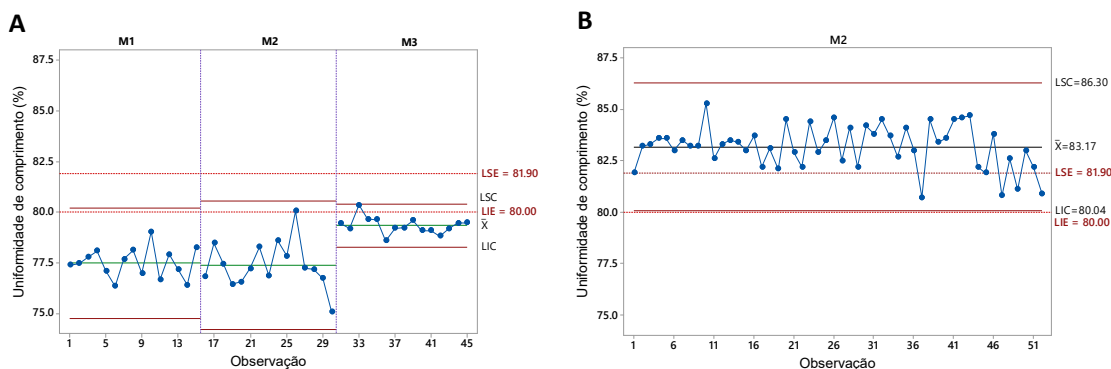


Figura 12. Carta de controle de valores individuais para Uniformidade de comprimento (%) nas etapas 1 (A) e 2 (B) da pesquisa.

Assim como no caso da uniformidade de comprimento, em que nos EUA serve como parâmetro de ágio e deságio, mas no Brasil não, acontece também com o índice de fibras curtas (SFI), que é usado para a aplicação de descontos no Brasil, porém nos EUA não.

Observa-se os resultados para índices de fibras curtas na Figura 13 e nota-se que houve grande variabilidade no processo nos métodos 1 e 2, enquanto no M3 a variabilidade foi menor demonstrando melhor qualidade no processo de descaroçamento, apesar de ter apresentado um ponto fora de controle.

Industrialmente, admite-se um índice de fibras curtas abaixo de 10%, pois acima disso a quantidade de fibras curtas se torna problemáticas para a fiação (Wakeham, 1955). Nota-se que em todos os métodos de coleta (com exceção da obs. 45), o indicador de qualidade SFI ficou acima do aceitável para a indústria podendo sofrer desconto na hora da venda. Entretanto, na etapa 2, observa-se que não houve pontos fora de controle e nenhum dos módulos que ultrapassaram o limite de fibras curtas.

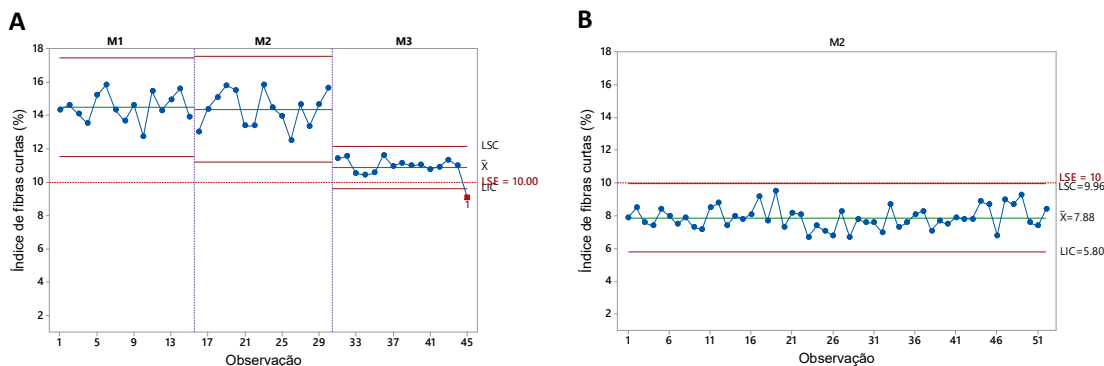


Figura 13. Carta de controle de valores individuais para Índices de Fibras Curtas (%). nas etapas 1 (A) e 2 (B) da pesquisa.

Por fim, o algodão tem um sistema de classificação que se refere ao grau de cor e de folha, comumente denominado de Color Grade (CG). Nesse sistema, os dois primeiros números se referem à cor, e o terceiro à carga de folha. Por exemplo, um algodão 31-4 é um algodão de cor 31, e de grau 4 em folha (Morais et al., 2021).

Na Figura 14, observa-se o indicador de qualidade Color Grade para etapa 1 da presente pesquisa e nota-se que a variabilidade no processo para os três métodos foi similar, e os valores variaram entre 31-1,31-2,31-3,31-4, 41-1, 41-2, 41-3 e 41-4.

Verificou-se que o M1 e M3 apresentaram pontos fora de controle (acima do LSC) quando resultaram em amostras de CG 41-1 (M1) e 41-4 (M3), indicando a ocorrência de causas especiais. O processo instável ou fora dos limites de controle pode ser analisado ou melhorado utilizando as premissas dos fatores 6 M's (máquina, mão-de-obra, medida, método, matéria-prima e meio).

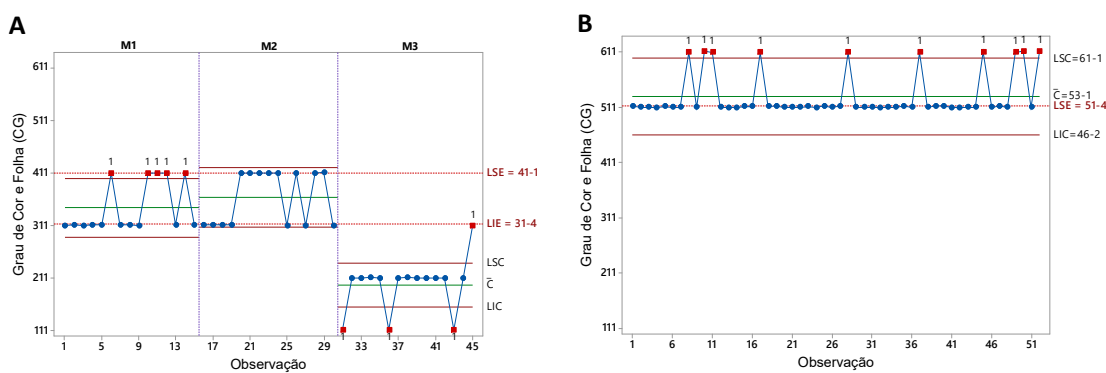


Figura 14. Carta de controle de atributos do tipo C para Grade de Cor (CG – Color Grade) nas etapas 1 (A) e 2 (B) da pesquisa.

No método 2, apesar de observarmos resultados 41-1 e 41-2, este método de coleta foi o único que apresentou todos os pontos dentro dos limites superior e inferior de controle, demonstrando melhor qualidade no processo de acordo com o conjunto de dados do tratamento M2.

Os resultados na etapa 2 (EUA) apresentaram dez pontos fora de controle indicando que o processo de classificação não foi estável, e os resultados de classificação CG variou entre 51-1, 51-2, 51-3, 51-4, 61-1 e 61-3, considerados de má qualidade, e serão explanados sobre na seção de mapeamento da qualidade da fibra nesta pesquisa.

Como o CG expressa o resultado do cruzamento dos valores do grau de refletância (%Rd) e do grau de amarelamento (+b) determinado no diagrama de cores de Nickerson/Hunter, os resultados foram altos devido à resultados indesejáveis tanto em folha (impurezas superficiais) como em amarelecimento e reflectância.

Diferentes mercados podem solicitar diferentes Grau de Cores (CGs). Por exemplo, o tipo base de algodão para o mercado interno brasileiro e americano é o 41-4. Enquanto isso, o tipo base para o mercado externo é o 31-4, ou seja, um algodão mais limpo (Bolsa Brasileira de Mercadorias, 2021), logo, pode-se dizer que os pontos fora de controle nesta pesquisa não estão aptos para exportação e sim apenas para o mercado interno.

Mapeamento da qualidade da fibra

Quanto ao mapeamento dos parâmetros de qualidade da fibra, antes da tecnologia RFID estar presentes nas colhedoras, e em colheita de algodão utilizando colhedoras de cesto, prensa e bass boy, duas pesquisas pontuais com uso de protótipos foram realizadas por Ge et al. (2008) no estado do Texas, USA, e Ge et al. (2012). Os autores, em propriedade agrícolas no Kansas, California e Texas, USA, construíram um sistema online com Wireless e GPS na tentativa de mapear os parâmetros de classificação da fibra de acordo com a área georreferenciada, módulo de algodão colhido e sua localização no talhão, possibilitando a geração de mapas de qualidade da fibra. Os autores observaram diferenças significativas na maioria das propriedades de fibra de algodão entre os módulos mapeados.

Em relação ao mapeamento da qualidade da fibra utilizando o RFID, estudos iniciados por Fuhrer et al. (2020) e Fuhrer (2022), ressalta que, com a implantação do sistema HID, os produtores podem ver espacialmente os dados de qualidade da fibra e que para a maioria das culturas o monitoramento da produtividade é suficiente para tomadas de decisões, mas no algodão, o produto colhido gera a fibra, que não é avaliada em campo.

Ward (2020) e Ward et al. (2022) tiveram os mesmos objetivos de mapear os parâmetros de qualidade da fibra em propriedades agrícolas em Carolina do Norte, EUA, e por meio deste gerar mapas de prêmios e desconto nos campos comerciais, e mapas de valores finais de venda dos módulos.

Visando alcançar o objetivo desta pesquisa em responder se é possível mapear os parâmetros de qualidade da fibra em campos comerciais, abaixo apresenta-se os mapas gerados (Figuras 15 a 20) nas etapas 1 (Brasil) e 2 (EUA) desta pesquisa, e

graduados de acordo com a Classificação apresentada na Instrução Normativa 24/2016 (Regulamento Técnico da Fibra de Algodão no Brasil - Brasil, 2016) e Classificação de Algodão dos EUA (USDA, 2021b).

Os dados de mapeamento apresentados da etapa 1 (Brasil) são inéditos, pois estes resultados são os primeiros a nível de pesquisa e em campos comerciais. Nos EUA trabalhos já vinham sendo realizados nessa temática desde 2020.

Para facilitar a interpretação dos resultados, na Tabela 2 apresenta-se a classificação dos parâmetros de qualidade da fibra mais importantes na comercialização de algodão produzido no Brasil e nos EUA, e que serviu de padronização para as legendas dos mapas.

Tabela 2. Classificação dos principais parâmetros de qualidade da fibra avaliados para comercialização em fibras produzidas no Brasil e nos EUA. Fonte: Brasil (2016) e USDA (2018).

Parâmetro	Classificação	Brasil (Brasil, 2016)	EUA (USDA, 2018)
Micronaire	Muito fina	≤ 3,0	≤ 3,0
	Fina	3,1 - 3,6	3,1 - 3,6
	Média	3,7 - 4,7	3,7 - 4,7
	Grossa	4,8 - 5,4	4,8 - 5,4
	Muito Grossa	≥ 5,5	≥ 5,5
Resistência (g/tex)	Fraca	≤ 24,9	≤ 23,4
	Intermediária	25,0 - 26,9	23,5 - 25,4
	Média	27,0 - 28,9	25,5 - 28,4
	Resistente	29,0 - 31,9	28,5 - 30,4
	Muito Resistente	≥ 32,0	≥ 30,5
Uniformidade de comprimento (%)	Muito baixa	≤ 77,0	≤ 76,5
	Baixa	77,0 - 79,0	76,5 - 79,4
	Regular	80,0 - 82,4	79,5 - 82,4
	Alta	82,5 - 85,5	82,5 - 85,4
	Muito alta	≥ 85,5	≥ 85,4
Índice de fibras curtas (%)	Muito baixa	≤ 6,0	-
	Baixa	6,0 a 7,9	-
	Regular	8,0 a 10,0	-
	Alta	10,1 a 13,0	-
	Muito Alta	≥ 13,0	-

Na Figura 15 apresenta-se os mapas de micronaire da fibra nas etapas 1 e 2 da presente pesquisa. A classificação do parâmetro micronaire é a mesma para o algodão produzido no Brasil e nos EUA, conforme consta na Tabela acima.

Nota-se que em todos os módulos avaliados na etapa 1, o micronaire apresentou fibras de classificação média (o que atende os padrões de comercialização sem sofrer ágio ou deságio), demonstrando uniformidade na área.



Figura 15. Mapa da classificação do parâmetro micronaire da fibra nas etapas 1 (Brasil) e 2 (EUA) da pesquisa.

Na etapa 2 houve variação de micronaire da fibra com classificação média e grossa, o que dependendo do contrato poderia haver ágios em alguns fardos comercializados.

Para o parâmetro Resistência, observa-se que as fibras variaram entre média, resistente (etapa 1 e 2) e muito resistente (etapa 2), atendendo aos parâmetros mínimos exigidos, e com recebimento de prêmios para os fardos com fibras acima da média.



Figura 16. Mapa da classificação do parâmetro resistência da fibra nas etapas 1 (Brasil) e 2 (EUA) da pesquisa.

Para o parâmetro comprimento da fibra (mm), observa-se na Figura 17 que nas duas etapas a classificação foi acima do mínimo exigido na comercialização, resultando em ágios para os produtores. O comprimento é um dos parâmetros mais importantes para a cadeia de valor do algodão em seus diferentes segmentos (Morais et al., 2021).



Figura 17. Mapa da classificação do parâmetro comprimento da fibra nas etapas 1 (Brasil) e 2 (EUA) da pesquisa.

No Brasil o comprimento é utilizado como base para ágio e deságio, e nos EUA não, entretanto, o parâmetro que eles utilizam para prêmios e descontos é a uniformidade do comprimento, conforme observado na Figura 18.

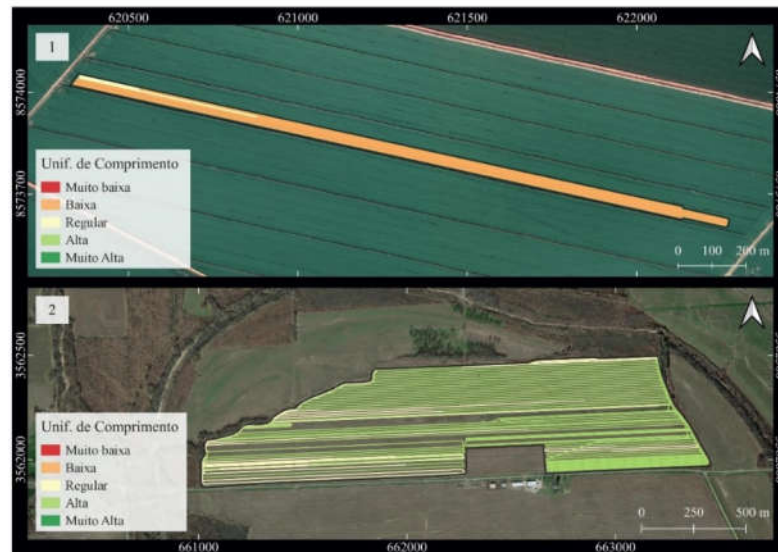


Figura 18. Mapa da classificação do parâmetro Uniformidade de comprimento nas etapas 1 (Brasil) e 2 (EUA) da pesquisa.

No Brasil, a uniformidade de comprimento foi baixa na maioria dos módulos e em dois deles como regular. No entanto, como citado anteriormente, a Uniformidade não é parâmetro base de ágio e deságio no Brasil, apenas nos EUA. Observa-se que na etapa 2 a uniformidade de comprimento da fibra variou entre regular e alta (maioria), estando acima do valor padrão, resultando em prêmio na comercialização dos fardos deste talhão.

O parâmetro índice de fibras curtas (FSI) é utilizado para ágio e deságio apenas no Brasil, não sendo levado em consideração para prêmio e desconto nos EUA. Observa-se na Figura 19 que o lote de fardos da área avaliada na etapa 1 sofreria deságio por não atender o máximo de 10% para FSI. Na etapa 2 apesar de não ter ágio ou prêmio, todo o talhão originou fibras com FSI aceitável.

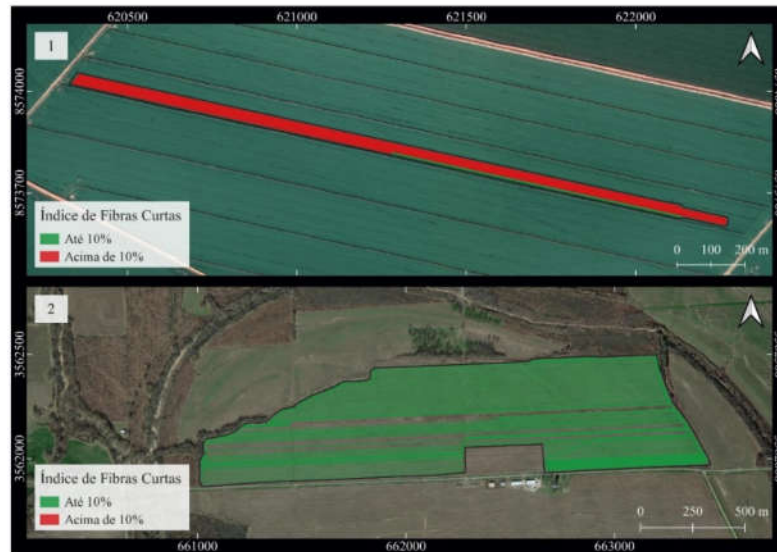


Figura 19. Mapa da classificação do parâmetro índices de fibras curtas nas etapas 1 (Brasil) e 2 (EUA) da pesquisa.

O Grau de cor e de folha (CG) é um parâmetro de classificação de algodão que se refere à cor e quantidade de impureza, e na etapa 2 (EUA) os resultados foram uma classificação insatisfatória, entretanto, esses resultados são explicados devido à forma de descaroçamento que as 52 amostras foram submetidas nesta etapa da pesquisa, que foi uma descaroçadora de mesa de 10 serras, que, infelizmente não descaroça e limpa a fibra com a qualidade de uma descaroçadora industrial, vindo a resultar nas fibras tipo 51-1 a 61-3 observadas nos resultados da etapa 2.

Diferentes mercados podem solicitar diferentes Grau de Cores (CGs). Por exemplo, o tipo base de algodão para o mercado interno brasileiro e americano é o 41-4. Enquanto isso, o tipo base para o mercado externo é o 31-4, ou seja, um algodão mais limpo (Bolsa Brasileira de Mercadorias, 2021), logo, pode-se dizer que os pontos fora de controle nesta pesquisa não estão aptos para exportação e sim apenas para o mercado interno.

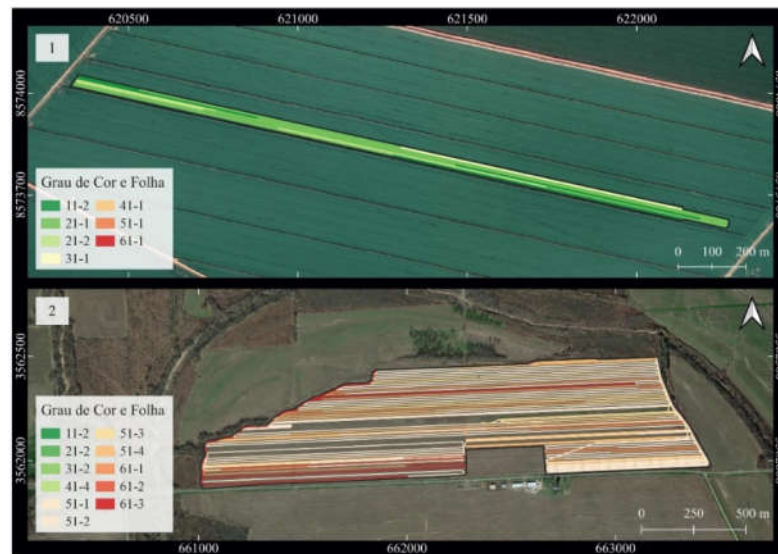


Figura 20. Mapa da classificação do parâmetro grau de cor e de folha nas etapas 1 e 2 da pesquisa.

Nesse sentido observa-se que o campo do Brasil resultou em fibras que atendem o mercado interno (41-4) e externo (31-4), porém no campo dos EUA todas as fibras não atendem nem o mercado interno devido a classificação 51-1 ou inferior, mas esse resultado é justificado devido ao método que essas fibras foram descaroçadas antes de serem classificadas. Como utilizou-se o descaroçador de 10 serras (utilizado em pesquisas e pequenas amostras) a limpeza não foi adequada e todas as amostras resultaram em uma fibra de qualidade inferior.

Os resultados de mapeamento dos parâmetros de qualidade da fibra aqui apresentados corroboram com as classificações mapeadas e apresentadas por Fuhrer (2022), Ward (2020) e Ward et al. (2022), comprovando que os mapas temáticos auxiliam na visualização da variabilidade nos resultados da classificação da fibra que ocorre em campos comerciais, se tornando importante aliado para os profissionais relacionados às tomadas de decisões.

Conclusões

É possível mapear os parâmetros de qualidade da fibra nas duas etapas da presente pesquisa desde que se vincule os dados de identificação do módulo com a etiqueta criada na algodoeira para os fardos beneficiados.

Esta técnica pode ser utilizada como ferramenta para os produtores visualizarem a variabilidade espacial da qualidade da fibra em seus campos comerciais, por meio de mapas, e tomarem decisões mais assertivas quanto ao manejo da lavoura.

A produtividade pode ser mapeada e classificada à nível de módulo.

Na etapa 1 (Brasil) o método de coleta do fardo (algodoeira) originou classificação de fibra com menor variabilidade no processo.

Os parâmetros de qualidade da fibra apresentam variabilidade em cada módulo e em cada método de coleta, na mesma área.

O processo de classificação ficou dentro dos limites superior e inferior de controle, apresentando variabilidade apenas entre os métodos de coleta, com exceção do parâmetro de classificação grau de cor e folha, que apresentaram diversos pontos fora de controle indicando que o processo não foi estável, coincidindo com a produção de fibras de padrão com qualidade inferior, fora da faixa de comercialização para o mercado internacional (31-4 ou inferior).

Referências

Ayele AG, Kelly BR, Hequet EF (2018) Evaluating Within-Plant Variability of Cotton Fiber Length and Maturity. **Agron. J.** 110:47–55. Doi:10.2134/agronj2017.06.0359

Bachelier B, Gourlot JP (2018) A fibra de algodão: origem, estrutura, composição e caracterização. In: AMPA & IMAMT. **Manual de qualidade da fibra da AMPA**. Ed. Jean-Louis Belot. Cuiabá. p.194-204. Disponível em: <http://www.ampa.com.br/arquivos/publicacoes/manual_qualidade_parte2.pdf>. Acesso em: 18 mar. 2019.

Barnes E (2020) Cotton Module Traceability. Disponível em: <<https://www.cottoninc.com/cotton-production/ag-resources/harvest-systems/rfid-tracking/>>. Acesso em: 01 ago. 2022.

Basf Sementes de algodão - FM 985GLTP (2022). Disponível em: <<https://agriculture.basf.com/br/pt/protecao-de-cultivos-e-sementes/produtos/fibermax/variedades-fibermax/FM-985GLTP.html>>. Acesso em 01 set. 2022.

Belot JL, Vilela PMC, Galbieri RDS (2017) Variedades comerciais e pré-comerciais de algodão para o Mato Grosso: síntese de resultados agronômicos e de qualidade de fibra obtidos pelo IMAmt até a safra 2015/16. IMAmt – Instituto Mato-grossense do Algodão.

Brasil, Ministério da Agricultura, Pecuária e Abastecimento (2016). Instrução Normativa nº 24, de 14 de julho de 2016. Regulamento Técnico do algodão em pluma. Disponível em: <http://www.inmetro.gov.br/barreirastecnicas/pontofocal/textos/regulamentos/BRA_683.PDF>.

Bolsa Brasileira de Mercadorias (2021) Tabela de ágios e deságios (mercado interno e externo) Disponível em: <https://www.bbmnet.com.br/upload/Tabela_de_Agios_e_desagios_19_10_2021.pdf>. Acesso em: 22 set. 2020 dão em Pluma. Diário Oficial da União, Brasília, 15 de julho de 2016.

Conab, Companhia Nacional de Abastecimento (2021) Acomp. safra bras. grãos, v. 7 - Safra 2020/21 - n. 12 – décimo segundo levantamento, Brasília. Disponível em: <<https://www.conab.gov.br/info-agro/safras/graos>>. Acesso em 10 jul 2022.

Conab/MOC – Companhia Nacional de Abastecimento - Manual de Operações (2022) Título 41 - Normas Específicas de Algodão – Safras 2021/2022 e 2022. Comunicado CONAB/MOC N.º 008 de 14/04/2022. Disponível em: <https://www.conab.gov.br/images/arquivos/moc/41_NORMAS_ESPECIFICAS_DE_ALGODAO_SAFRA_2021_2022_E_2022.pdf>. Acesso em 13 out. 2022.

Cotton Incorporated (2018) The Classification of Cotton. 32 p. Disponível em: <<https://www.cottoninc.com/wp-content/uploads/2017/02/Classification-of-Cotton.pdf>>. Acesso em 02 set. 2022.

Franco BC (2020) Tama e o algodão. Disponível em: <<https://soudealgodao.com.br/blog/apoiadores/tama/>>. Acesso em: 20 out. 2022.

Fuhrer L, Porter WM, Rains G, Snider JL, Barnes EM (2020) Mapping in-field cotton fiber quality utilizing John Deere Harvest Identification System. **Proceedings...** Beltwide Cotton Conferences, Austin, TX, January, 297-301.

Fuhrer L (2022) **Mapping of in-field cotton fiber quality utilizing John Deere's Harvest Identification System (HID)**. 73 f. Thesis (Master of Science). University of Georgia, Tifton.

Fryxell PA (1971). Phenetic analysis and the phylogeny of the diploid species of *Gossypium* L. (Malvaceae). **Evolution**, v. 25, n. 3, p. 554-562. DOI: 10.1111/j.1558-5646.1971.tb01916.x.

Ge Y, Thomasson JA, Sui R, Morgan CL, Searcy, WS, Parnell, CB (2008) Spatial variation of fiber quality and associated loan rate in a dryland cotton field. **Precision Agric.**, 9(4), 181–194. DOI: <https://doi.org/10.1007/s11119-008-9064-2>.

Ge Y, Thomasson JA, Sui, R. (2012) Wireless-and-GPS system for cotton fiber-quality mappings. **Precision Agric.** n. 13, p. 90-103. DOI: <https://doi.org/10.1007/s11119-011-9225-6>.

Gourlot JP, Drieling A, Qaud M, Gordon S, Knowlton J, Matusiak, M.; Sluijs, R. Van Der; Martin, V.; Froese, K.; Delhom, C (2020) Interpretation and use of instrument measured cotton characteristics - A guideline by ITMF International Committee on Cotton Testing Methods (ICCTM) and by ICAC task force on commercial standardization of instrument testing of cotton (CSITC). Washington, DC [s.n.]. Disponível em: www.csitc.org. Acesso em: 18 Jun. 2020.

Hardin IV RG, Barnes EM, Delhom CD, Wanjura, JD, Ward, JK (2022). Internet of things: Cotton harvesting and processing. *Computers and Electronics in Agriculture*, 202, 107294. DOI: <https://doi.org/10.1016/j.compag.2022.107294>

John Deere. Colhedora de Algodão CP690. Disponível em: <https://www.deere.com.br/pt/magazines/publication.html?id=4f26a16d#5>. Acesso em: 01 mar. 2020.

Johnson RM, Downer RG, MBradow JM, J Bauer, PJ, Sadler EJ (2002) Variability in Cotton Fiber Yield, Fiber Quality, and Soil Properties in a Southeastern Coastal Plain. **Agronomy Journal**, v. 94, p. 1305-1316. Doi: 10.2134 / agronj2002.1305

Kothari N, Deve J, Hague S, Hequet E (2015) Evaluating intraplant cotton fiber variability. *Crop Sci.* 55:564–570. Doi: 10.2135/cropsci2014.01.0077.

Kottek MJ, Grieser C, Beck B Rudolf, Rubel F(2006) World Map of Köppen Geiger Climate Classification updated. *Meteorol. Z.*, 15, 259-263.

Maldaner LMJ, Spekken M, Molin JP (2019) Software Map Filter 2. Piracicaba: USP Laboratório de Agricultura de Precisão. Disponível em: <<https://www.agriculturadeprecisao.org.br/software/>>. Acesso em: 28 out. 2022.

Montgomery DC (2013) *Statistical Quality Control*. 7th Edition. New York: John Wiley & Sons, Inc.

Morais JPS, Farias FJC, Belot J, Martins RSA, Mizoguchi ET (2021). Interpretação das características avaliadas no SITC para qualidade de fibra de algodão – uma abordagem prática. Campina Grande: Embrapa Algodão. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1134241/1/PUBLICA_R-AINFO-DOC285.pdf>. Acesso em 05 ago. 2022.

IoP Journal (2020) ID-Cotton guarantees cotton quality from the agriculture. Disponível em: <<https://iopjournal.com.br/id-cotton-guarantees-cotton-quality-from-the-agriculture/?lang=en>>. Acesso em 01 ago. 2022.

Porter WM et al. Tracking Cotton Fiber Quality and Foreign Matter through a Stripper Harvester. *The Journal of Cotton Science.* 21:29-39 (2017). Disponível em: <<https://www.cotton.org/journal/2017-21/1/upload/JCS21-029.pdf>>. Acesso em 02 mar. 2020.

Santos HG, JacominE PKT, Anjo LHC, Oliveira VA, Lumbreras JF, Coelho MR, Almeida JA, Araújo Filho, JC; Oliveira, JB, Cunha, TJF (2018) Sistema brasileiro de classificação de solos. Brasília, DF: Embrapa. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/199517/1/SiBCS-2018-ISBN-9788570358004.pdf>>. Acesso em: 10 maio. 2021.

Silva ORRF, Sofiatti V (2017) Colheita e Beneficiamento In Sistemas de Produção 2 ed. Embrapa Algodão. ISSN 1678-8710. Disponível em: <https://www.spo.cnptia.embrapa.br/conteudo?p_p_lifecycle=0&p_p_id=conteudopor_tlet_WAR_sistemasdeproducao1f6_1ga1ceportlet&p_p_col_count=1&p_p_col_id=column-1&p_p_state=normal&p_r_p_-76293187_sistemaProducaold=7718&p_r_p_-996514994_topicold=7993&p_p_mode=view>.

Silva RP, Corrêa CF, Cortez JW, Furlani CEA (2008). Statistical control applied to the mechanized harvesting process of sugarcane. **Engenharia Agrícola**, v. 28, p. 292-304.

Soil Survey Staff (2014) Keys to Soil Taxonomy (12th ed.) Washington DC: USDA-Natural Resources Conservation Service.

Stackhouse P. (2010) Prediction of worldwide energy resource. Hampton: NASA Langley Research Center. Disponível em: <<https://power.larc.nasa.gov/>>.

Taylor RA (1990) Estimating the size of cotton trash with video images. Textile Research Journal, v. 60, n. 4, p. 185-193, 1990. Doi: 10.1177/004051759006000401.

United States, Department of Agriculture (2020) Washington, DC, 1982. Disponível em: <https://naldc.nal.usda.gov/download/CAT87208951/PDF>. Acesso em: 17 jun. 2022

USDA - United States Department of Agriculture (2018) Cotton Classification: Understanding the Data. Disponível em: <<https://www.ams.usda.gov/sites/default/files/media/CottonDBUnderstandingtheData.pdf>>. Acesso em: 04 ago. 2022.

USDA - United States Department of Agriculture (2021a) National Agricultural Statistics Service. Crop Production – December. ISSN: 1936-3737. Disponível em: <https://www.nass.usda.gov/Publications/Todays_Reports/reports/crop1221.pdf>. Acesso em: 10 mar 2022.

USDA - United States Department of Agriculture (2021b) Premiums and Discounts for Grade, Staple Length, and Leaf Content of 2021 - Crop American Upland Cotton 1. 5 p. Disponível em: <<https://www.fsa.usda.gov/Assets/USDA-FSA-Public/usdfiles/Price-Support/pdf/2021/2021%20Cotton%20%20Premiums%20and%20Discounts.pdf>>.

USDA – United State Department of Agriculture (2020) World Agricultural Production. Circular Series WAP 2-20 February. Disponível em: <<https://apps.fas.usda.gov/psdonline/circulars/production.pdf>>. Acesso em: 22 fev. 2020

Wakeham H (1955) Cotton fiber length distribution: an important quality factor. Textile Research Journal, v. 25, n. 5, p. 422-429, 1955. DOI: 10.1177/004051755502500509.

Wanjura, JD, Barnes EM, Pelletier, MG, Holt GA (2017) New Technologies for Managing Cotton Modules. **Proceedings...** Beltwide Cotton Conference, Dallas, TX. Memphis, TN: National Cotton Council of America. 420-432.

Wanjura JD, Barnes EM, Holt GA, Pelletier MG (2018) New technologies for managing cotton modules and harvest information. **Proceedings...** Beltwide Cotton Conf. pp. 841-856. National Cotton Council of America, Memphis, TN.

Wanjura JD, Holt GA, Pelletier MG, Barnes EM (2020). Advances in managing cotton modules using RFID technology - system development update. **Proceedings...** Beltwide Cotton Conference National Cotton Council Beltwide Cotton Conference. 588-609.

Ward JK, Mitchell L, Barnes EM. Fiber Quality Mapping for Yield and Quality Interaction **Proceedings...** Beltwide Cotton Conference, San Antonio TX. Disponível em: <<https://ncc.confex.com/ncc/2022/meetingapp.cgi/Paper/21911>>. Acesso em 05 jul. 2022.

Ward JK (2020) Cotton Quality Mapping. Doi: <https://doi.org/10.1094/GROW-COT-12-20-275>.

CAPÍTULO 4 – Considerações finais

No capítulo 2 da presente tese, observou-se que os índices de vegetação avaliados em todas as datas de aquisição das imagens apresentaram correlação significativa e positiva. O destaque foi para o índice SR que foi o de maior correlação e o índice selecionado pela Stepwise Regression, em dois dos três campos de estudo como a variável de maior relevância para a predição da produtividade de algodão.

O cenário 1 (40 inputs) apresentou as melhores métricas em todos os campos de estudo e os algoritmos Multilayer Perceptron e Random Forest apresentaram os melhores desempenhos, ressaltando que neste cenário necessitamos de todos os inputs preditores que foram gerados ao longo de todo o ciclo da cultura (aprox. 180 dias).

De modo geral e para atingir o objetivo neste estudo, o KNN apresentou a melhor performance. A diferença encontrada nas métricas usando 1 input dentre 40 inputs na predição foi mínima por se tratar de área comercial com talhões de agricultura de larga escala no Mato Grosso, destacando que no cenário 2 reduzimos 39 variáveis de entrada e predizemos com mínima diferença no RMSE, possibilitando sua utilização como técnica rápida que auxiliará produtores e gerenciadores de dados na predição de rendimento, resultando em ganho de tempo no planejamento, seja da comercialização ou na intervenção no manejo e colheita do algodão.

Além do melhor desempenho apresentado pelo algoritmo KNN, outro aspecto importante observado no estudo foi que esse algoritmo exigiu menor demanda computacional nos ajustes dos hiperparâmetros e calibração dos algoritmos, quando comparado a MLP e RF por exemplo.

No capítulo 3, através do desenvolvimento da metodologia proposta, um processo de utilização dos arquivos do rastreador de colheita (HID cotton files) para vincular aos dados de qualidade de fibra por meio da trajetória de colheita da máquina mostraram com sucesso que o rastreamento e mapeamento da qualidade da fibra é possível, mostrando que essa pode ser uma ferramenta valiosa para produtores entenderem mais sobre sua colheita e a qualidade de seu produto, a fibra.

Esta técnica pode ser utilizada como ferramenta para os produtores visualizarem a variabilidade espacial da qualidade da fibra em seus campos

comerciais, por meio de mapas, e tomarem decisões mais assertivas quando ao manejo da lavoura.

Observamos que o processo de descaroçamento influencia fortemente a qualidade da fibra. Assim, na etapa 1 e 2, em que estes métodos de coleta foram manuais das colhedoras e fardos, as amostras tiveram que ser descaroçadas em uma máquina de 10 serras, não podendo ser na industrial da algodoeira devido à pequena quantidade, o que originou fibras de qualidade inferior quando comparada ao método 3 da algodoeira.

Os métodos coletados na etapa 1 nos permitiu escolher um método na etapa 2, portanto, o método escolhido foi a coleta do fardo ao invés da coleta na colhedora, pois não precisa estar com o operador durante a colheita e parar a máquina de subir e coletar a amostra do cesto da colhedora.

Após a avaliação geral das etapas 1 e 2, a presente pesquisa nos permitiu escolher um método, e recomendamos que o método 2 e 3 podem ser utilizados. O método 2 (coleta no módulo em campo) é manual e mais moroso, a nível de pesquisa.

Já o método 3, com as informações de RFID e classificação da fibra vinculados, torna possível o mapeamento completo em unidades de produção sem ter que ir a campo coletar amostras manuais, pois por meio do número ID do módulo e os resultados de classificação, consegue-se realizar o mapeamento da qualidade da fibra.

Caso essa vinculação não seja feita, nada pode ser feito em relação ao mapeamento, uma vez que, se tem as análises, porém não se tem a originação delas em relação aos módulos, para saber qual a localização deste. Essa realidade é mais comum que acontece na maioria das propriedades cotonicultoras e infelizmente foi o que ocorreu na etapa 2 (EUA) deste estudo.

O produtor nos informou que o número de identificação dos módulos estava relacionado com a algodoeira e que depois nos enviaria as informações para o mapeamento, mas infelizmente foi gerado novas etiquetas de identificação dos módulos pela algodoeira e sem vinculação aos dados de identificação dos módulos, o que nos impossibilitou de realizar o mapeamento pelo método 3 na etapa 2 desta pesquisa.

Para futuras pesquisas, os métodos atuais precisarão ser adaptados para tornar o processo de coleta de dados e criação de mapas mais automatizado e fácil de usar, pois até o momento no Brasil, a coleta ocorrendo no fardo em campo, tendo que ir até eles anotar número de identificação do módulo e coletar amostras de todas as partes do módulo, se torna inviável para todas as áreas comerciais da fazenda.

Para satisfazer plenamente o objetivo do terceiro capítulo, mais trabalhos devem ser feitos para adicionar as visualizações da análise econômica do lucro líquido e desenvolver para o usuário uma ferramenta ou interface amigável na geração dos mapas para facilitar as tomadas de decisões.

Será possível mapear os parâmetros de qualidade da fibra de forma completa em todos os talhões de uma fazenda desde que se vincule os dados de identificação do módulo com a etiqueta criada na algodoeira para os fardos beneficiados.

APÊNDICES

Apêndice A. Valores dos hiperparâmetros utilizados pelos algoritmos avaliados na predição de produtividade em cada cenário e campo de estudo do segundo capítulo desta tese.

	Hiperparâmetros	GridSearch
MLP	hidden_layer_sizes	3,5,8,10
	learning_rate	adaptive,constant
	activation	relu
	solver	Lbfgs, adam
	random_state	0,1,50,100
RF	max_depth	2,3,4,5,6
	n_estimators	1,2,3,5,7,9,10
	max_features	sqrt, auto
	min_samples_split	2,3,4,5,6
	min_samples_leaf	2,3,4,5,6
	bootstrap	True,False
KNN	leaf_size	range(1,10)
	n_neighbors	range(1,10)
	p=list	range(1,5)
	algorithm	auto
GLM	alpha	0, 0.2, 0.4, 0.6, 0.8, 1
AutoML	activation	RectifierWithDropout
	epochs	10000
	epsilon	1e-6, 1e-7, 1e-8, 1e-9
	hidden:	Grid1: 20,50,100 Grid2: {20, 20}, {50, 50}, {100, 100} Grid3: {20, 20, 20}, {50, 50, 50}, {100, 100, 100}
	Hidden dropout ratios	Grid1: {0.1}, {0.2}, {0.3}, {0.4}, {0.5} Grid2: {0.1, 0.1}, {0.2, 0.2}, {0.3, 0.3}, {0.4, 0.4}, {0.5, 0.5} Grid3: {0.1, 0.1, 0.1}, {0.2, 0.2, 0.2}, {0.3, 0.3, 0.3}, {0.4, 0.4, 0.4}, {0.5, 0.5, 0.5}
	input_dropout_ratio	0, 0.05, 0.1, 0.15, 0.2
	rho	{0.9, 0.95, 0.99}

RLM: Regressão Linear Múltipla, **MLP:** Multilayer Perceptron, **RF:** Random Forest Regressor, **KNN:** K-Nearest Neighbor Regressor; **AutoML:** Aprendizado de máquina automático.

Apêndice B. Tabela de ágios e deságios (R\$/kg) em função Da Classificação Universal para comercialização de fibras de algodão upland produzidas no Brasil. Fonte: CONAB/MOC (2022).

- 6) **PREÇOS MÍNIMOS:** Tomando por base o Preço Mínimo básico fixado pela Portaria MAPA N.º 201, de 01/07/2021, os preços foram estabelecidos em R\$/kg, conforme TABELAS I, II, III e IV:

a) algodão em pluma:

TABELA I – Algodão em Pluma – Branco (Safrá 2021/2022 e 2022)
Tabela de ágios e deságios (em R\$/kg)

GRADE	TIPO	FOLHA					
		1 & 2	3	4	5	6	7
GM - Good Middling	11	5,6390	5,6059	5,5728	n	n	n
SM - Strict Middling	21	5,6390	5,6059	5,5728	n	n	n
M - Middling	31	5,6059	5,5728	5,5398	5,5067	n	n
SLM - Strict Low Middling	41	5,5398	5,5287	5,5067	5,4516	5,4185	n
LM - Low Middling	51	5,5067	5,4516	5,4185	5,3854	5,3524	5,3193
SGO - Strict Good Ordinary	61	5,4185	5,3854	5,3524	5,3193	5,2862	5,2532

MICRONAIRE		RESISTÊNCIA		COMPRIMENTO DA FIBRA		
-	R\$/kg	gf/tex	R\$/kg	Polegadas	Cód. Universal	R\$/kg
5.0/5.2	(0,0772)	25.0/26.9	(0,0441)	1.1/8" ou acima	36 ou acima	0,0220
3.5/4.9 (*)	0,0000 (*)	27.0/29.9 (*)	0,0000 (*)	1.3/32" (*)	35 (*)	0,0000 (*)
3.3/3.4	(0,0661)	30.0/31.9 ou acima	0,0331	1.1/16"	34	(0,0772)

(*) Preço Mínimo Básico.

Obs.: Exemplo de cálculo de Preço Mínimo (R\$/Kg):
Classificação universal (obtida no certificado): 21337; micronaire: 3.39 e resistência: 26.1;
Preço Mínimo (R\$/kg): 5,6059 + 0,0220 - 0,0661 - 0,0441 = 5,5177 - Alíquota INSS vigente (Título 20 do MOC).

TABELA II – Algodão em Pluma – Ligeiramente Creme (Safrá 2021/2022 e 2022)
Tabela de ágios e deságios (em R\$/kg)

GRADE	TIPO	FOLHA					
		1 & 2	3	4	5	6	7
GM - Good Middling	12	5,6170	5,5839	5,5508	n	n	n
SM - Strict Middling	22	5,6170	5,5839	5,5508	n	n	n
M - Middling	32	5,5839	5,5508	5,5178	n	n	n
SLM - Strict Low Middling	42	5,5178	5,4847	5,4516	5,4186	5,3855	n
LM - Low Middling	52	5,4516	5,4186	5,3855	5,3524	5,3194	5,2863
SGO - Strict Good Ordinary	62	5,3855	5,3524	5,3194	5,2863	5,2532	5,2201

MICRONAIRE		RESISTÊNCIA		COMPRIMENTO DA FIBRA		
-	R\$/kg	gf/tex	R\$/kg	Polegadas	Cód. Universal	R\$/kg
5.0/5.2	(0,0551)	25.0/26.9	(0,0496)	1.1/8" ou acima	36 ou acima	0,0331
3.5/4.9 (*)	0,0000 (*)	27.0/29.9 (*)	0,0000 (*)	1.3/32" (*)	35 (*)	0,0000 (*)
3.3/3.4	(0,0220)	30.0/31.9 ou acima	0,0496	1.1/16"	34	(0,0661)

(*) Preço Mínimo Básico.

Obs.: Exemplo de cálculo de Preço Mínimo (R\$/Kg):
Classificação universal (obtida no certificado): 52435; micronaire: 5.1 e resistência: 25.5;
Preço Mínimo (R\$/kg): 5,3855 + 0,0000 - 0,0551 - 0,0496 = 5,2808 - Alíquota INSS vigente (Título 20 do MOC).

b) algodão em caroço: calculado com base no Preço Mínimo Básico, sendo que os tipos e classificação, baseiam-se no certificado de classificação;

b.1) finalidade: comprovação do pagamento do Preço Mínimo a produtores e suas cooperativas;

Apêndice C. Tabela de prêmios e descontos (pontos por libra) em função dos Padrões Universais dos Estados Unidos para comercialização de algodão upland em 2021. Fonte: USDA (2021).

2021-CROP UPLAND COTTON SCHEDULE OF PREMIUMS AND DISCOUNTS

Extraneous Matter 2021-Crop Upland Cotton		
	Level 1	Level 2
	Points of discounts	
Tex-NM-Oklahoma-KS Bark	-345	-525
Prep. All Locations	-205	-675
Other 1/	-445	-630
Plastic	-3130	-3240

1/ Bark in locations other than TX/NM/OK/KS. Extraneous matter other than bark and preparation, in all locations.

Fiber Strength 2021-Crop Upland Cotton	
Strength	Points
17.9 & lower	-600
18.0 - 18.9	-530
19.0 - 19.9	-530
20.0 - 20.9	-525
21.0 - 21.9	-520
22.0 - 22.9	-485
23.0 - 23.9	-470
24.0 - 24.9	-460
25.0 - 25.9	-410
26.0 - 26.9	0
27.0 - 27.9	0
28.0 - 28.9	0
29.0 - 29.9	5
30.0 - 30.9	20
31.0 - 32.9	40
33.0 & higher	45

Micronaire Differences 2021-Crop Upland Cotton	
Micronaire Reading	Points
2.4 & lower	-1905
2.5 through 2.6	-1415
2.7 through 2.9	-920
3.0 through 3.2	-665
3.3 through 3.4	-495
3.5 through 3.6	0
3.7 through 4.2 a/	10
4.3 through 4.9	0
5.0 through 5.2	-240
5.3 & higher	-390

a/ Premium applies only to white grades 11-41, leaf 1-6; 51, leaf 1-5; light spotted grades 12-32, leaf 1-5; 42, leaf 1-4; and 52, leaf 1-3.

Length Uniformity 2021-Crop Upland Cotton	
Uniformity	Points
77.9 & lower	-105
78.0 - 78.9	-65
79.0 - 79.9	-55
80.0 - 80.9	0
81.0 - 81.9	0
82.0 - 82.9	5
83.0 - 83.9	10
84.0 - 84.9	15
85.0 - 85.9	20
86.0 & higher	25