

UNIVERSIDADE FEDERAL DE VIÇOSA

**MODELAGEM DA INFESTAÇÃO DE LIANAS E DA BIOMASSA DO FUSTE DE
ÁRVORES EM FRAGMENTOS DE MATA ATLÂNTICA.**

Ivaldo da Silva Tavares Júnior
Doctor Scientiae

**VIÇOSA - MINAS GERAIS
2024**

IVALDO DA SILVA TAVARES JÚNIOR

**MODELAGEM DA INFESTAÇÃO DE LIANAS E DA BIOMASSA DO FUSTE DE
ÁRVORES EM FRAGMENTOS DE MATA ATLÂNTICA.**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Doctor Scientiae*.

Orientador: Jose Marinaldo Gleriani

Coorientador: Carlos M M Eleto Torres

**VIÇOSA - MINAS GERAIS
2024**

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

T

T231m
2024 Tavares Júnior, Ivaldo da Silva, 1991-
Modelagem da infestação de lianas e da biomassa do fuste
de árvores em fragmentos de Mata Atlântica / Ivaldo da Silva
Tavares Júnior. – Viçosa, MG, 2024.
1 tese eletrônica (133 f.): il. (algumas color.).

Inclui apêndices.

Orientador: José Marinaldo Gleriani.

Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Engenharia Florestal, 2024.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2024.713>

Modo de acesso: World Wide Web.

1. Plantas trepadeiras - Mata Atlântica. 2. Mapeamento
florestal - Mata Atlântica. 3. Florestas - Manejo. 4. Aprendizado
do computador. I. Gleriani, José Marinaldo, 1966-.
II. Universidade Federal de Viçosa. Departamento de Engenharia
Florestal. Programa de Pós-Graduação em Ciência Florestal.
III. Título.

GFDC adapt. CDD 634.9441

IVALDO DA SILVA TAVARES JÚNIOR

**MODELAGEM DA INFESTAÇÃO DE LIANAS E DA BIOMASSA DO FUSTE DE
ÁRVORES EM FRAGMENTOS DE MATA ATLÂNTICA.**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência Florestal, para obtenção do título de *Doctor Scientiae*.

APROVADA: 18 de julho de 2024.

Assentimento:

Ivaldo da Silva Tavares Júnior
Autor

Jose Marinaldo Gleriani
Orientador

Essa tese foi assinada digitalmente pelo autor em 21/11/2024 às 01:24:37 e pelo orientador em 25/11/2024 às 16:24:45. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **5P83.CBF7.NYMV** e clique no botão 'Validar documento'.

AGRADECIMENTOS

Primeiramente a Deus pelo dom da vida e por ter me dado saúde para superar as dificuldades até o presente momento.

A minha esposa Jianne Mazzini pelos incentivos, conselhos e carinho durante o doutorado e ao meu filho Matheus Tavares Mazzini que foi meu maior presente. Aos meus pais Maria Lucia Pereira Tavares e Ivaldo da Silva Tavares pelo amor, carinho, atenção e ensinamentos durante toda a vida.

Ao meu orientador Prof. José Marinaldo Gleriani e coorientador Prof. Carlos Moreira Miquelino Eleto Torres pela amizade, apoio, dedicação e orientações para realização deste trabalho.

Aos amigos: Jadson Abreu, Gabriela Costa, Fernanda Lambrecht, Lucas Lopes, Júlia Silveira, Gabriel Soares, Duberlí Gonzáles, Cátia Silva, Sindy Abranches, Epifanía Vega, Debora Christo, Gustavo Abreu, Bruna Guirardi, Thales Martins e Agust Sales, que foram minha grande família em Viçosa.

Ao Prof. Ângelo Augusto Ebling, Prof. Jonas Elias Castro da Rocha e Prof^a Gládis de Oliveira Jucoski pelo apoio fundamental no início desta caminhada e que não serão esquecidos.

Ao Wendelo Silva Costa pela amizade, apoio e acolhimento fundamental durante as primeiras semanas em Viçosa.

Ao amigo Samuel Rocha pela amizade e incentivo na busca de novas ferramentas para o manejo de florestas nativas.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão de bolsa para realização da pesquisa.

A todos os funcionários do departamento de Engenharia Florestal, em especial Alexandre, Dilson, Chiquinho da floresta e Elizabeth por ajudarem sempre que precisamos, e a Universidade Federal de Viçosa por proporcionar toda a estrutura necessária.

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

RESUMO

JÚNIOR, Ivaldo da Silva Tavares, D.Sc., Universidade Federal de Viçosa, julho de 2024. **MODELAGEM DA INFESTAÇÃO DE LIANAS E DA BIOMASSA DO FUSTE DE ÁRVORES EM FRAGMENTOS DE MATA ATLÂNTICA..** Orientador: Jose Marinaldo Gleriani. Coorientador: Carlos Moreira Miquelino Eleto Torres.

A Mata Atlântica, uma das florestas tropicais mais biodiversas do mundo, atualmente sofre com a degradação e fragmentação, restando apenas cerca de 12% de sua extensão original. A fragmentação florestal impacta negativamente a biomassa florestal, reduzindo a densidade e a diversidade de árvores e aumentando a abundância de lianas. Portanto, compreender e modelar a ocorrência e a abundância de lianas, assim como a biomassa das árvores, é essencial para a conservação e manejo sustentável da Mata Atlântica. Assim, os objetivos deste trabalho foram classificar a ocorrência de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando técnicas de aprendizagem de máquina (Capítulo I); projetar a abundância de árvores infestadas por lianas utilizando técnicas de aprendizagem de máquina (Capítulo II); e prever a biomassa do fuste de árvores integrando descritores texturais da matriz de coocorrência em níveis de cinza (GLCM) e aprendizagem de máquina, incluindo a abundância de árvores infestadas por lianas nos modelos (Capítulo III). Na modelagem da ocorrência de árvores infestadas por lianas, a análise de variáveis como altura total das árvores e diâmetro à 1,30 m de altura indicam que árvores maiores oferecem mais suporte para o crescimento das lianas. Na modelagem da abundância de árvores infestadas por lianas, a área basal por hectare, grupo ecológico, fósforo remanescente no solo e a concentração de alumínio trocável foram variáveis cruciais. A maior área basal sugere uma floresta mais madura, proporcionando o fechamento do dossel e redução das lianas. Já a alta concentração de alumínio trocável indica um ambiente competitivo mais favorável para as lianas devido à toxicidade para muitas plantas. Na predição da biomassa do fuste de árvores utilizando descritores texturais GLCM e aprendizagem de máquina utilizando dados do sensor Landsat 7, foram extraídos descritores texturais durante períodos secos e chuvosos. As técnicas de aprendizagem de máquina, como a regressão vetorial de suporte (RVS) e redes neurais artificiais (RNA), mostraram robustez na validação, especialmente a RVS

durante o período seco. A inclusão de variáveis texturais e a consideração da abundância de árvores infestadas por lianas melhoraram a precisão dos modelos, destacando a importância de adaptar os modelos às condições sazonais e estruturais da vegetação. Esses resultados ressaltam a eficácia das técnicas de aprendizagem de máquina na modelagem da infestação de lianas, bem como na predição da biomassa do fuste de árvores. Além disso a inclusão de fatores ambientais e antrópicos proporcionam uma compreensão detalhada dos fragmentos florestais, promovendo um manejo sustentável da Mata Atlântica.

Palavras-chave: distúrbios em florestas; infestação; manejo florestal

ABSTRACT

JÚNIOR, Ivaldo da Silva Tavares, D.Sc., Universidade Federal de Viçosa, July, 2024. **MODELING LIANA INFESTATION AND TREE STEM BIOMASS IN ATLANTIC FOREST FRAGMENTS..** Adviser: Jose Marinaldo Gleriani. Co-adviser: Carlos Moreira Miquelino Eleto Torres.

The Atlantic Forest, one of the world's most biodiverse tropical forests, currently faces severe degradation and fragmentation, with only about 12% of its original extent remaining. Forest fragmentation negatively impacts forest biomass, reducing tree density and diversity while increasing liana abundance. Therefore, understanding and modeling liana occurrence and abundance, as well as tree biomass, are essential for the conservation and sustainable management of the Atlantic Forest. The objectives of this study were to classify the occurrence of liana-infested trees in Atlantic Forest fragments using machine learning techniques (Chapter I); project the abundance of liana-infested trees using machine learning techniques (Chapter II); and predict tree stem biomass by integrating texture descriptors from the Gray-Level Co-occurrence Matrix (GLCM) and machine learning, including liana-infested tree abundance in the models (Chapter III). In modeling the occurrence of liana-infested trees, the analysis of variables such as total tree height and diameter at breast height (DBH, 1.3 m above the ground) indicated that larger trees provide more support for liana growth. In modeling the abundance of liana-infested trees, key variables included basal area per hectare, ecological group, residual phosphorus in the soil, and exchangeable aluminum concentration. Higher basal area suggests a more mature forest, contributing to canopy closure and reducing liana growth. On the other hand, high exchangeable aluminum concentrations indicate a competitive environment more favorable for lianas due to its toxicity to many plant species. In predicting tree stem biomass using GLCM texture descriptors and machine learning, based on data from the Landsat 7 sensor, texture descriptors were extracted during dry and wet seasons. Machine learning techniques, such as Support Vector Regression (SVR) and Artificial Neural Networks (ANN), demonstrated robustness during validation, particularly SVR during the dry period. The inclusion of texture variables and the consideration of liana-infested tree abundance improved model accuracy, highlighting the importance of adapting models to the seasonal and structural conditions of vegetation. These results underscore

the effectiveness of machine learning techniques in modeling liana infestation as well as in predicting tree stem biomass. Additionally, incorporating environmental and anthropogenic factors provided a detailed understanding of forest fragments, promoting the sustainable management of the Atlantic Forest.

Keywords: forest disturbances; infestation; forest management

SUMÁRIO

INTRODUÇÃO GERAL	9
Capítulo I: Modelagem da ocorrência de infestação de lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina	15
RESUMO.....	15
1 INTRODUÇÃO	16
2 MATERIAL E MÉTODOS	17
3 RESULTADOS.....	27
4 DISCUSSÃO	38
5 CONCLUSÃO	41
Capítulo II: Modelagem da abundância de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina	49
RESUMO.....	49
1 INTRODUÇÃO	50
2 MATERIAL E MÉTODOS	51
3 RESULTADOS.....	58
4 DISCUSSÃO	68
5 CONCLUSÃO	71
Capítulo III: Modelagem da biomassa do fuste de árvores em fragmentos de Mata Atlântica integrando descritores texturais GLCM e aprendizagem de máquina.....	76
RESUMO.....	76
1 INTRODUÇÃO	77
2 MATERIAL E MÉTODOS	79
3 RESULTADOS.....	86
4 DISCUSSÃO	97
5 CONCLUSÃO	102
CONCLUSÕES GERAIS	110
APÊNDICE A	111
APÊNDICE B	121
APÊNDICE C	131

INTRODUÇÃO GERAL

A dinâmica florestal é resultado da interação entre os processos demográficos das árvores, fatores ambientais (como clima, solo e topografia), distúrbios temporários de diversas escalas e origens (como queda de árvores, incêndios e exploração madeireira) e interações bióticas (ZHOU et al., 2021; STURTEVANT; FORTIN, 2021). A relevância desses fatores na dinâmica das florestas pode variar conforme os contextos ecológicos locais e regionais, os regimes de distúrbios e as respostas dos ecossistemas às mudanças nesses fatores (WORTHY; SWENSON, 2019).

Nesse contexto, a dinâmica florestal da Mata Atlântica é fortemente afetada por distúrbios tanto naturais, como antropogênicos (GREIG; ROBERTSON; LACERDA, 2018; GRELLE; RAJÃO; MARQUES, 2021). Atualmente, após sucessivos ciclos econômicos na região, restam aproximadamente 12% de sua extensão original (DIAS; SILVEIRA; FRANCISCO, 2023). A cobertura remanescente da Mata Atlântica consiste principalmente em pequenos fragmentos de floresta secundária com baixa conectividade, vulneráveis aos distúrbios naturais e antrópicos, e aos efeitos de borda (DINIZ et al., 2024; NASCIMENTO et al., 2014; LIRA; PORTELA; TAMBOSI, 2021). Uma das consequências desses distúrbios é o aumento da proliferação de lianas, que por sua vez inibem a regeneração das árvores (CAMPBELL; LAURANCE; MAGRACH, 2015; MARSHALL et al., 2020). As lianas desempenham um papel importante na dinâmica das florestas tropicais, muitas vezes atuando como colonizadoras precoces de locais perturbados (BECKNELL et al., 2022; PARTHASARATHY et al., 2015).

Lianas são formas de vida vegetal que dependem de suporte para alcançar o dossel da floresta, sendo denominadas parasitas estruturais (DURÁN; SÁNCHEZ-AZOFEIFA, 2015; EWERS; ROSELL; OLSON, 2015; VARGAS; GROMBONE-GUARATINI; MORELLATO, 2021). As lianas ocorrem principalmente nas florestas tropicais, podendo representar aproximadamente 25% da densidade de caules lenhosos, 35% das espécies lenhosas e 3% da área basal lenhosa (LIU et al., 2021; PANDI, 2023; SCHNITZER et al., 2012; YANG et al., 2018).

As lianas competem intensamente por luz, água e nutrientes (RODRÍGUEZ-RONDEROS et al., 2016), podendo prejudicar as espécies arbóreas devido à redução na área foliar e mudanças na estrutura da copa das árvores nos estratos

médio e superior da floresta (RODRÍGUEZ-RONDEROS et al., 2016); diminuem a reprodução das copas das árvores (GARCÍA LEÓN et al., 2018); reduzem o crescimento radial e as taxas de sequestro de carbono nas espécies colonizadas, modificando a anatomia do xilema e a densidade da madeira nas árvores hospedeiras (VENEGAS-GONZÁLEZ et al., 2020); diminuem o crescimento em biomassa da floresta (LAI et al., 2017); e reduzem a velocidade de transporte de seiva nas árvores hospedeiras (TOBIN et al., 2012).

Nesse contexto, a previsão da infestação de lianas e da biomassa de árvores sob a influência dessas plantas torna-se de grande importância para auxiliar no manejo das florestas e reduzir os impactos negativos na biomassa dos indivíduos arbóreos (KAUFFMAN et al., 2009; GUSTAFSON et al., 2010; POWELL et al., 2010). Ainda, a introdução de variáveis ambientais nos modelos para previsão da infestação das lianas pode aumentar a acurácia das previsões (GUARIGUATA; OSTERTAG, 2001; LAI et al., 2017; VAN DER SANDE et al., 2017; BERINI et al., 2018). Contudo, a modelagem da infestação de lianas e da influência na biomassa é desafiadora devido aos diversos fatores ambientais e antrópicos que atuam no ambiente e que podem aumentar as incertezas das previsões.

Desse modo, o objetivo deste trabalho foi modelar a ocorrência e abundância de árvores infestadas por lianas em fragmentos de Mata Atlântica, bem como modelar a biomassa do fuste de árvores utilizando técnicas de aprendizado de máquina. Para atender a este objetivo, o estudo foi dividido em três capítulos, conforme descrito a seguir:

Capítulo I: Modelagem da ocorrência de infestação de lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina;

Capítulo II: Modelagem da abundância de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina;

Capítulo III: Modelagem da biomassa do fuste de árvores em fragmentos de Mata Atlântica integrando descritores texturais GLCM e aprendizagem de máquina.

REFERÊNCIAS

BECKNELL, J. M.; VARGAS G., G.; WRIGHT, L. A.; WOODS, N.-F.; MEDVIGY, D.; POWERS, J. S. Increasing Liana Abundance and Associated Reductions in Tree Growth in Secondary Seasonally Dry Tropical Forest. **Frontiers in Forests and Global Change**, v. 5, p. 838357, 18 abr. 2022.

CAMPBELL, M.; LAURANCE, W. F.; MAGRACH, A. Ecological effects of lianas in fragmented forests. *Em*: SCHNITZER, S. A.; BONGERS, F.; BURNHAM, R. J.; PUTZ, F. E. **Ecology of Lianas**. 1. ed. [s.l.] Wiley, 2015. p. 443–450.

CASTAGNERI, D.; GARBARINO, M.; NOLA, P. Host Preference and Growth Patterns of Ivy (*Hedera Helix* L.) in a Temperate Alluvial Forest. **Plant Ecology**, v. 214, n. 1, p. 1–9, 1 jan. 2013.

CORNEJO, N. S.; BECKER, J. N.; HEMP, A.; HERTEL, D. Effects of Land-Use Change and Disturbance on the Fine Root Biomass, Dynamics, Morphology, and Related C and N Fluxes to the Soil of Forest Ecosystems at Different Elevations at Mt. Kilimanjaro (Tanzania). **Oecologia**, v. 201, n. 4, p. 1089–1107, 1 abr. 2023.

DIAS, T. da C.; SILVEIRA, L. F.; FRANCISCO, M. R. Spatiotemporal Dynamics Reveals Forest Rejuvenation, Fragmentation, and Edge Effects in an Atlantic Forest Hotspot, the Pernambuco Endemism Center, Northeastern Brazil. **PLOS ONE**, v. 18, n. 9, p. e0291234, 8 set. 2023.

DINIZ, É. S.; MOTA, P. H. S.; REIS, J. P.; DA SILVA COSTA, W.; DE PAIVA, E. V.; DE LANA, J. M.; LAGE, G. B.; DO AMARAL, C. H. Connectivity Value of Atlantic Forest Fragments: Pathways towards Enhancing Biodiversity Conservation. **Brazilian Journal of Botany**, v. 47, n. 1, p. 249–259, 1 mar. 2024.

DO NASCIMENTO, L. M.; DE SÁ BARRETTO SAMPAIO, E. V.; RODAL, M. J. N.; LINS-E-SILVA, A. C. B. Secondary Succession in a Fragmented Atlantic Forest Landscape: Evidence of Structural and Diversity Convergence along a Chronosequence. **Journal of Forest Research**, v. 19, n. 6, p. 501–513, 1 dez. 2014.

DURÁN, S. M.; SÁNCHEZ-AZOFEIFA, G. A. Liana Effects on Carbon Storage and Uptake in Mature and Secondary Tropical Forests. *Em*: PARTHASARATHY, N. **Biodiversity of Lianas**. Cham: Springer International Publishing, 2015. p. 43–55.

EWERS, F. W.; ROSELL, J. A.; OLSON, M. E. Lianas as Structural Parasites. *Em*: HACKE, U. **Functional and Ecological Xylem Anatomy**. Cham: Springer International Publishing, 2015. p. 163–188.

GALETTI, M.; GONÇALVES, F.; VILLAR, N.; ZIPPARRO, V. B.; PAZ, C.; MENDES, C.; LAUTENSCHLAGER, L.; SOUZA, Y.; AKKAWI, P.; PEDROSA, F.; BULASCOSCHI, L.; BELLO, C.; SEVÁ, A. P.; SALES, L.; GENES, L.; ABRA, F.; BOVENDORP, R. S. Causes and Consequences of Large-Scale Defaunation in the Atlantic Forest. *Em*: MARQUES, M. C. M.; GRELE, C. E. V. **The Atlantic Forest: History, Biodiversity, Threats and Opportunities of the Mega-diverse Forest**. Cham: Springer International Publishing, 2021. p. 297–324.

GARCÍA LEÓN, M. M.; MARTÍNEZ IZQUIERDO, L.; MELLO, F. N. A.; POWERS, J. S.; SCHNITZER, S. A. Lianas Reduce Community-Level Canopy Tree Reproduction in a Panamanian Forest. **Journal of Ecology**, v. 106, n. 2, p. 737–745, 2018.

GREIG, C.; ROBERTSON, C.; LACERDA, A. E. B. Spectral-temporal modelling of bamboo-dominated forest succession in the Atlantic Forest of Southern Brazil. **Ecological Modelling**, v. 384, p. 316–332, 24 set. 2018.

GRELLE, C. E. V.; RAJÃO, H.; MARQUES, M. C. M. The Future of the Brazilian Atlantic Forest. *Em*: MARQUES, M. C. M.; GRELLE, C. E. V. **The Atlantic Forest: History, Biodiversity, Threats and Opportunities of the Mega-diverse Forest**. Cham: Springer International Publishing, 2021. p. 487–503.

JONES, I. L.; PERES, C. A.; BENCHIMOL, M.; BUNNEFELD, L.; DENT, D. H. Woody Lianas Increase in Dominance and Maintain Compositional Integrity across an Amazonian Dam-Induced Fragmented Landscape. **PLOS ONE**, v. 12, n. 10, p. e0185527, 17 out. 2017.

LAI, H. R.; HALL, J. S.; TURNER, B. L.; VAN BREUGEL, M. Liana effects on biomass dynamics strengthen during secondary forest succession. **Ecology**, v. 98, n. 4, p. 1062–1070, 1 abr. 2017.

LIRA, P. K.; PORTELA, R. de C. Q.; TAMBOSI, L. R. Land-Cover Changes and an Uncertain Future: Will the Brazilian Atlantic Forest Lose the Chance to Become a Hopespot? *Em*: MARQUES, M. C. M.; GRELLE, C. E. V. **The Atlantic Forest: History, Biodiversity, Threats and Opportunities of the Mega-diverse Forest**. Cham: Springer International Publishing, 2021. p. 233–251.

LIU, Q.; STERCK, F. J.; ZHANG, J.-L.; SCHEIRE, A.; KONINGS, E.; CAO, M.; SHA, L.-Q.; POORTER, L. Traits, Strategies, and Niches of Liana Species in a Tropical Seasonal Rainforest. **Oecologia**, v. 196, n. 2, p. 499–514, 1 jun. 2021.

MARQUES, M. C. M.; TRINDADE, W.; BOHN, A.; GRELLE, C. E. V. The Atlantic Forest: An Introduction to the Megadiverse Forest of South America. *Em*: MARQUES, M. C. M.; GRELLE, C. E. V. **The Atlantic Forest: History, Biodiversity, Threats and Opportunities of the Mega-diverse Forest**. Cham: Springer International Publishing, 2021. p. 3–23.

MARSHALL, A. R.; PLATTS, P. J.; CHAZDON, R. L.; SEKI, H.; CAMPBELL, M. J.; PHILLIPS, O. L.; GEREAU, R. E.; MARCHANT, R.; LIANG, J.; HERBOHN, J.; MALHI, Y.; PFEIFER, M. Conceptualising the Global Forest Response to Liana Proliferation. **Frontiers in Forests and Global Change**, v. 3, p. 35, 24 abr. 2020.

PANDI, V. Ecology of Lianas: Diversity and Distribution. *Em*: PANDI, V. **Taxonomy and Ecology of Climbers: Climbing Plants of India**. Singapore: Springer Nature, 2023. p. 57–70.

PARTHASARATHY, N.; VIVEK, P.; MUTHUMPERUMAL, C.; MUTHURAMKUMAR, S.; AYYAPPAN, N. Biodiversity of Lianas and Their Functional Traits in Tropical Forests of Peninsular India. *Em*: PARTHASARATHY, N. **Biodiversity of Lianas**. Cham: Springer International Publishing, 2015. p. 123–148.

RODRÍGUEZ-RONDEROS, M. E.; BOHRER, G.; SANCHEZ-AZOFEIFA, A.; POWERS, J. S.; SCHNITZER, S. A. Contribution of Lianas to Plant Area Index and Canopy Structure in a Panamanian Forest. **Ecology**, v. 97, n. 12, p. 3271–3277, 2016.

SCHNITZER, S. A. A Mechanistic Explanation for Global Patterns of Liana Abundance and Distribution. **The American Naturalist**, v. 166, n. 2, p. 262–276, ago. 2005.

SCHNITZER, S. A.; MANGAN, S. A.; DALLING, J. W.; BALDECK, C. A.; HUBBELL, S. P.; LEDO, A.; MULLER-LANDAU, H.; TOBIN, M. F.; AGUILAR, S.; BRASSFIELD, D.; HERNANDEZ, A.; LAO, S.; PEREZ, R.; VALDES, O.; YORKE, S. R. Liana Abundance, Diversity, and Distribution on Barro Colorado Island, Panama. **PLOS ONE**, v. 7, n. 12, p. e52114, 21 dez. 2012.

SINGH, M.; HUANG, Z. Analysis of Forest Fire Dynamics, Distribution and Main Drivers in the Atlantic Forest. **Sustainability**, v. 14, n. 2, p. 992, jan. 2022.

STURTEVANT, B. R.; FORTIN, M.-J. Understanding and Modeling Forest Disturbance Interactions at the Landscape Level. **Frontiers in Ecology and Evolution**, v. 9, p. 653647, 4 out. 2021.

TOBIN, M. F.; WRIGHT, A. J.; MANGAN, S. A.; SCHNITZER, S. A. Ind. have a greater competitive effect than trees of similar biomass on tropical canopy trees. **Ecosphere**, v. 3, n. 2, p. art20, 1 fev. 2012.

TOLEDO-ACEVES, T.; SWAINE, M. D. Above- and below-Ground Competition between the Liana *Acacia kamerunensis* and Tree Seedlings in Contrasting Light Environments. **Plant Ecology**, v. 196, n. 2, p. 233–244, 1 jun. 2008.

USCANGA, A.; BARTLEIN, P. J.; SILVA, L. C. R. Local and Regional Effects of Land-Use Intensity on Aboveground Biomass and Tree Diversity in Tropical Montane Cloud Forests. **Ecosystems**, v. 26, n. 8, p. 1734–1752, 1 dez. 2023.

VARGAS, B. da C.; GROMBONE-GUARATINI, M. T.; MORELLATO, L. P. C. Lianas Research in the Neotropics: Overview, Interaction with Trees, and Future Perspectives. **Trees**, v. 35, n. 2, p. 333–345, abr. 2021.

VENEGAS-GONZÁLEZ, A.; MELLO, F. N. A.; SCHNITZER, S. A.; CÉSAR, R. G.; TOMAZELLO-FILHO, M. The Negative Effect of Lianas on Tree Growth Varies with Tree Species and Season. **Biotropica**, v. 52, n. 5, p. 836–844, 2020.

VILLAGRA, M.; TRENTINI, C. P.; DI FRANCESCANTONIO, D.; ELEUTERIO, A. A.; FOLETTO, F.; MONTTI, L. F.; CAMPANELLO, P. I. Disturbance as a Driver of Trait Assembly in Liana Communities in a Semi-Deciduous Atlantic Forest. **Plant Ecology**, v. 222, n. 7, p. 773–790, 1 jul. 2021.

WORTHY, S. J.; SWENSON, N. G. Functional Perspectives on Tropical Tree Demography and Forest Dynamics. **Ecological Processes**, v. 8, n. 1, p. 1, dez. 2019.

YANG, S.-Z.; FAN, H.; LI, K.-W.; KO, T.-Y. How the Diversity, Abundance, Size and Climbing Mechanisms of Woody Lianas Are Related to Biotic and Abiotic Factors in a Subtropical Secondary Forest, Taiwan. **Folia Geobotanica**, v. 53, n. 1, p. 77–88, 1 mar. 2018.

ZHAO, D.; WANG, Z.; WU, X.; QIU, T. Response of vegetation dynamics in environmentally sensitive and fragile areas to natural and anthropogenic factors: A case study in Inner Mongolia Autonomous Region, China. **Anthropocene**, v. 44, p. 100414, 1 dez. 2023.

ZHOU, T.; ZHANG, J.; QIN, Y.; JIANG, M.; QIAO, X. The Effects of Biotic and Abiotic Factors on the Community Dynamics in a Mountain Subtropical Forest. **Forests**, v. 12, n. 4, p. 427, 1 abr. 2021.

Capítulo I: Modelagem da ocorrência de infestação de lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina

RESUMO

As lianas desempenham um papel fundamental na estrutura florestal, na diversidade de espécies, bem como, aspectos funcionais das florestas tropicais. Apesar do crescente interesse no estudo de lianas nos trópicos nos últimos anos, a modelagem da ocorrência de lianas em fragmentos de Mata Atlântica ainda é escassa. Desse modo, o objetivo desse capítulo foi classificar a ocorrência de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando técnicas de aprendizagem de máquina. Três fragmentos florestais da Mata Atlântica em Minas Gerais, Brasil, foram analisados, estando localizadas nos municípios de Guanhães (Cachoeira das Pombas), Ipaba (Ipaba) e Coronel Fabriciano (São José). Em cada fragmento, parcelas permanentes retangulares de 0,05 ha foram estabelecidas. As árvores com circunferência à 1,30 m de altura (CAP) de pelo menos 15 cm foram mapeadas, sendo coletados o nome científico das espécies, a CAP, a altura total e o grau de infestação por lianas nos anos de 2002, 2012 e 2017. A infestação de lianas foi classificada em campo em quatro categorias: 1 - nenhuma liana; 2 - lianas apenas no tronco; 3 - lianas apenas na copa; 4 - e lianas no tronco e na copa. Posteriormente as árvores foram reclassificadas como infestadas, suscetíveis e mortas. Para a modelagem da infestação as redes neurais artificiais (RNA), a *random forest* (RF) e máquina de vetor de suporte (SVM) foram utilizadas. Variáveis explicativas como atributos da floresta, climáticas, paisagem, antropogênicas e de solo foram testadas. As variáveis quantitativas foram transformadas com normalização linear e os dados foram divididos em conjuntos de treinamento (70%) e validação (30%). O número de árvores infestadas por lianas aumentou 20,8% de 2012 para 2017, enquanto o número de árvores mortas aumentou 48% e o de árvores suscetíveis reduziu 6,6%. A RNA foi a técnica que apresentou o melhor desempenho, principalmente no período 2012-2017 na abordagem com modelo completo, com acurácia de 0,7257, sensibilidade de 0,5392 e especificidade de 0,7788. Esses resultados destacam que o uso das técnicas de aprendizagem de máquina é promissor para modelagem da ocorrência de lianas e podem ser ferramentas importantes na gestão e conservação da Mata Atlântica.

Palavras-chave: Floresta tropical; Perturbação da floresta; Trepadeiras lenhosas.

1 INTRODUÇÃO

A Mata Atlântica, um dos biomas mais biodiversos do planeta, tem sofrido intensa fragmentação devido à expansão urbana, agrícola e industrial (HADDAD et al., 2015; JOLY et al., 2010; RIBEIRO et al., 2009). Essa fragmentação resulta em paisagens com remanescentes florestais isolados e que impactam na dinâmica sucessional das florestas (GIRÃO et al., 2007; LEAL et al., 2012; PORTELA; DIRZO, 2020). A sucessão florestal na Mata Atlântica é caracterizada pela regeneração de espécies vegetais após distúrbios, contudo, a fragmentação dificulta esse processo, reduzindo a conectividade entre os fragmentos e alterando a composição e estrutura das comunidades vegetais (ANDRONACHE, 2024; SIMINSKI et al., 2021).

Nesse contexto, o aumento da ocorrência de lianas tem sido associado à fragmentação da Mata Atlântica, em que a perda de habitat e a criação de bordas florestais favorecem a proliferação dessas plantas devido ao aumento da disponibilidade de luz e espaço (CAMPBELL; LAURANCE; MAGRACH, 2015). As lianas são plantas lenhosas que utilizam outras plantas como suporte para alcançar a copa das árvores e possuem algumas características como os caules flexíveis, adaptação à luz intensa e rápido crescimento (HARRISON; GUZMÁN Q.; SÁNCHEZ-AZOFEIFA, 2021; VAN DER HEIJDEN; SCHNITZER; MEUNIER, 2023). As lianas impactam na dinâmica florestal ao competir com as árvores por recursos, aumentar a mortalidade das árvores e reduzir a taxa de crescimento e reprodução desses indivíduos (MARSHALL et al., 2020). Esses efeitos resultam em mudanças na estrutura e composição da floresta, influenciando a biodiversidade e a funcionalidade do ecossistema (MEUNIER et al., 2021).

Diante da influência das lianas na dinâmica florestal, a previsão da ocorrência de lianas se torna de grande importância para a gestão e conservação da Mata Atlântica. Contudo, essa previsão é um grande desafio devido à complexidade de fatores que influenciam a sua distribuição (ADDO-FORDJOUR; RAHMAD; BURNHAM, 2016). A heterogeneidade ambiental, a variação na disponibilidade de recursos e as interações bióticas são algumas das variáveis que dificultam essa previsão. Além disso, a ocorrência de fenômenos climáticos extremos aumenta os níveis de incerteza na modelagem da ocorrência de lianas (URIARTE; MUSCARELLA; ZIMMERMAN, 2018). Para contornar esses problemas, uma alternativa é a utilização de técnicas de aprendizagem de máquina (AM). Essas técnicas permitem analisar

grandes quantidades de dados e identificar padrões que podem ser utilizados para prever a ocorrência de lianas com maior precisão (BOCHENEK; USTRNUL, 2022; CHANDLER et al., 2021).

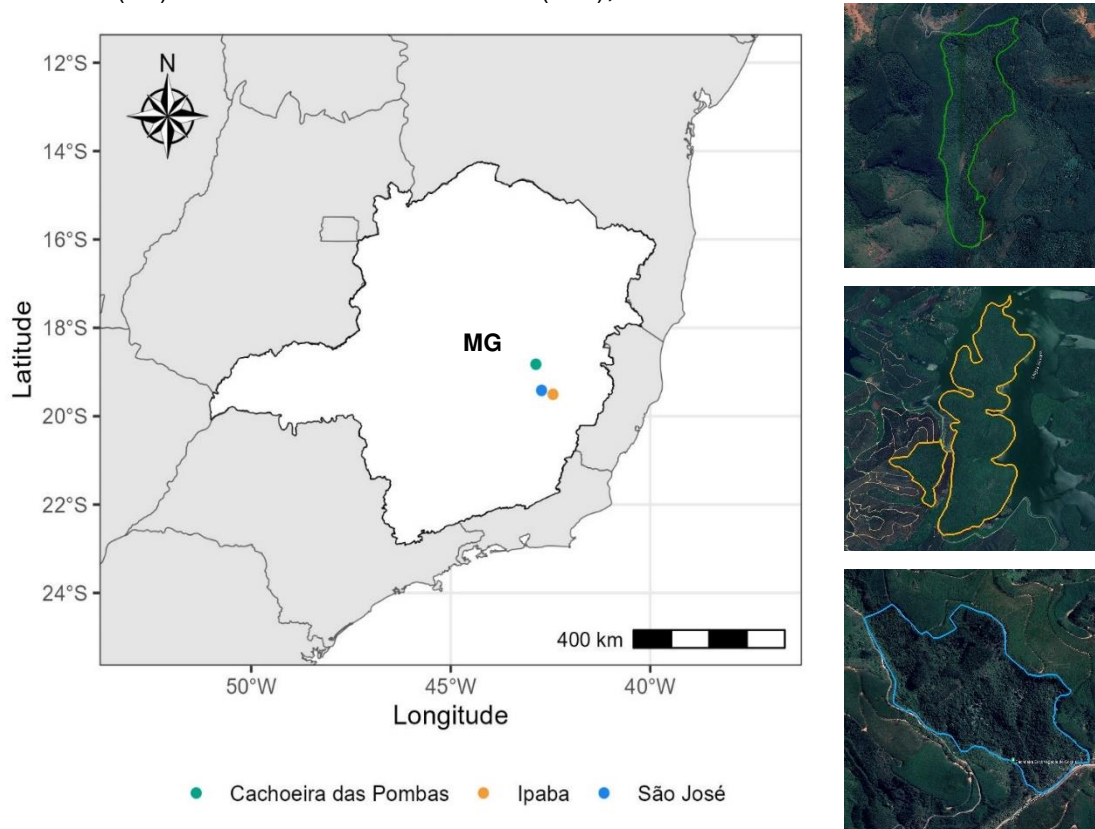
Embora as aplicações com AM para predição de atributos em florestas tropicais, como sobrevivência e mortalidade de árvores (ROCHA et al., 2018), crescimento de árvores (ROCHA et al., 2024; TAVARES JÚNIOR et al., 2020), biomassa e carbono (GHOSH; BEHERA, 2018; MASCARO et al., 2014) entre outras sejam recorrentes na literatura, a modelagem de lianas em fragmentos de Mata Atlântica ainda não foi abordada a partir dessas técnicas. Neste capítulo introduziu-se uma abordagem inédita para classificação de árvores infestadas por lianas utilizando redes neurais artificiais (RNA), o *random forest* (RF) e máquina de vetor de suporte (SVM). Diante desse contexto, o objetivo deste capítulo foi classificar a ocorrência de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando técnicas de aprendizagem de máquina.

2 MATERIAL E MÉTODOS

2.1 Caracterização das áreas de estudo

Três fragmentos florestais de Mata Atlântica denominados Cachoeira das Pombas, Ipaba (2 fragmentos) e São José foram analisados (Figura 1). Todas as áreas pertencem à região fitoecológica de Floresta Estacional Semidecidual (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2012).

Figura 1 - Mapa de localização dos fragmentos Cachoeira das Pombas (CP), Ipaba (IP) e São José (SJ) no Estado de Minas Gerais (MG), Brasil.



Fonte: O autor (2024).

2.1.1 Cachoeira das Pombas

A área possui 106 hectares e está situada no município de Guanhães, Minas Gerais, Brasil. Segundo a classificação climática de Köppen, o clima predominante é do tipo Cwa, caracterizado como tropical de altitude, com chuvas no verão (ALVARES et al., 2013). A temperatura média anual e a precipitação média anual no período de 2000 a 2017 foram de 20,85 °C e 1.112,74 mm, respectivamente (TORRES et al., 2023). A altitude na área varia entre 800 e 1.162 metros. A evapotranspiração potencial anual varia de 850 a 1.050 mm, conforme registros de uma estação meteorológica localizada em Guanhães, com longitude de 42° 58', latitude de 18° 48' e altitude de 801 metros (SOCIEDADE DE INVESTIGAÇÕES FLORESTAIS, 2008).

2.1.2 Ipaba

O fragmento está situado no município de Ipaba, Minas Gerais, Brasil. O

fragmento possui duas áreas florestais: Mata 1, com 264 hectares e altitude entre 200 e 450 metros, e Mata 2, com 37,30 hectares e altitude entre 150 e 300 metros. De acordo com a classificação climática de Köppen, o clima predominante na região é o Aw – tropical úmido, com inverno seco e verão chuvoso (ALVARES et al., 2013). Entre 2005 e 2017, a temperatura média anual foi de 24,15 °C e a precipitação média anual foi de 1.194,38 mm (TORRES et al., 2023). A evapotranspiração potencial anual varia entre 950 e 1.200 mm (SOCIEDADE DE INVESTIGAÇÕES FLORESTAIS, 2008). Os dados de ambas as matas foram compilados para aumentar o número de observações e permitir a execução das etapas subsequentes.

2.1.3 São José

A área de São José possui 38,40 hectares e está localizada no município de Coronel Fabriciano, Minas Gerais, Brasil. A altitude que varia entre 450 e 950 metros. De acordo com a classificação climática de Köppen, o clima predominante é o Cwb, típico de áreas tropicais de altitude (ALVARES et al., 2013). Entre 2000 e 2017, a temperatura média anual foi de 19,51 °C, com uma precipitação média anual de 1.263,92 mm (TORRES et al., 2023). A evapotranspiração potencial anual varia entre 750 e 950 mm, conforme dados de uma estação meteorológica em Antônio Dias, situada a uma longitude de 42° 51', latitude de 19° 27' e altitude de 950 metros.

2.2 Coleta dos dados

Parcelas permanentes retangulares com área fixa de 10 × 50 m (0,05 ha) foram estabelecidas e sinalizadas para facilitar a localização das árvores em futuras ocasiões de monitoramento e coleta de material botânico. Cada parcela foi subdividida em 20 subparcelas de inventário de 5 × 5 m, e as árvores com circunferência à altura do peito (CAP) de pelo menos 15 cm foram mapeadas. O nome científico das espécies, a CAP (cm), a altura total (Ht, em m) e o grau de infestação por lianas em árvores com CAP ≥ 15 cm foram registrados nos anos de 2002, 2012 e 2017, em cada fragmento. Posteriormente, o CAP foi convertido em diâmetro à 1,3 m de altura (DAP, em cm).

A classificação da infestação por lianas nas árvores em campo foi realizada de forma visual de acordo com as seguintes classes: 1 - Nenhuma liana; 2 - Lianas



apenas no tronco; 3 - Lianas apenas na copa; e 4 - Lianas no tronco e na copa.


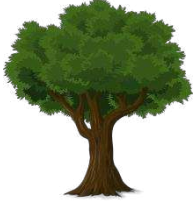


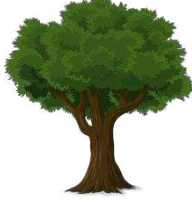


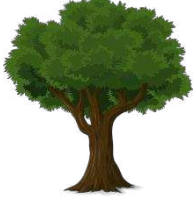



2.3 Modelagem da infestação de lianas


A classificação de campo da infestação de lianas (classes de 1 a 4) foi a base para a definição das novas classes utilizadas na modelagem. De forma prévia, a modelagem com diferentes classes de infestação foi testada de modo a otimizar o tempo de processamento. Assim, as classes definitivas para a modelagem em nível de árvore individual foram as seguintes: infestada, livre e morta, sendo que a classificação foi referente ao ano final de observação (a_2) (Figura 2). As árvores **infestadas** são aquelas que apresentam ou não lianas no ano a_1 e obrigatoriamente possuem lianas no ano a_2 . A classe **livre** são as árvores que apresentam ou não lianas no ano a_1 e obrigatoriamente não possuem lianas no ano a_2 , ou seja, estão livres de lianas. As árvores mortas são aquelas com ou sem lianas no ano a_1 e que foram registradas como mortas no ano a_2 . Além disso, as árvores que não possuíam registro no ano a_1 e que passaram a ser registradas com a presença de lianas no ano a_2 foram classificadas como **infestadas**. Já as árvores que passaram a ser registradas no ano a_2 sem a presença de lianas foram classificadas como **livres**.

Posteriormente, as classes de infestação de árvores por lianas no ano a_2 (2012, por exemplo) foram modeladas com base nas variáveis referentes ao ano a_1 (2002, por exemplo). Assim, os períodos de classificação foram 2002-2012 (10 anos), 2002-2017 (15 anos) e 2012-2017 (5 anos).

Figura 2 - Classificação da infestação das árvores com base na presença ou ausência de lianas nos anos a_1 e a_2

Ano a_1	Ano a_2	Classe
		Infestada
Com lianas	Com lianas	

Ano a_1	Ano a_2	Classe
 Com lianas	 Sem lianas	Suscetível
 Com lianas	 Morta	Morta
 Sem lianas	 Morta	Morta
 Sem lianas	 Sem lianas	Suscetível
 Sem lianas	 Com lianas	Infestada
Sem registro	 Sem lianas	Suscetível

Ano a_1	Ano a_2	Classe
Sem registro		Infestada
	Com lianas	

Fonte: O autor (2024).

2.3.1 Variáveis explicativas

Variáveis de solo, topográficas, climáticas, antropogênicas e atributos da floresta foram testadas como variáveis explicativas para a modelagem das classes de infestação de lianas. As Tabelas A1 e A2 contém as siglas, descrições e estatísticas descritivas de todas as variáveis explicativas utilizadas referentes aos anos de 2002 e 2012, respectivamente (APÊNDICE A).

2.3.1.1 Atributos da floresta e variáveis antropogênicas

Os atributos da floresta considerados foram a área basal por hectare, número de fustes por hectare em cada parcela dos fragmentos, número de espécies, DAP, Ht das árvores e a classe de lianas registrada em campo no ano a_1 . As variáveis antropogênicas utilizadas incluíram o histórico de uso e mudança de uso da terra. Com base em dados coletados de fotografias aéreas (aproximadamente de 1960, 1980 e 1990), imagens de satélite Landsat de 1985, entrevistas com proprietários e registros de propriedade, foi determinado o histórico de uso da terra das áreas. As categorias de uso da terra consideradas foram: desmatamento, produção agrícola, plantio de eucalipto e extração seletiva de madeira (TORRES et al., 2023).

2.3.1.2 Variáveis climáticas

A precipitação anual, número de meses com menos de 100 mm de precipitação, precipitação nos três meses mais secos e temperatura média anual para cada fragmento da estação climatológica mais próxima (ROCHA et al. 2020). Em seguida, o déficit hídrico climático (CWD) foi estimado – um balanço hídrico entre

precipitação e evapotranspiração – como uma aproximação para condições de seca seguindo Lutz et al. (2010) e usando a função do R CWD e AET (evapotranspiração real) de Redmond (2019). O CWD reflete as condições de seca com mais precisão do que a precipitação total anual (CHAVE et al., 2014). O CWD mais negativo indica condições de alto estresse hídrico e valores próximos a 0 (zero) indicam não estresse hídrico (POORTER et al., 2017). A média da precipitação anual total e CWD para um, dois, três e quatro anos antes do ano de medição também foi calculado (TORRES et al., 2023).

2.3.1.3 Variáveis do solo e da paisagem

Cerca de 20-30 amostras do solo foram coletadas por parcela para obter uma amostra composta. Informações para as profundidades de 0-20 cm e 20-40 cm foram obtidas. Algumas das variáveis coletadas foram: pH do solo em H₂O, cátions trocáveis (Ca²⁺, Mg²⁺ e Al³⁺), acidez total (H⁺ + Al³⁺), capacidade de troca catiônica (CTC), saturação por bases (V), fósforo disponível (P), P remanescente na solução (P-rem) e matéria orgânica do solo (MO) foram determinados usando métodos padrão (TEIXEIRA et al., 2017). As análises de solo foram realizadas no laboratório de Fertilidade do Solo da Universidade Federal de Viçosa (UFV), Minas Gerais, Brasil.

As variáveis de paisagem foram a distância da borda, elevação, inclinação e o nome dos fragmentos. A elevação e inclinação foram calculados para cada parcela usando *Spatial Analyst Tools* de superfície no ArcGIS® 10.3.1 1 (ESRI, 2015). O *Shuttle Radar Topography Mission* (SRTM) e *Digital Elevation Models* (DEMs) foram utilizados para a análise (TORRES et al., 2023). A distância até a borda mais próxima da floresta foi calculada com a ferramenta "*near*" no ArcGIS® 10.3.1 (ESRI, 2015)

2.3.2 Seleção das variáveis explicativas

Após a etapa anterior, duas abordagens para seleção das variáveis explicativas foram testadas: a) abordagem com modelo completo: seleção das variáveis quantitativas que não possuíam correlação linear maior que 0,8 com outras variáveis. As variáveis qualitativas foram posteriormente compiladas com as variáveis selecionadas e; b) abordagem com modelo reduzido: seleção das variáveis quantitativas e/ou qualitativas mais importantes utilizando o algoritmo *Random Forest*

(RF). A colinearidade entre as variáveis explicativas pode não afetar severamente os métodos não paramétricos, contudo, a exclusão dessas variáveis favorece a escolha de modelos mais parcimoniosos e sem redundância (CHAN et al., 2022).

Por meio da técnica RF, as variáveis mais importantes são determinadas através do aumento percentual do erro quadrático médio quando uma variável preditora é permutada aleatoriamente. Assim, quanto maior esse erro, mais importante uma determinada variável é em relação as demais. A função utilizada nessa etapa foi a *VSURF* do pacote “*VSURF*” no ambiente R (R CORE TEAM, 2022). Após a aplicação da RF, as variáveis mais importantes para predição foram selecionadas para o treinamento das redes neurais artificiais (RNA), regressão por vetores de suporte (RVS) e das próprias RF.

2.4 Pré-processamento dos dados

As variáveis quantitativas foram escalonadas, sendo que a amplitude das variáveis de entrada e saída ficaram no intervalo de 0 a 1 (Equação 1). Essa transformação possibilita acelerar a taxa de convergência e reduzir o processo de iteração no treinamento das técnicas de aprendizagem de máquina (TAVARES JÚNIOR et al., 2020).

$$v' = \frac{v - \min(v)}{\max(v) - \min(v)} (\max_{\text{novo}} - \min_{\text{novo}}) + \min_{\text{novo}} \quad (1)$$

Em que: v' = valor da observação transformado; v = valor da observação não transformado; $\min(v)$ = menor valor da variável no banco de dados; $\max(v)$ = maior valor da variável no banco de dados; \max_{novo} = maior valor da transformado; e \min_{novo} = menor valor da transformado.

Os dados foram divididos de forma aleatória em conjuntos de treinamento e validação, sendo que 70% dos registros de cada classe de infestação foram utilizados nos treinamentos e 30% nas validações.

2.5 Treinamento das *Random Forests* (RF)

A classificação por *Random Forest* (RF) é uma abordagem de modelagem que

utiliza um grande conjunto de árvores de decisão. Cada árvore é construída a partir de uma amostra aleatória dos dados de treinamento, permitindo substituições (BREIMAN, 2001). Em cada nó das árvores, um subconjunto aleatório das variáveis preditoras é empregado para determinar a divisão mais eficiente. Essa eficiência é definida pela identificação da variável preditora e do ponto de divisão que maximiza a redução na soma dos quadrados residuais entre as observações da amostra e a média do nó. Todas as árvores são cultivadas ao máximo, com o tamanho do nó controlado pelo usuário. O resultado é um conjunto (floresta) de baixa polarização e alta variância nas árvores de decisão, cujas previsões finais são obtidas pelo voto majoritário das árvores (BREIMAN, 2001).

No treinamento das RF três parâmetros foram configurados: o número de árvores de regressão aleatória (*ntree*), no qual foram testados de 500, 1000, 1500, 2000, 2500 e 3000 árvores; o número de variáveis de divisão (*mtry*), utilizado para determinar o número de variáveis disponíveis para dividir em cada nó da árvore, sendo que o número padrão de 1/3 das variáveis independentes foi utilizado; e o tamanho mínimo de nós terminais (*nodesize*), cujo valor foi 5.

A função *randomForest* do pacote de mesmo nome no *R* foi utilizada para os treinamentos das RFs (LIAW; WIENER, 2002).

2.6 Treinamento das redes neurais artificiais (RNA)

As RNA treinadas foram do tipo *Multilayer Perceptron* (MLP), com uma camada oculta (HAYKIN, 2009). Nessa camada, o intervalo do número de neurônios foi determinado pelo método de Fletcher-Gloss (SILVA et al., 2010): $2 \times n^{0.5} + n_2 \leq n_1 \leq 2 \times n + 1$; em que: n = número de entradas da rede; n_1 = número de neurônios na camada oculta; e n_2 = número de neurônios na camada de saída.

Os algoritmos testados foram o *Resilient Propagation* (*Rprop*) e o *Scaled Conjugate Gradient* (*SCG*). As funções de ativação testadas foram a tangente hiperbólica, logística e exponencial. Os pesos iniciais das RNA foram gerados aleatoriamente, sendo os valores atualizados com base na função do erro. O número máximo de ciclos foi definido em 1000. As RNA foram treinadas utilizando a função *mlp* do pacote "*RSNNS*" no programa *R* (BERGMEIR; BENITEZ, 2012).

2.7 Treinamento das máquinas de vetores de suporte (SVM)

Nos treinamentos das SVM foram utilizadas quatro configurações, representadas por quatro funções kernel, em que a seguinte função de erro foi otimizada:

$$\text{minimize } \frac{1}{2} \|w\|^2 + \frac{C}{m} \sum_{i=1}^m (\xi_i + \xi_i^*)$$

$$\text{sujeito a } (\langle \Phi(x_i), w \rangle + b) - y_i \leq \epsilon - \xi_i^*$$

$$y_i - (\langle \Phi(x_i), w \rangle + b) \leq \epsilon - \xi_i$$

$$\xi_i \geq 0 \quad (i = 1, \dots, m)$$

em que: y_i = dados de saída e; ϵ = erro máximo admitido; w = vetor de coeficientes; i = casos de treinamento; m = número total de casos de treinamento; C = parâmetro de penalidade do erro; ξ_i, ξ_i^* = variáveis de folga que caracterizam o erro acima e abaixo do ϵ -tubo, respectivamente; $\Phi(x_i)$ = kernel utilizado; b = bias.

As funções kernel avaliadas foram as do tipo linear, polinomial, base radial e sigmóide (Tabela 1).

Tabela 1 - Funções kernel testadas nos treinamentos das regressões vetoriais de suporte

Tipo de kernel	Função
Linear	$u' \times v$
Polinomial	$(\gamma \times u' \times v + coef0)^d$
Base radial	$e^{(-\gamma \times u-v ^2)}$
Sigmoidal	$\tanh(\gamma \times u' \times v + coef0)$

Em que: γ = parâmetro de forma necessário para todos os kernels, exceto linear [padrão: $1/(\text{dimensão de dados})$]; d = grau do polinômio; $coef0$ = parâmetro de penalidade do erro; \tanh = tangente hiperbólica.

A função utilizada para os treinamentos das SVM foi a *svm* do pacote “*e1071*” no R (MEYER et al., 2018).

2.8 Seleção das melhores configurações das técnicas

Após os treinamentos e validações, matrizes de confusão foram construídas

para cada configuração das técnicas, sendo essas matrizes a base para o cálculo das seguintes estatísticas: acurácia (Equação 2), sensibilidade (Equação 3) e especificidade (Equação 4). A curva ROC (*Receiver Operating Characteristic*) posteriormente também foi construída para avaliar e selecionar o melhor classificador.

$$\text{Acurácia} = \frac{VP + VN}{VP + FN + VN + FP} \quad (2)$$

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (3)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (4)$$

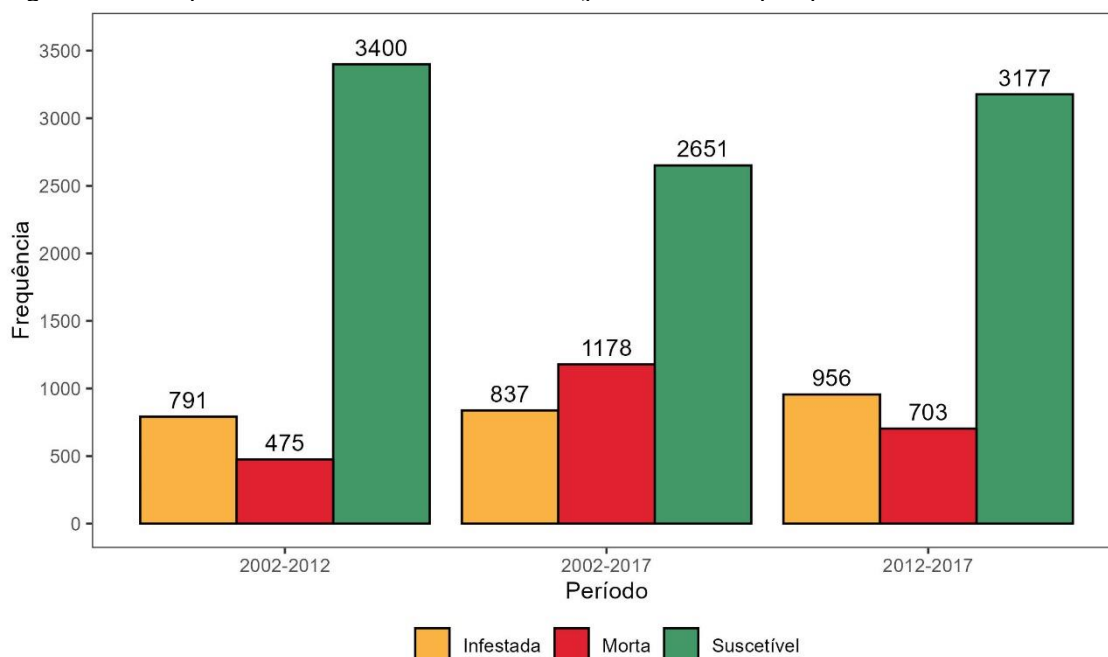
Em que: VP = verdadeiros positivos; VN = verdadeiros negativos; FN = falsos negativos e; FP = falsos positivos.

A escolha da melhor técnica foi baseada no melhor desempenho das estatísticas e curvas ROC na validação.

3 RESULTADOS

Analisando o número de árvores em cada *status* de infestação, o número de árvores infestadas por lianas aumentou, aproximadamente, 20,8% de 2012 para 2017 (Figura 3). Nesse mesmo período, o número de árvores mortas aumentou cerca de 48%, enquanto o número de árvores suscetíveis reduziu 6,6%.

Figura 3 - Frequência das classes de infestação de lianas por período.



Fonte: O autor (2024).

3.1 Variáveis selecionadas

Na seleção das variáveis através da abordagem do modelo completo, 22 variáveis foram selecionadas para classificar a ocorrência de árvores infestadas por lianas nos três períodos (Tabela 2). Por outro lado, na abordagem com modelo reduzido, 3, 6 e 2 variáveis foram selecionadas para classificação da infestação de lianas nos períodos 2002-2012, 2002-2017 e 2012-2017, respectivamente. As correlações das variáveis quantitativas selecionadas em cada período para a abordagem com modelo completo podem ser observadas nas Figuras A1 e A2. As etapas de seleção das variáveis mais importantes pela RF na abordagem com modelo reduzido estão apresentadas nas Figuras A3 a A5.

Tabela 2 - Variáveis selecionadas por abordagem (modelo completo e reduzido) para os três períodos avaliados

Abordagem	Período	Variáveis
Modelo completo	2002-2012; 2002-2017; 2012-2017	Al3_20_40
		BA_ha
		Changes
		Lianas_2002
		CWD
		DAP_2002
		Declivity_perc
		Edge_distance
		Fragmentos
		GE
		Ht_2002
		LUH
		m_20_40
		MO_20_40
		P_20_40
		pH_H2O_20_40
		Precp_avrg
Precp_dry		
Slope_Angle		
Species		
Stems_ha		
t_20_40		
Modelo reduzido	2002-2012	Lianas_2002
		Ht_2002
		DAP_2002
	2002-2017	BA_ha
		Lianas_2002
		GE
		P_Rem_0_20
		Species
	T_20_40	
	2012-2017	Lianas_2012
Elevation		

Em que: Al3_20_40 = Alumínio trocável (cmolc/dm³) 20cm - 40 cm; Changes = Mudanças no uso da terra; CWD = Déficit hídrico climático (ano de medição); Declivity_perc = Declividade (%); Edge_distance = Distância da borda (m); LUH = Histórico de uso da terra; m_20_40 = Saturação por alumínio (%) de 20 cm - 40 cm; MO_20_40 = Matéria orgânica (dag/kg) de 20 cm - 40 cm; P_20_40 = Fósforo (cmolc/dm³) de 20 cm - 40 cm; pH_H2O_20_40 = pH em água de 20 cm - 40 cm; Precp_avrg = Precipitação média anual (de 1989 até 1 ano antes da medição); Precp_dry = Precipitação total dos três meses mais secos; Slope_Angle = Ângulo de inclinação (°); Species = Número de espécies; t_20_40 = CTC efetiva (cmolc/dm³) de 20 cm - 40 cm; T_20_40 = CTC potencial (cmolc/dm³) de 20 cm - 40 cm; Lianas_2002 = Classe de infestação de lianas registrada em campo no ano de 2002; Lianas_2012 = Classe de infestação de lianas registrada em campo no ano de 2012; GE = Grupo ecológico; P_Rem_0_20 = Fósforo remanescente (mg/L) de 0 cm - 20 cm; Elevation = Elevação (m); BA_ha = área basal por hectare.

3.2 Avaliação da classificação da infestação de lianas pelas técnicas

Nas estatísticas do treinamento, a rede neural artificial (RNA) apresentou maior acurácia em todas as abordagens e períodos (com exceção do período 2012-2017 na abordagem com modelo reduzido), com valores de acurácia variando entre 0,6922 e 0,8151 (Tabela 3). O *random forest* (RF), embora tenha uma acurácia ligeiramente inferior a RNA e SVM, apresentou desempenho balanceado entre sensibilidade e especificidade. A quantidade de árvores de decisão na RF variou entre os períodos e abordagens, variando de 500 a 2500.

Em relação a abordagem com modelo completo *versus* reduzido, a abordagem com modelo completo apresentou valores maiores na maior parte das estatísticas. No geral, as RNA com o algoritmo *Resilient Propagation* e função de ativação exponencial apresentaram maiores valores de acurácia ($> 0,65$), sensibilidade ($> 0,51$) e especificidade ($> 0,76$).

Tabela 3 - Estatísticas de acurácia, sensibilidade e especificidade das melhores configurações das técnicas de aprendizagem de máquina no treinamento

Período	Abordagem	Modelo	Acurácia	Sensibilidade	Especificidade	Kernel ^a	Algoritmo ^b	Função ^c	Neurônios ^d	Árvores ^e
2002-2012	Completo	SVM	0,7900	0,5070	0,7618	Polinomial	-	-	-	-
		RF	0,7393	0,4326	0,7264	-	-	-	-	1000
		RNA	0,8151	0,6064	0,8143	-	Rprop	Exponencial	22	-
	Reduzido	SVM	0,7417	0,3834	0,6990	Polinomial	-	-	-	-
		RF	0,7320	0,3741	0,6948	-	-	-	-	2500
		RNA	0,7479	0,4535	0,7471	-	Rprop	Logística	14	-
2002-2017	Completo	SVM	0,7173	0,5946	0,8046	Polinomial	-	-	-	-
		RF	0,6278	0,5030	0,7564	-	-	-	-	1000
		RNA	0,7240	0,6184	0,8159	-	Rprop	Exponencial	19	-
	Reduzido	SVM	0,6342	0,4809	0,7457	Polinomial	-	-	-	-
		RF	0,6155	0,4935	0,7565	-	-	-	-	2000
		RNA	0,6522	0,5138	0,7611	-	Rprop	Exponencial	26	-
2012-2017	Completo	SVM	0,7636	0,5823	0,7972	Polinomial	-	-	-	-
		RF	0,7211	0,5212	0,7729	-	-	-	-	500
		RNA	0,7839	0,6246	0,8181	-	Rprop	Exponencial	29	-
	Reduzido	SVM	0,6990	0,4649	0,7524	Radial	-	-	-	-
		RF	0,6949	0,4482	0,7417	-	-	-	-	1500
		RNA	0,6922	0,4583	0,7508	-	Rprop	Tangente hiperbólica	6	-

Em que: Rprop = *resilient propagation*; SCG = *scaled conjugate gradient*; RNA = rede neural artificial; SVM = máquina de vetor de suporte; RF = *random forest*.
^a Função Kernel utilizada nas SVM. ^b Algoritmo aplicado na camada oculta ^c Função ativação na camada oculta da RNA. ^d Número de neurônios na camada oculta da RNA. ^e Número de árvores de decisão na RF.

Para as estatísticas de validação das técnicas na abordagem com modelo completo do período 2002-2012, a RNA apresentou maiores valores de acurácia, sensibilidade e especificidade (Tabela 4). Na abordagem com modelo reduzido, os valores das estatísticas, de modo geral diminuíram, principalmente na sensibilidade, com valores iguais a 0,37 para a SVM e RF, e de 0,40 para a RNA.

No período 2002-2017, a RNA novamente obteve a maior acurácia, sensibilidade e especificidade na abordagem com modelo completo. Na abordagem com modelo reduzido, todos os modelos apresentam uma queda de desempenho, com a SVM apresentando novamente a menor acurácia.

No período 2012-2017 na abordagem completa, as técnicas apresentaram performance mais equilibrada entre acurácia e sensibilidade na abordagem completa. A RNA apresentou novamente os maiores valores em acurácia, sensibilidade e especificidade, com valores de 0,72, 0,53 e 0,77, respectivamente. Na abordagem com modelo reduzido, as estatísticas de todas as técnicas reduziram, com a SVM apresentando os maiores valores de sensibilidade e especificidade.

Tabela 4 - Estatísticas de acurácia, sensibilidade e especificidade das melhores configurações das técnicas de aprendizagem de máquina na validação

Período	Abordagem	Modelo	Acurácia	Sensibilidade	Especificidade	Kernel ^a	Algoritmo ^b	Função ^c	Neurônios ^d	Árvores ^e
2002-2012	Completo	SVM	0,7459	0,4050	0,7093	Radial	-	-	-	-
		RF	0,7530	0,4445	0,7334	-	-	-	-	1000
		RNA	0,7566	0,4651	0,7405	-	Rprop	Logística	21	-
	Reduzido	SVM	0,7380	0,3732	0,6935	Polinomial	-	-	-	-
		RF	0,7380	0,3743	0,6946	-	-	-	-	500
		RNA	0,7259	0,4014	0,7188	-	SCG	Exponencial	15	-
2002-2017	Completo	SVM	0,6238	0,4888	0,7491	Polinomial	-	-	-	-
		RF	0,6246	0,5041	0,7554	-	-	-	-	500
		RNA	0,6367	0,5150	0,7640	-	SCG	Logística	22	-
	Reduzido	SVM	0,5939	0,4347	0,7227	Polinomial	-	-	-	-
		RF	0,6188	0,4939	0,7583	-	-	-	-	1000
		RNA	0,6124	0,4665	0,7378	-	Rprop	Exponencial	18	-
2012-2017	Completo	SVM	0,7236	0,5262	0,7687	Polinomial	-	-	-	-
		RF	0,7229	0,5229	0,7762	-	-	-	-	2000
		RNA	0,7257	0,5392	0,7788	-	Rprop	Logística	30	-
	Reduzido	SVM	0,7009	0,4715	0,7535	Radial	-	-	-	-
		RF	0,7085	0,4636	0,7482	-	-	-	-	500
		RNA	0,6912	0,4568	0,7506	-	Rprop	Exponencial	6	-

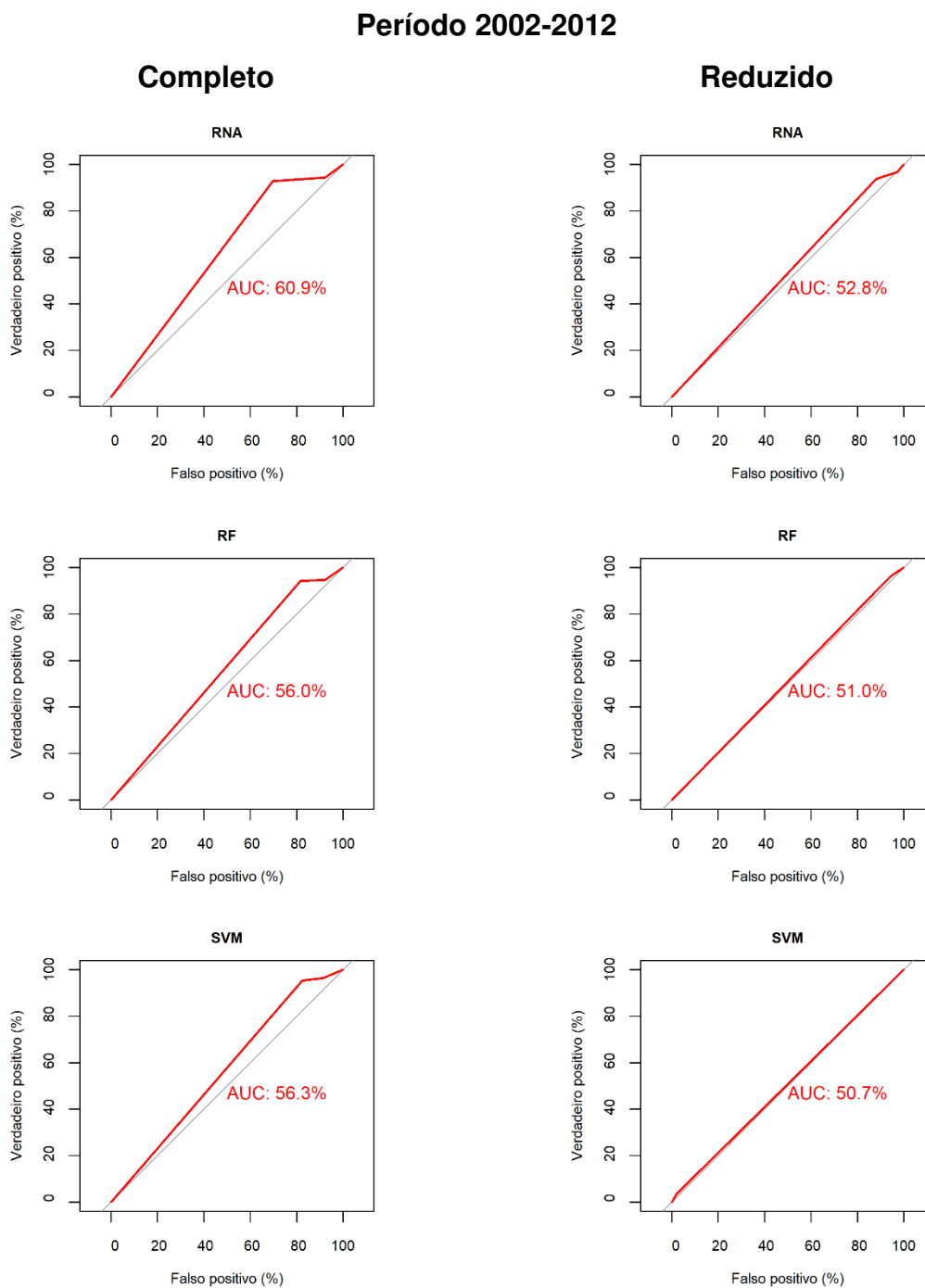
Em que: Rprop = *resilient propagation*; SCG = *scaled conjugate gradient*; RNA = rede neural artificial; SVM = máquina de vetor de suporte; RF = *random forest*. ^a Função Kernel utilizada nas SVM. ^b Algoritmo aplicado na camada oculta ^c Função ativação na camada oculta da RNA. ^d Número de neurônios na camada oculta da RNA. ^e Número de árvores de decisão na RF.

No período de 2002 a 2012, ao avaliar as curvas ROC das técnicas na validação, a RNA na abordagem com modelo completo apresentou maior valor de área sob a curva (AUC) em relação às outras técnicas (Figura 4). A RF e SVM apresentaram desempenho similar entre si, com AUC próximas. Na abordagem com modelo reduzido, os valores de AUC diminuíram para todas as técnicas. Por outro lado, a redução nos valores de AUC para a RF e SVM foram menores e relativamente próximos aos valores da abordagem com modelo completo. A RNA apresentou a maior redução na AUC.

No período de 2002 a 2017, na abordagem com modelo completo, a RNA manteve a maior AUC (Figura 5). A RF apresentou uma AUC próxima ao da RNA, enquanto o SVM teve a menor AUC. Na abordagem reduzida, a AUC de todas as técnicas diminuiu em relação a abordagem completa. A RNA permaneceu com o maior AUC, seguida pelo RF e depois pela SVM, com o menor valor de AUC. Na abordagem reduzida, os valores de AUC também diminuíram para todas as técnicas, com a SVM novamente apresentando a menor AUC.

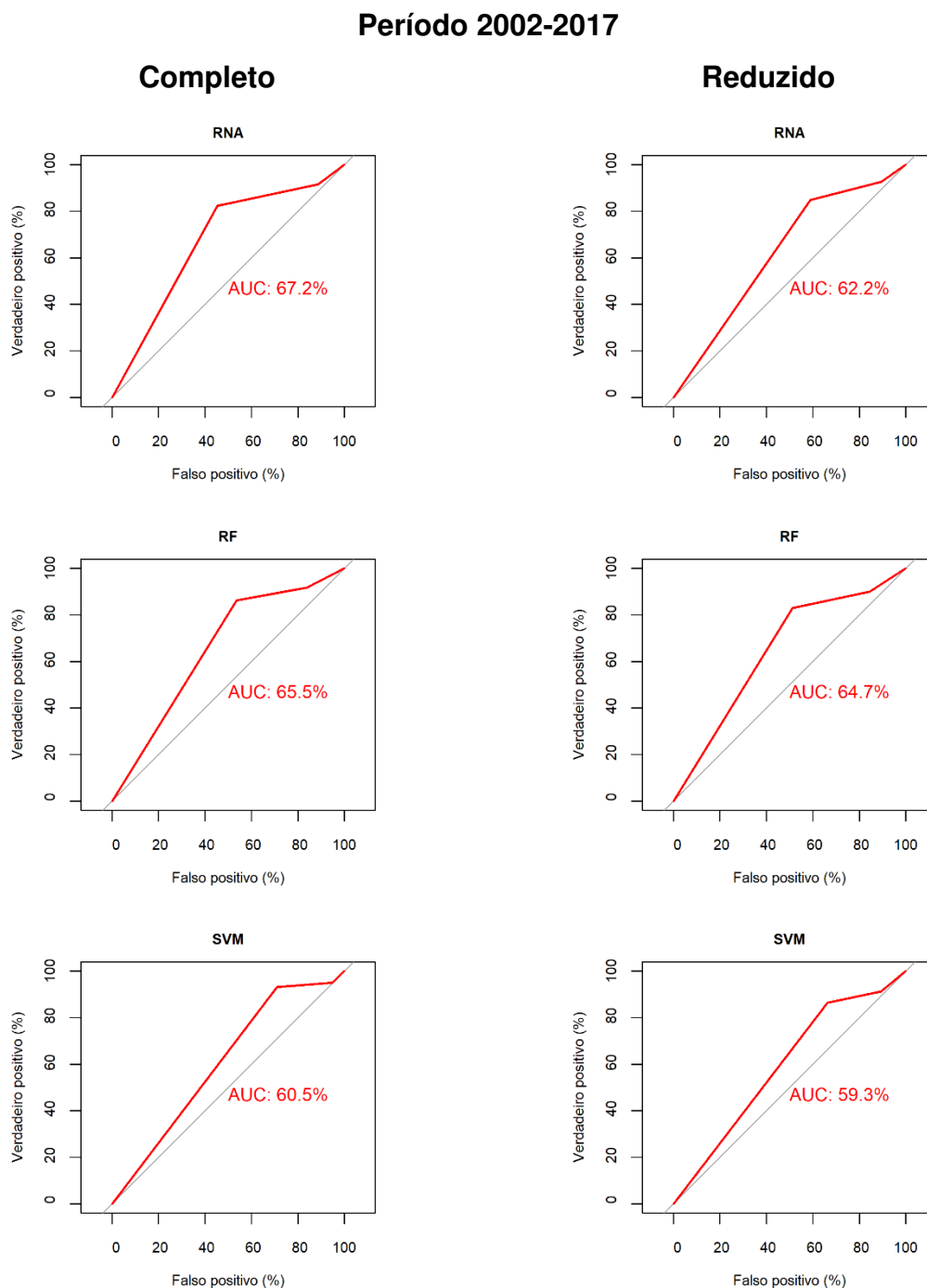
No período 2012 a 2017, na abordagem com modelo completo, as técnicas apresentaram valores de AUC equilibrados (Figura 6). A RNA continuou com a maior área sob a curva (AUC), seguida pelo RF e pelo SVM, que também apresentaram valores próximos de AUC. Na abordagem reduzida, os valores de AUC reduziram para todas as técnicas. Apesar dessa redução, a SVM se destacou por apresentar o maior valor de AUC.

Figura 4 - Curvas ROC das redes neurais artificiais (RNA), máquina de vetor de suporte (SVM) e *random forest* (RF) nas abordagens com modelo completo e reduzido no período 2002-2012.



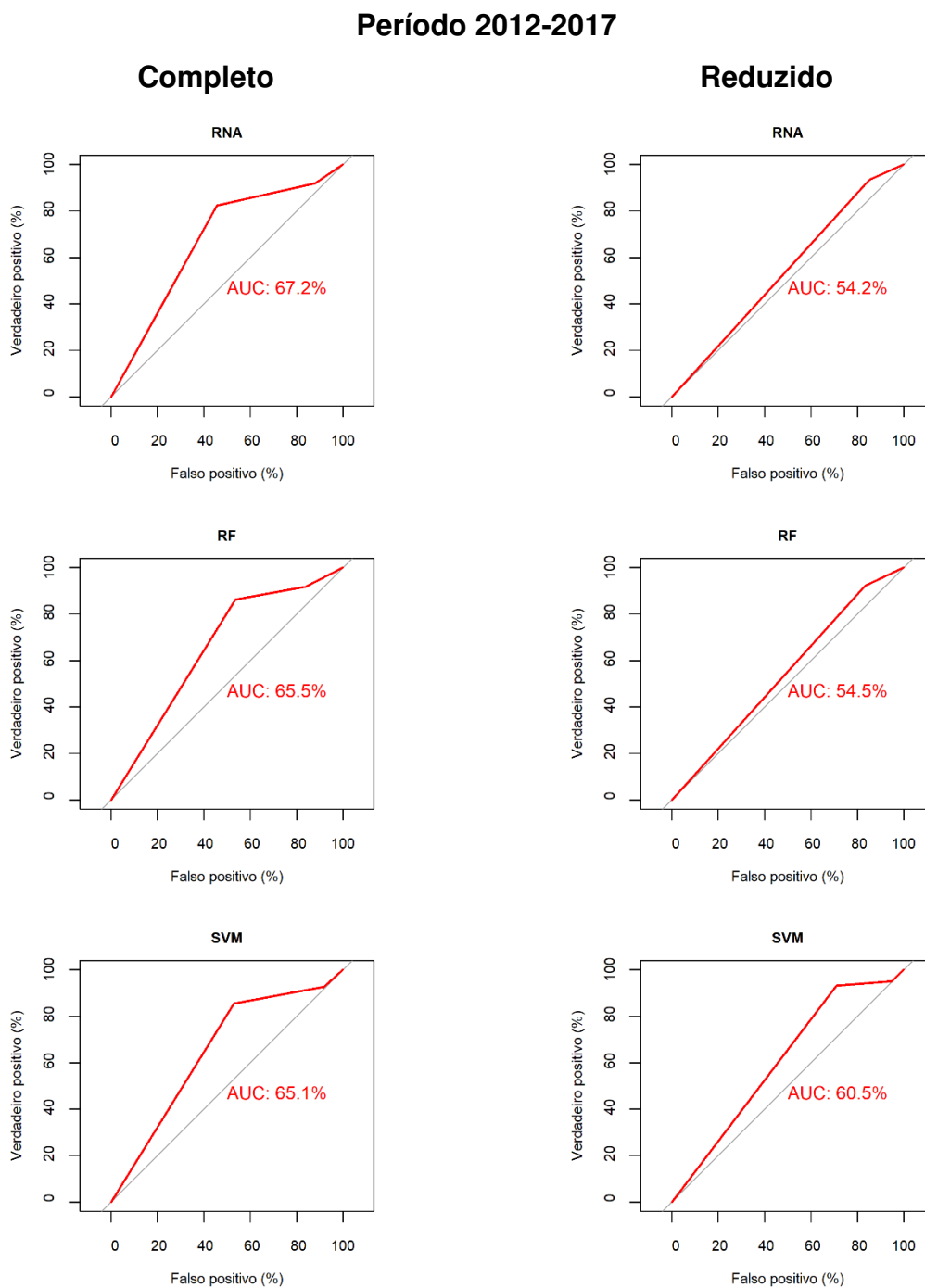
Fonte: O autor (2024).

Figura 5 - Curvas ROC das redes neurais artificiais (RNA), máquina de vetor de suporte (SVM) e *random forest* (RF) nas abordagens com modelo completo e reduzido no período 2002-2017.



Fonte: O autor (2024).

Figura 6 - Curvas ROC das redes neurais artificiais (RNA), máquina de vetor de suporte (SVM) e *random forest* (RF) nas abordagens com modelo completo e reduzido no período 2012-2017.



Fonte: O autor (2024).

4 DISCUSSÃO

Neste estudo, avaliou-se de forma inédita a aplicação de técnicas de aprendizagem de máquina (RNA, SVM e RF) na modelagem de ocorrência de árvores infestadas por lianas em fragmentos da Mata Atlântica para diferentes períodos (2002-2012, 2002-2017 e 2012-2017) e abordagens (modelo completo e reduzido).

O aumento de aproximadamente 20,8% no número de árvores infestadas por lianas de 2012 para 2017 sugere que as lianas estão aumentando a sua distribuição e afetando um número crescente de árvores. Esse aumento pode estar relacionado a distúrbios florestais e a fenômenos climáticos (MARSHALL et al., 2020; VOGADO et al., 2022; YUAN et al., 2019). Um desses fenômeno pode ter sido o El Niño no período de 2015/2016, que foi um dos mais fortes registrados, trazendo secas severas para muitas regiões tropicais (ESQUIVEL-MUELBERT et al., 2019; THOMPSON et al., 2017). Essas condições de seca podem ter contribuído para o aumento da mortalidade de árvores e a proliferação de lianas. Durante os períodos de seca, as árvores são mais vulneráveis ao estresse hídrico, tornando esses indivíduos mais suscetíveis a infestações (ALLEN et al., 2010; HAMMOND et al., 2022). As lianas, por outro lado, possuem adaptações que lhes permitem sobreviver em condições de seca, como sistemas radiculares profundos que podem acessar água em camadas mais profundas do solo (DIAS et al., 2019; LIU et al., 2021).

O aumento da mortalidade das árvores em 48% no período analisado pode ser atribuído ao estresse hídrico causado pela seca prolongada durante o El Niño. Além disso, as lianas aumentam esse estresse ao competir com as árvores, além da carga física sobre as árvores hospedeiras, levando a um aumento na mortalidade (MARSHALL et al., 2020; MEDINA-VEGA et al., 2021).

A redução de 6,6% no número de árvores suscetíveis ocorreu devido ao aumento da mortalidade e da infestação. À medida que mais árvores são infestadas e morrem, o número de árvores que permanecem suscetíveis diminui. Este fenômeno pode indicar que a população de árvores está sendo progressivamente afetada e que as árvores restantes estão cada vez mais sob pressão das lianas (BECKNELL et al., 2022; ESTRADA-VILLEGAS et al., 2022).

4.1 Importância das variáveis

No período de 2002 a 2012, as variáveis mais importantes foram altura das árvores (Ht_2002), o diâmetro à 1,3 m de altura (DAP_2002) e a classe de infestação de lianas em 2002 (Lianas_2002). A seleção das variáveis diâmetro e altura indica que árvores maiores fornecem mais suporte estrutural para o crescimento das lianas, aumentando a probabilidade de infestação (LEICHT-YOUNG et al., 2010). Além disso, a presença das classes prévias de lianas no ano de 2002 (Lianas_2002) auxilia na indicação de quais são as árvores mais favoráveis a infestação de lianas.

Para o período de 2002 a 2017, além da classe de infestação de lianas em 2002, as variáveis mais importantes foram área basal por hectare (BA_ha), grupo ecológico (GE), fósforo remanescente no solo (P_Rem_0_20), número de espécies (Species) e a concentração de alumínio trocável no solo a 20-40 cm de profundidade (Al3_20_40). A importância da BA_ha sugere que uma maior área basal pode representar uma floresta mais madura e estruturada, proporcionando suporte físico adicional para as lianas crescerem e se estabelecerem (CEBALLOS; MALIZIA, 2017).

A importância do alumínio trocável sugere que concentrações mais altas podem ser tóxicas para muitas plantas, inibindo o crescimento radicular e a absorção de nutrientes essenciais, afetando negativamente as árvores hospedeiras e criando um ambiente competitivo mais favorável para as lianas (ZHANG et al., 2023a). A importância do fósforo remanescente no solo pode estar relacionada à fertilidade do solo e à disponibilidade de nutrientes, que podem beneficiar tanto as árvores como as lianas (CAI; BONGERS, 2007).

Em relação ao grupo ecológico, geralmente as pioneiras são mais suscetíveis à infestação de lianas devido à sua alta taxa de crescimento e por serem encontradas muitas vezes em áreas perturbadas (SONG et al., 2023). Já para o número de espécies (Species), sugere que quanto maior a riqueza do fragmento florestal, maior será a competição por recursos como luz, água e nutrientes, o que pode favorecer a infestação de lianas (CEBALLOS; MALIZIA, 2017; LI et al., 2022). Além disso, a diversidade de espécies pode proporcionar uma variedade de estruturas de suporte para lianas, aumentando a probabilidade de infestação em áreas com maior riqueza de espécies (MADRIGAL-GONZÁLEZ et al., 2018).

No período de 2012 a 2017, além da variável Lianas_2012, a variável elevação (Elevation) também foi importante. A elevação pode ter impacto no microclima e nas

condições edáficas, influenciando a distribuição e o crescimento das lianas (CLARK; HURTADO; SAATCHI, 2015). A importância dessa variável sugere que a variação topográfica dentro dos fragmentos florestais pode afetar a ocorrência de lianas devido a diferenças na disponibilidade de luz, água e nutrientes ao longo dos gradientes de altitude (KUSUMOTO; ENOKI; WATANABE, 2008; LI et al., 2022).

4.2 Desempenhos das técnicas de aprendizagem de máquina

Avaliando as estatísticas de validação das técnicas na abordagem completa do período de 10 anos (2002-2012), os maiores valores de acurácia, sensibilidade, especificidade e a maior área sob a curva (AUC) nas curvas ROC para a RNA indicam que essa técnica é a mais eficiente na discriminação correta das classes, com maior capacidade de identificar verdadeiros positivos e negativos em comparação a SVM e RF. Esses resultados corroboram com estudos que demonstram que as RNA possuem grande capacidade de aprender padrões e interações complexas entre as variáveis (LECUN; BENGIO; HINTON, 2015; SCHMIDHUBER, 2015). A maior AUC observada nas RNA indica uma melhor performance geral na classificação, estando de acordo com os resultados das estatísticas (MA et al., 2019).

Por outro lado, os menores valores das estatísticas na abordagem com modelo reduzido foram devido à redução no número de variáveis de entrada, impactando na precisão das técnicas. Essa redução foi maior nas técnicas RF e SVM, que podem ser mais sensíveis à disponibilidade e variedade de dados de entrada para otimizar suas classificações (BIAU; SCORNET, 2016; CORTES; VAPNIK, 1995). Embora a redução de variáveis simplifique o modelo e diminua a colinearidade, isso também pode remover características importantes que representam a variabilidade do fenômeno que está sendo classificado (CHAN et al., 2022). A redução da precisão das técnicas com menor número de variáveis de entrada sugere a necessidade de uma abordagem equilibrada que considere tanto a parcimônia do modelo quanto um número mínimo de entradas para evitar a perda de informações (VANDEKERCKHOVE; MATZKE; WAGENMAKERS, 2015).

Diante do exposto, a aplicação de técnicas de aprendizado de máquina na classificação de árvores infestadas por lianas é promissora. Essas técnicas permitem a análise de múltiplas variáveis complexas simultaneamente, capturando padrões não lineares e interações entre fatores ambientais e antropogênicos (MILLER;

PANNEERSELVAM; LIU, 2022). A inclusão de variáveis como altura das árvores, área basal, grupos ecológicos, elevação, dentre outros, reforça a importância de uma abordagem mais abrangente. Contudo, a seleção das variáveis deve equilibrar a parcimônia do modelo e evitar a perda de informações (PAN et al., 2023; ZHANG et al., 2023b). De modo geral, as técnicas apresentaram uma classificação precisa em nível de árvore individual, contribuindo para a compreensão e gestão da dinâmica florestal.

5 CONCLUSÃO

As técnicas de aprendizado de máquina (RNA, SVM e RF) são indicadas para classificar as árvores quanto a infestação por lianas.

O período 2012-2017 (5 anos) é indicado para classificação de árvores infestadas por lianas, em que as técnicas apresentaram performance mais equilibrada entre acurácia, sensibilidade e especificidade.

A melhor abordagem é a de modelo completo, com maiores valores no geral de acurácia, sensibilidade e especificidade em todas as técnicas comparadas à abordagem com modelo reduzido.

A melhor técnica para classificação da infestação de lianas é a RNA, que apresenta de forma geral os maiores valores de acurácia, sensibilidade e especificidade, principalmente na abordagem com modelo completo.

REFERÊNCIAS

ADDO-FORDJOUR, P.; RAHMAD, Z. B.; BURNHAM, R. J. Intercontinental Comparison of Liana Community Assemblages in Tropical Forests of Ghana and Malaysia. **Journal of Plant Ecology**, p. rtw082, 17 ago. 2016.

ALLEN, C. D.; MACALADY, A. K.; CHENCHOUNI, H.; BACHELET, D.; MCDOWELL, N.; VENNETIER, M.; KITZBERGER, T.; RIGLING, A.; BRESHEARS, D. D.; HOGG, E. H. (Ted); GONZALEZ, P.; FENSHAM, R.; ZHANG, Z.; CASTRO, J.; DEMIDOVA, N.; LIM, J.-H.; ALLARD, G.; RUNNING, S. W.; SEMERCI, A.; COBB, N. A Global Overview of Drought and Heat-Induced Tree Mortality Reveals Emerging Climate Change Risks for Forests. **Forest Ecology and Management**, v. 259, n. 4, p. 660–684, fev. 2010.

AMARO, M. A.; SOARES, C. P. B.; SOUZA, A. L. D.; LEITE, H. G.; SILVA, G. F. D. Estoque volumétrico, de biomassa e de carbono em uma floresta estacional semidecidual em viçosa, minas gerais. **Revista Árvore**, v. 37, n. 5, p. 849–857, out.

2013.

ANDRONACHE, I. Analysis of Forest Fragmentation and Connectivity Using Fractal Dimension and Succolarity. **Land**, v. 13, n. 2, p. 138, 25 jan. 2024.

BECKNELL, J. M.; VARGAS G., G.; WRIGHT, L. A.; WOODS, N.-F.; MEDVIGY, D.; POWERS, J. S. Increasing Liana Abundance and Associated Reductions in Tree Growth in Secondary Seasonally Dry Tropical Forest. **Frontiers in Forests and Global Change**, v. 5, p. 838357, 18 abr. 2022.

BERGMEIR, C.; BENÍTEZ, J. M. Neural Networks in *R* Using the Stuttgart Neural Network Simulator: **RSNNS. Journal of Statistical Software**, v. 46, n. 7, 2012. Disponível em: <<http://www.jstatsoft.org/v46/i07/>>.

BIAU, G.; SCORNET, E. A Random Forest Guided Tour. **TEST**, v. 25, n. 2, p. 197–227, 1 jun. 2016.

BOCHENEK, B.; USTRNUL, Z. Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. **Atmosphere**, v. 13, n. 2, p. 180, fev. 2022.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, p. 5–32, 2001.

BROCKERHOFF, E. G.; BARBARO, L.; CASTAGNEYROL, B.; FORRESTER, D. I.; GARDINER, B.; GONZÁLEZ-OLABARRIA, J. R.; LYVER, P. O.; MEURISSE, N.; OXBROUGH, A.; TAKI, H.; THOMPSON, I. D.; VAN DER PLAS, F.; JACTEL, H. Forest Biodiversity, Ecosystem Functioning and the Provision of Ecosystem Services. **Biodiversity and Conservation**, v. 26, n. 13, p. 3005–3035, dez. 2017.

CAI, Z.; BONGERS, F. Contrasting Nitrogen and Phosphorus Resorption Efficiencies in Trees and Lianas from a Tropical Montane Rain Forest in Xishuangbanna, South-West China. **Journal of Tropical Ecology**, v. 23, n. 1, p. 115–118, jan. 2007.

CAMPBELL, M.; LAURANCE, W. F.; MAGRACH, A. Ecological effects of lianas in fragmented forests. *Em*: SCHNITZER, S. A.; BONGERS, F.; BURNHAM, R. J.; PUTZ, F. E. **Ecology of Lianas**. 1. ed. [s.l.] Wiley, 2015. p. 443–450.

CEBALLOS, S. J.; MALIZIA, A. Liana Density Declined and Basal Area Increased over 12 y in a Subtropical Montane Forest in Argentina. **Journal of Tropical Ecology**, v. 33, n. 4, p. 241–248, jul. 2017.

CHAN, J. Y.-L.; LEOW, S. M. H.; BEA, K. T.; CHENG, W. K.; PHOONG, S. W.; HONG, Z.-W.; CHEN, Y.-L. Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. **Mathematics**, v. 10, n. 8, p. 1283, jan. 2022.

CHANDLER, C. J.; VAN DER HEIJDEN, G. M. F.; BOYD, D. S.; CUTLER, M. E. J.; COSTA, H.; NILUS, R.; FOODY, G. M. Remote Sensing Liana Infestation in an Aseasonal Tropical Forest: Addressing Mismatch in Spatial Units of Analyses. **Remote Sensing in Ecology and Conservation**, v. 7, n. 3, p. 397–410, 2021.

CHAVE, J.; RÉJOU-MÉCHAIN, M.; BÚRQUEZ, A.; CHIDUMAYO, E.; COLGAN, M. S.; DELITTI, W. B. C.; DUQUE, A.; EID, T.; FEARNSIDE, P. M.; GOODMAN, R. C.; HENRY, M.; MARTÍNEZ-YRÍZAR, A.; MUGASHA, W. A.; MULLER-LANDAU, H. C.;

MENCUCCINI, M.; NELSON, B. W.; NGOMANDA, A.; NOGUEIRA, E. M.; ORTIZ-MALAVASSI, E.; PÉLISSIER, R.; PLOTON, P.; RYAN, C. M.; SALDARRIAGA, J. G.; VIEILLEDENT, G. Improved Allometric Models to Estimate the Aboveground Biomass of Tropical Trees. **Global Change Biology**, v. 20, n. 10, p. 3177–3190, out. 2014.

CLARK, D. B.; HURTADO, J.; SAATCHI, S. S. Tropical Rain Forest Structure, Tree Growth and Dynamics along a 2700-m Elevational Transect in Costa Rica. **PLOS ONE**, v. 10, n. 4, p. e0122905, 9 abr. 2015.

CORTES, C.; VAPNIK, V. Support-Vector Networks. **Machine Learning**, v. 20, n. 3, p. 273–297, 1 set. 1995.

DIAS, A. S.; OLIVEIRA, R. S.; MARTINS, F. R.; BONGERS, F.; ANTEN, N. P. R.; STERCK, F. How do lianas and trees change their vascular strategy in seasonal versus rain forest? **Perspectives in Plant Ecology, Evolution and Systematics**, v. 40, p. 125465, 1 out. 2019.

ESQUIVEL-MUELBERT; BENNETT; SULLIVAN; BAKER; GAVISH; JOHNSON; WANG; CHAMBERS-OSTLER; GIANNICHI; GOMES; KALAMANDEEN; PATTNAYAK; FAUSET. A Spatial and Temporal Risk Assessment of the Impacts of El Niño on the Tropical Forest Carbon Cycle: Theoretical Framework, Scenarios, and Implications. **Atmosphere**, v. 10, n. 10, p. 588, 27 set. 2019.

ESRI, 2015. ArcGIS Desktop: Release 10.3.1. Environmental Systems Research Institute, Redlands, CA.

ESTRADA-VILLEGAS, S.; PEDRAZA NARVAEZ, S. S.; SANCHEZ, A.; SCHNITZER, S. A. Lianas Significantly Reduce Tree Performance and Biomass Accumulation Across Tropical Forests: A Global Meta-Analysis. **Frontiers in Forests and Global Change**, v. 4, p. 812066, 11 fev. 2022.

GHOSH, S. M.; BEHERA, M. D. Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. **Applied Geography**, v. 96, p. 29–40, 1 jul. 2018.

GIRÃO, L. C.; LOPES, A. V.; TABARELLI, M.; BRUNA, E. M. Changes in Tree Reproductive Traits Reduce Functional Diversity in a Fragmented Atlantic Forest Landscape. **PLoS ONE**, v. 2, n. 9, p. e908, 19 set. 2007.

HADDAD, N. M.; BRUDVIG, L. A.; CLOBERT, J.; DAVIES, K. F.; GONZALEZ, A.; HOLT, R. D.; LOVEJOY, T. E.; SEXTON, J. O.; AUSTIN, M. P.; COLLINS, C. D.; COOK, W. M.; DAMSCHEN, E. I.; EWERS, R. M.; FOSTER, B. L.; JENKINS, C. N.; KING, A. J.; LAURANCE, W. F.; LEVEY, D. J.; MARGULES, C. R.; MELBOURNE, B. A.; NICHOLLS, A. O.; ORROCK, J. L.; SONG, D.-X.; TOWNSHEND, J. R. Habitat Fragmentation and Its Lasting Impact on Earth's Ecosystems. **Science Advances**, v. 1, n. 2, p. e1500052, 6 mar. 2015.

HAMMOND, W. M.; WILLIAMS, A. P.; ABATZOGLOU, J. T.; ADAMS, H. D.; KLEIN, T.; LÓPEZ, R.; SÁENZ-ROMERO, C.; HARTMANN, H.; BRESHEARS, D. D.; ALLEN, C. D. Global Field Observations of Tree Die-off Reveal Hotter-Drought Fingerprint for Earth's Forests. **Nature Communications**, v. 13, n. 1, p. 1761, 5 abr. 2022.

HARRISON, D.; GUZMÁN Q., J. A.; SÁNCHEZ-AZOFEIFA, G. A. Leaf Anatomical Traits of Lianas and Trees at the Canopy of Two Contrasting Lowland Tropical Forests in the Context of Leaf Economic Spectrum. **Frontiers in Forests and Global Change**, v. 4, p. 720813, 17 ago. 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, I. **Manual técnico da vegetação brasileira**. Rio de Janeiro: IBGE, 2012.

JOLY, C. A.; RODRIGUES, R. R.; METZGER, J. P.; HADDAD, C. F. B.; VERDADE, L. M.; OLIVEIRA, M. C.; BOLZANI, V. S. Biodiversity Conservation Research, Training, and Policy in São Paulo. **Science**, v. 328, n. 5984, p. 1358–1359, 11 jun. 2010.

JONES, I. L.; PERES, C. A.; BENCHIMOL, M.; BUNNEFELD, L.; DENT, D. H. Woody Lianas Increase in Dominance and Maintain Compositional Integrity across an Amazonian Dam-Induced Fragmented Landscape. **PLOS ONE**, v. 12, n. 10, p. e0185527, 17 out. 2017.

KUSUMOTO, B.; ENOKI, T.; WATANABE, Y. Community Structure and Topographic Distribution of Lianas in a Watershed on Okinawa, South-Western Japan. **Journal of Tropical Ecology**, v. 24, n. 6, p. 675–683, nov. 2008.

LEAL, I. R.; FILGUEIRAS, B. K. C.; GOMES, J. P.; IANNUZZI, L.; ANDERSEN, A. N. Effects of Habitat Fragmentation on Ant Richness and Functional Composition in Brazilian Atlantic Forest. **Biodiversity and Conservation**, v. 21, n. 7, p. 1687–1701, jun. 2012.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep Learning. **Nature**, v. 521, n. 7553, p. 436–444, maio 2015.

LEICHT-YOUNG, S. A.; PAVLOVIC, N. B.; FROHNAPPLE, K. J.; GRUNDEL, R. Liana habitat and host preferences in northern temperate forests. **Forest Ecology and Management**, v. 260, n. 9, p. 1467–1477, 30 set. 2010.

LI, B.; ZHANG, Y.; LUAN, F.; YUAN, Z.; ALI, A.; CHU, C.; BIN, Y. Habitat Conditions and Tree Species Shape Liana Distribution in a Subtropical Forest. **Forests**, v. 13, n. 9, p. 1358, set. 2022.

LIAW, A.; WIENER, M. Classification and Regression by randomForest. v. 2, 2002.

LIU, Q.; STERCK, F. J.; ZHANG, J.-L.; SCHEIRE, A.; KONINGS, E.; CAO, M.; SHA, L.-Q.; POORTER, L. Traits, Strategies, and Niches of Liana Species in a Tropical Seasonal Rainforest. **Oecologia**, v. 196, n. 2, p. 499–514, 1 jun. 2021.

LUTZ, J. A.; VAN WAGTENDONK, J. W.; FRANKLIN, J. F. Climatic Water Deficit, Tree Species Ranges, and Climate Change in Yosemite National Park. **Journal of Biogeography**, v. 37, n. 5, p. 936–950, maio 2010.

MA, L.; LIU, Y.; ZHANG, X.; YE, Y.; YIN, G.; JOHNSON, B. A. Deep learning in remote sensing applications: A meta-analysis and review. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 152, p. 166–177, 1 jun. 2019.

MADRIGAL-GONZÁLEZ, J.; RIOS, R. S.; ARAGÓN, C. F.; GIANOLI, E. Indirect

facilitation by a liana might explain the dominance of a small tree in a temperate forest. **Journal of Plant Ecology**, v. 11, n. 4, p. 604–612, 23 maio 2018.

MARSHALL, A. R.; PLATTS, P. J.; CHAZDON, R. L.; SEKI, H.; CAMPBELL, M. J.; PHILLIPS, O. L.; GEREAU, R. E.; MARCHANT, R.; LIANG, J.; HERBOHN, J.; MALHI, Y.; PFEIFER, M. Conceptualising the Global Forest Response to Liana Proliferation. **Frontiers in Forests and Global Change**, v. 3, p. 35, 24 abr. 2020.

MASCARO, J.; ASNER, G. P.; KNAPP, D. E.; KENNEDY-BOWDOIN, T.; MARTIN, R. E.; ANDERSON, C.; HIGGINS, M.; CHADWICK, K. D. A Tale of Two “Forests”: Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. **PLOS ONE**, v. 9, n. 1, p. e85993, 28 jan. 2014.

MEDINA-VEGA, J. A.; BONGERS, F.; SCHNITZER, S. A.; STERCK, F. J. Lianas Explore the Forest Canopy More Effectively than Trees under Drier Conditions. **Functional Ecology**, v. 35, n. 2, p. 318–329, 2021.

MEUNIER, F.; VAN DER HEIJDEN, G. M. F.; SCHNITZER, S. A.; DE DEURWAERDER, H. P. T.; VERBEECK, H. Lianas Significantly Reduce Aboveground and Belowground Carbon Storage: A Virtual Removal Experiment. **Frontiers in Forests and Global Change**, v. 4, p. 663291, 22 jun. 2021.

MILLER, A.; PANNEERSELVAM, J.; LIU, L. A Review of Regression and Classification Techniques for Analysis of Common and Rare Variants and Gene-Environmental Factors. 15 jun. 2022. Disponível em: <https://figshare.le.ac.uk/articles/journal_contribution/A_review_of_regression_and_classification_techniques_for_analysis_of_common_and_rare_variants_and_gene-environmental_factors/25632696/1>.

OLIVEIRA NETO, R. R. de; LEITE, H. G.; GLERIANI, J. M.; STRIMBU, B. M. Estimation of Eucalyptus Productivity Using Efficient Artificial Neural Network. **European Journal of Forest Research**, v. 141, n. 1, p. 129–151, 1 fev. 2022.

PAN, J.; LI, W.; LIU, L.; JIA, K.; LIU, T.; CHEN, F. Variable Selection Using Deep Variational Information Bottleneck with Drop-Out-One Loss. **Applied Sciences**, v. 13, n. 5, p. 3008, jan. 2023.

POORTER, L.; VAN DER SANDE, M. T.; ARETS, E. J. M. M.; ASCARRUNZ, N.; ENQUIST, B. J.; FINEGAN, B.; LICONA, J. C.; MARTÍNEZ-RAMOS, M.; MAZZEI, L.; MEAVE, J. A.; MUÑOZ, R.; NYTCH, C. J.; DE OLIVEIRA, A. A.; PÉREZ-GARCÍA, E. A.; PRADO-JUNIOR, J.; RODRÍGUEZ-VELÁZQUES, J.; RUSCHEL, A. R.; SALGADO-NEGRET, B.; SCHIAVINI, I.; SWENSON, N. G.; TENORIO, E. A.; THOMPSON, J.; TOLEDO, M.; URIARTE, M.; HOUT, P. V. D.; ZIMMERMAN, J. K.; PEÑA-CLAROS, M. Biodiversity and Climate Determine the Functioning of Neotropical Forests. **Global Ecology and Biogeography**, v. 26, n. 12, p. 1423–1434, dez. 2017.

PORTELA, R. D. C. Q.; DIRZO, R. Forest Fragmentation and Defaunation Drive an Unusual Ecological Cascade: Predation Release, Monkey Population Outburst and Plant Demographic Collapse. **Biological Conservation**, v. 252, p. 108852, dez. 2020.

Redmond, M. D., 2019. CWD and AET function.

RIBEIRO, M. C.; METZGER, J. P.; MARTENSEN, A. C.; PONZONI, F. J.; HIROTA, M. M. The Brazilian Atlantic Forest: How Much Is Left, and How Is the Remaining Forest Distributed? Implications for Conservation. **Biological Conservation**, v. 142, n. 6, p. 1141–1153, jun. 2009.

ROCHA, S. J. S. S. D.; TORRES, C. M. M. E.; JACOVINE, L. A. G.; LEITE, H. G.; GELCER, E. M.; NEVES, K. M.; SCHETTINI, B. L. S.; VILLANOVA, P. H.; SILVA, L. F. D.; REIS, L. P.; ZANUNCIO, J. C. Artificial Neural Networks: Modeling Tree Survival and Mortality in the Atlantic Forest Biome in Brazil. **Science of The Total Environment**, v. 645, p. 655–661, dez. 2018.

ROCHA, S. J. S. S. D.; TORRES, C. M. M. E.; VILLANOVA, P. H.; TAVARES JÚNIOR, I. D. S.; RUFINO, M. P. M. X.; ROMERO, F. M. B.; JACOVINE, L. A. G.; DE MORAIS JUNIOR, V. T. M.; FRANÇA, L. C. D. J.; SCHETTINI, B. L. S.; REIS, L. P.; VIANA, Á. B. T.; ALBUQUERQUE, T. P.; VERLY, O. M.; SOARES, C. P. B.; LEITE, H. G. Machine Learning Methods: Modeling Net Growth in the Atlantic Forest of Brazil. **Ecological Informatics**, v. 81, p. 102564, jul. 2024.

SCHMIDHUBER, J. Deep learning in neural networks: An overview. **Neural Networks**, v. 61, p. 85–117, 1 jan. 2015.

SILVA, I. N., SPATTI, D. H., FLAUZIN, R. A., 2010. **Redes neurais artificiais: para engenharia e ciências aplicadas**. São Paulo: Artliber, 396 p.

Meyer, D., DIMITRIADOU, E., HORNIK, K., WEINGESSEL, A., LEISCH, F., 2018. e1071: Misc functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.

SIMINSKI, A.; ZAMBIAZI, D. C.; DOS SANTOS, K. L.; FANTINI, A. C. Dynamics of Natural Regeneration: Implications for Landscape Restoration in the Atlantic Forest, Brazil. **Frontiers in Forests and Global Change**, v. 4, p. 576908, 11 mar. 2021.

SOCIEDADE DE INVESTIGAÇÕES FLORESTAIS – SIF. **Monitoramento da flora em áreas preservadas pela CELULOSE NIPO-BRASILEIRA S.A. – CENIBRA**. Viçosa – MG: 2008. 395 p.

SONG, S.; SCHNITZER, S. A.; DING, Y.; WANG, G.; CHEN, L.; LIU, J.; WEN, R.; LUAN, F.; FANG, X.; YANG, Q.; SONG, Q. Light-Demanding Tree Species Are More Susceptible to Lianas than Shade-Tolerant Tree Species in a Subtropical Secondary Forest. **Journal of Ecology**, v. 111, n. 8, p. 1656–1669, 2023.

SOUZA, A. L. de; BOINA, A.; SOARES, C. P. B.; VITAL, B. R.; GASPAR, R. de O.; LANA, J. M. de. Estrutura fitossociológica, estoques de volume, biomassa, carbono e dióxido de carbono em Floresta Estacional Semidecidual. **Revista Árvore**, v. 36, p. 169–179, fev. 2012.

TAVARES JÚNIOR, I. da S.; TORRES, C. M. M. E.; LEITE, H. G.; CASTRO, N. L. M. de; SOARES, C. P. B.; CASTRO, R. V. O.; FARIAS, A. A. Machine Learning: Modeling Increment in Diameter of Individual Trees on Atlantic Forest Fragments. **Ecological Indicators**, v. 117, p. 106685, out. 2020.

TEIXEIRA, P. C., DONAGEMMA, G. K., FONTANA, A., TEIXEIRA, W. G. **Manual de métodos de análise de solo**. 3ª ed., Embrapa, Brasília, Brasil, 2017.

THOMPSON, L. G.; DAVIS, M. E.; MOSLEY-THOMPSON, E.; BEAUDON, E.; PORTER, S. E.; KUTUZOV, S.; LIN, P.-N.; MIKHALENKO, V. N.; MOUNTAIN, K. R. Impacts of Recent Warming and the 2015/2016 El Niño on Tropical Peruvian Ice Fields. **Journal of Geophysical Research: Atmospheres**, v. 122, n. 23, p. 12,688-12,701, 2017.

TORRES, C. M. M. E.; JACOVINE, L. A. G.; OLIVEIRA NETO, S. N. de; SOUZA, A. L. de; CAMPOS, R. A.; SCHETTINI, B. L. S. Análise Fitossociológica e Valor de Importância em Carbono para uma Floresta Estacional Semidecidual. **Floresta e Ambiente**, v. 24, p. e00099714, 17 ago. 2017.

TORRES, C. M. M. E.; MEDINA-VEGA, J. A.; ROCHA, S. J. S. S. D.; COSTA, W. D. S.; SOARES, C. P. B.; SOUZA, A. L. D.; JACOVINE, L. A. G.; LANA, J. M. D.; PEÑA-CLAROS, M. Drivers of Tree Demographic Processes in Forest Fragments of the Brazilian Atlantic Forest. **Forest Ecology and Management**, v. 534, p. 120893, abr. 2023.

URIARTE, M.; MUSCARELLA, R.; ZIMMERMAN, J. K. Environmental Heterogeneity and Biotic Interactions Mediate Climate Impacts on Tropical Forest Regeneration. **Global Change Biology**, v. 24, n. 2, fev. 2018. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1111/gcb.14000>>.

VAN DER HEIJDEN, G. M. F.; SCHNITZER, S. A.; MEUNIER, F. Editorial: Lianas, Ecosystems, and Global Change. **Frontiers in Forests and Global Change**, v. 6, p. 1079620, 2 fev. 2023.

VANDEKERCKHOVE, J.; MATZKE, D.; WAGENMAKERS, E.-J. Model Comparison and the Principle of Parsimony. *Em*: BUSEMEYER, J. R.; WANG, Z.; TOWNSEND, J. T.; EIDELS, A. **The Oxford Handbook of Computational and Mathematical Psychology**. [s.l.] Oxford University Press, 2015. p. 0.

VOGADO, N. O.; ENGERT, J. E.; LINDE, T. L.; CAMPBELL, M. J.; LAURANCE, W. F.; LIDDELL, M. J. Climate Change Affects Reproductive Phenology in Lianas of Australia's Wet Tropics. **Frontiers in Forests and Global Change**, v. 5, p. 787950, 9 jun. 2022.

WONGCHAI, W.; ONSREE, T.; SUKKAM, N.; PROMWUNGKWA, A.; TIPPAYAWONG, N. Machine Learning Models for Estimating above Ground Biomass of Fast Growing Trees. **Expert Systems with Applications**, v. 199, p. 117186, ago. 2022.

YUAN, X.; LAAKSO, K.; MARZAHN, P.; SANCHEZ-AZOFEIFA, G. A. Canopy Temperature Differences between Liana-Infested and Non-Liana Infested Areas in a Neotropical Dry Forest. **Forests**, v. 10, n. 10, p. 890, 9 out. 2019.

ZANNE, A. E.; LOPEZ-GONZALEZ, G.; COOMES, D. A.; ILIC, J.; JANSEN, S.; LEWIS, S. L.; MILLER, R. B.; SWENSON, N. G.; WIEMANN, M. C.; CHAVE, J. **Data from: Towards a worldwide wood economics spectrum** Dryad, 4 fev. 2009.

Disponível em: <<https://datadryad.org/stash/dataset/doi:10.5061/dryad.234>>.

ZHANG, J.; QU, X.; SONG, X.; XIAO, Y.; WANG, A.; LI, D. Spatial Variation in Soil Base Saturation and Exchangeable Cations in Tropical and Subtropical China. **Agronomy**, v. 13, n. 3, p. 781, mar. 2023a.

ZHANG, X.; XUE, J.; XIAO, Y.; SHI, Z.; CHEN, S. Towards Optimal Variable Selection Methods for Soil Property Prediction Using a Regional Soil Vis-NIR Spectral Library. **Remote Sensing**, v. 15, n. 2, p. 465, jan. 2023b.

Capítulo II: Modelagem da abundância de árvores infestadas por lianas em fragmentos de Mata Atlântica utilizando aprendizagem de máquina

RESUMO

A intensa fragmentação florestal da Mata Atlântica devido as atividades humanas como desmatamento, agricultura e urbanização tem levado à perda significativa de habitat e à alteração na dinâmica sucessional dessas florestas. Nesse contexto, as lianas desempenham um papel importante, influenciando a estrutura, composição e funcionamento dos ecossistemas florestais. O objetivo deste capítulo foi projetar a abundância de árvores infestadas por lianas em fragmentos da Mata Atlântica utilizando técnicas de aprendizagem de máquina. Três fragmentos florestais da Mata Atlântica em Minas Gerais, Brasil, foram analisados, estando localizadas nos municípios de Guanhães (Cachoeira das Pombas), Ipaba (Ipaba) e Coronel Fabriciano (São José). Para projetar a abundância de árvores infestadas por lianas, atributos florestais como área basal por hectare, número de fustes por hectare, número de espécies e o número de árvores infestadas por lianas por hectare no ano inicial foram considerados. As variáveis antropogênicas incluíram o histórico e mudanças no uso da terra. As técnicas de aprendizagem de máquina aplicadas foram redes neurais artificiais (RNA), regressão vetorial de suporte (RVS) e *random forest* (RF). O desempenho das técnicas foi avaliado pela correlação entre os valores projetados e observados de ind. ha⁻¹, raiz do erro quadrático médio (RMSE%), erro médio percentual (MAPE) e dispersão dos resíduos padronizados em função dos valores observados. No período 2002-2012, a área basal por hectare apresentou correlação negativa com a abundância de árvores infestadas por lianas ($r = -0.22$). Para o período 2002-2017, a correlação foi negativa ($r = -0.04$), assim como a correlação da matéria orgânica a 20-40 cm de profundidade ($r = -0.4$). A precipitação dos três meses mais secos teve correlação negativa ($r = -0.42$) no período 2002-2017 e positiva ($r = 0.45$) no período 2012-2017. Os resultados indicaram uma redução significativa na quantidade de ind. ha⁻¹ entre 2002 e 2012, seguida por um aumento entre 2012 e 2017, devido ao fenômeno climático El Niño de 2015/2016. A modelagem com RNA foi a mais eficaz, seguida pela RVS, enquanto a RF apresentou tendência de subestimação dos maiores valores de ind. ha⁻¹. De forma geral, a abordagem com modelo completo apresentou as melhores estatísticas nos períodos de 2002-2012 e 2002-2017, enquanto a abordagem reduzida teve melhor desempenho no período de 2012-2017. Conclui-se que variáveis como a área basal, matéria orgânica e precipitação são importantes para entender a dinâmica das lianas. As técnicas de aprendizagem de máquina, especialmente RNA, foram eficazes na modelagem da abundância de árvores infestadas por lianas, sendo ferramentas valiosas para a conservação e manejo sustentável da Mata Atlântica.

Palavras-chave: Floresta tropical; Infestação; Modelagem de lianas.

1 INTRODUÇÃO

A Mata Atlântica é um bioma de extrema importância para a manutenção da biodiversidade, abrigando uma vasta gama de espécies endêmicas de fauna e flora (JOLY et al., 2010). No entanto, este bioma enfrenta intensa pressão devido à expansão urbana, agrícola e industrial, resultando em significativa perda de florestas (HADDAD et al., 2015). A fragmentação da Mata Atlântica não só reduz a extensão dos habitats naturais, mas também altera a dinâmica sucessional florestal e pode influenciar na abundância de árvores infestadas por lianas (CAMPBELL; LAURANCE; MAGRACH, 2015; RIBEIRO et al., 2009).

As lianas, conhecidas também como trepadeiras lenhosas, são plantas que competem com as árvores por luz, água e nutrientes, afetando negativamente o crescimento e a sobrevivência das árvores, impactando a estrutura e composição da floresta (REIS et al., 2020; MEUNIER et al., 2021; RODRÍGUEZ et al., 2021). Fatores ambientais como clima, solo e topografia influenciam a abundância de árvores infestadas por lianas, especialmente no contexto das mudanças climáticas (ADDO-FORDJOUR; RAHMAD, 2015). Alterações nos padrões de temperatura e precipitação podem modificar a distribuição e abundância dessas plantas, enquanto a degradação do solo e a topografia influenciam a disponibilidade de nutrientes e a estrutura da floresta (LIU et al., 2021). Ainda, atividades humanas, como o desmatamento e práticas agrícolas podem favorecer a proliferação de lianas (NGUTE et al., 2024).

Desse modo, prever a abundância de árvores infestadas por lianas é um desafio complexo devido a heterogeneidade ambiental e a interação de múltiplos fatores que introduzem níveis de incerteza na modelagem da abundância de árvores infestadas por lianas (BAI et al., 2020; LI et al., 2022). Para contornar esse desafio, uma alternativa é a utilização de técnicas de aprendizagem de máquina (AM) que podem integrar múltiplas variáveis ambientais e antropogênicas, permitindo a criação de modelos mais acurados (ROCHA et al., 2024).

As técnicas de AM estão sendo cada vez mais utilizadas na predição de diversos atributos das florestas tropicais, como na estimativa da biomassa de árvores (SINGH et al., 2022), estimativa de atributos do solo (FERREIRA et al., 2022) e reconhecimento de espécies ameaçadas e invasoras na Mata Atlântica (CRISIGIOVANNI et al., 2021), contudo, o uso na modelagem da abundância de árvores infestadas por lianas ainda não foi abordado. Assim, introduzimos de forma

inédita a aplicação de redes neurais artificiais (RNA), regressão vetorial de suporte (RVS) e *random forest* (RF) na modelagem do número de indivíduos infestados por lianas. Desse modo, o objetivo deste capítulo foi projetar a abundância de indivíduos arbóreos infestados por lianas (ind. ha⁻¹) em fragmentos de Mata Atlântica utilizando técnicas de aprendizagem de máquina.

2 MATERIAL E MÉTODOS

2.1 Caracterização das áreas de estudo

Os fragmentos avaliados neste capítulo foram os mesmos do Capítulo I, estando localizados nos municípios de Guanhões (Cachoeira das Pombas), Ipaba (Ipaba) e Coronel Fabriciano (São José). As características dos fragmentos podem ser consultadas seção 2.1 do Capítulo I.

2.2 Coleta dos dados

As coletas dos dados em campo foram realizadas em parcelas permanentes retangulares de área fixa de 10 × 50 m (0,05 ha). Cada parcela foi subdividida em 20 subparcelas de 5 × 5 m, e as árvores com circunferência a 1,30 m de altura do solo (CAP; cm) ≥ 15 cm foram mapeadas. O nome científico das espécies, a CAP, a altura total – Ht (m) e o grau de infestação de lianas em árvores com CAP ≥ 15 cm foram registrados nos anos de 2002, 2012 e 2017, em cada fragmento. Posteriormente, a CAP foi convertida em diâmetro a 1,30 m de altura (DAP; cm). A classificação da infestação de lianas nas árvores foi feita da seguinte forma: 1 - Nenhuma liana; 2 - Lianas apenas no tronco; 3 - Lianas apenas na copa; e 4 - Lianas no tronco e na copa.

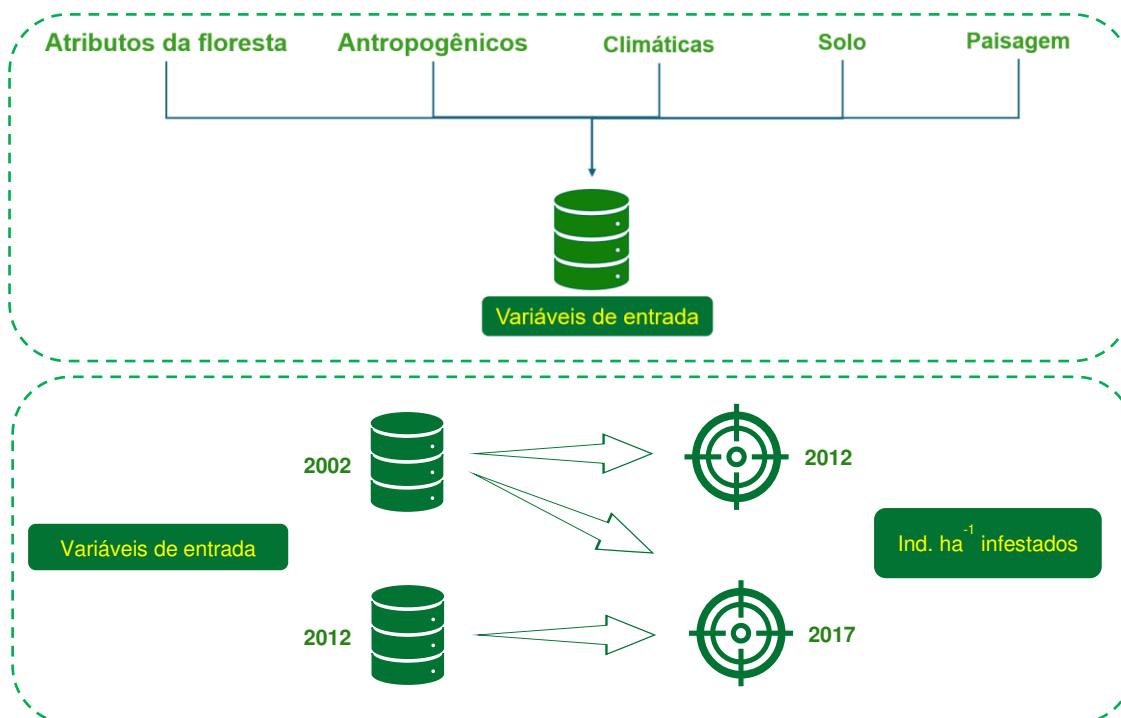
2.3 Projeção do número de indivíduos infestados por hectare

A abundância de indivíduos arbóreos infestados por lianas (Ind. ha⁻¹) foi determinada para cada parcela de todos os fragmentos (Cachoeira das Pombas, Ipaba e São José). Para isso, as árvores que possuíam a classe de infestação igual ou superior a 2, registradas em campo, foram filtradas e contadas em cada parcela dos fragmentos. Posteriormente, os valores de cada parcela foram extrapolados para

hectare, resultando no número de ind. ha⁻¹. Esse processo foi realizado para os anos de 2012 e 2017, que foram os anos em que a abundância foi projetada.

Dessa forma, os valores de ind. ha⁻¹ calculados para os anos de 2012 e 2017 foram projetados a partir das variáveis explicativas referentes ao ano de 2002 (Figura 1). Ainda, as variáveis explicativas referentes ao ano de 2012 também foram utilizadas para projetar os ind. ha⁻¹ infestados para o ano de 2017. Os dados dos três fragmentos foram unidos e formaram uma tabela única contendo todas as informações dos três fragmentos.

Figura 1 - Esquema da projeção de árvores infestadas por hectare (ind. ha⁻¹) para os anos 2012 e 2017 utilizando variáveis de entrada do ano 2002 ou 2012.



Fonte: O autor (2024).

2.3.1 Variáveis explicativas para projeção de árvores infestadas por hectare

Para projetar o número de árvores infestadas por hectare, as variáveis topográficas, de solo, climáticas, atributos da floresta e antropogênicas foram testadas como variáveis explicativas.

2.3.1.1 Variáveis climáticas

A precipitação anual, o número de meses com menos de 100 mm de chuva, a precipitação nos três meses mais secos e a temperatura média anual foram obtidos da estação climatológica mais próxima de cada fragmento (ROCHA et al., 2020). Posteriormente, o déficit hídrico climático (CWD) foi estimado – um balanço hídrico entre precipitação e evapotranspiração – como uma aproximação para condições de seca, seguindo a metodologia de Lutz et al. (2010) e utilizando a função R CWD e AET (evapotranspiração real) de Redmond (2019). O CWD reflete as condições de seca com mais precisão do que a precipitação total anual (CHAVE et al., 2014). Valores de CWD mais negativos indicam alto estresse hídrico, enquanto valores próximos a zero indicam ausência de estresse hídrico (POORTER et al., 2017). A média da precipitação anual total e do CWD para um, dois, três e quatro anos anteriores ao ano de medição também foi calculada (TORRES et al., 2023).

2.3.1.2 Variáveis do solo e da paisagem

Cerca de 20 a 30 amostras de solo foram coletadas por parcela para formar uma amostra composta. Informações para as profundidades de 0-20 e 20-40 cm foram obtidas. Algumas das variáveis analisadas incluíram: pH do solo em H₂O, cátions trocáveis (Ca²⁺, Mg²⁺ e Al³⁺), acidez total (H⁺ + Al³⁺), capacidade de troca catiônica (CTC), saturação por bases (V), fósforo disponível (P), P remanescente na solução (P-rem) e matéria orgânica do solo (MO), determinadas utilizando métodos padrão (TEIXEIRA et al., 2017).

As análises de solo foram realizadas no laboratório de Fertilidade do Solo da Universidade Federal de Viçosa (UFV), Minas Gerais, Brasil. As variáveis de paisagem incluíram a distância da borda, elevação, inclinação e o nome dos fragmentos. A elevação e inclinação foram calculadas para cada parcela usando *Spatial Analyst Tools* de superfície no ArcGIS® 10.3.1 (ESRI, 2015). O *Shuttle Radar Topography Mission* (SRTM) e Modelos Digitais de Elevação (DEMs) foram utilizados para a análise (TORRES et al., 2023). A distância até a borda mais próxima da floresta foi calculada com a ferramenta "*near*" no ArcGIS® 10.3.1 (ESRI, 2015).

2.3.1.3 Atributos da floresta e variáveis antropogênicas

Os atributos da floresta considerados foram a área basal por hectare, número

de fustes por hectare em cada parcela dos fragmentos, número de espécies, DAP, Ht das árvores e a classe de lianas registrada em campo no ano a_1 . As variáveis antropogênicas utilizadas incluíram o histórico de uso da terra e mudança no uso da terra. Com base em dados coletados de fotografias aéreas (aproximadamente de 1960, 1980 e 1990), imagens de satélite Landsat de 1985, entrevistas com proprietários e registros de propriedade, foi determinado o histórico de uso da terra das áreas. As categorias de uso da terra consideradas foram: desmatamento, produção agrícola, plantio de eucalipto e extração seletiva de madeira (TORRES et al., 2023).

Nas Tabelas B1 e B2 as estatísticas descritivas de todas as variáveis explicativas dos anos de 2002 e 2012 estão apresentadas.

2.3.2 Seleção das variáveis explicativas para projeção de indivíduos infestados por lianas por hectare

A seleção das variáveis explicativas foi através de duas abordagens chamadas de modelo completo e modelo reduzido. No modelo completo, as variáveis que não possuíam correlação linear maior que 0,8 com outras variáveis explicativas foram selecionadas. Já no modelo reduzido, as variáveis explicativas mais importantes foram selecionadas com base no algoritmo *Random Forest* (RF).

Por meio da técnica RF, as variáveis mais importantes são determinadas através do quanto o aumento percentual do erro quadrático médio é alterado quando uma variável preditora é permutada aleatoriamente. Assim, quanto maior esse erro, mais importante uma determinada variável é em relação as demais.

A função utilizada nessa etapa foi a *VSURF* do pacote de mesmo nome no ambiente R (R CORE TEAM, 2022). Após a aplicação da *Random Forest*, as variáveis que possuírem maior importância para projeção das ind. ha⁻¹ foram selecionadas para o treinamento das redes neurais artificiais (RNA), regressão por vetores de suporte (RVS) e *random forest* (RF).

A colinearidade de entre as variáveis explicativas pode não afetar severamente os métodos não paramétricos, contudo, a exclusão dessas variáveis favorece a escolha de modelos mais parcimoniosos e sem redundância.

2.4 Pré-processamento dos dados

As variáveis quantitativas foram escalonadas, sendo que a amplitude e magnitude das variáveis de entrada e saída ficaram entre o intervalo de 0 a 1. Essa transformação possibilitou acelerar a taxa de convergência e reduzir o processo de iteração no treinamento (Equação 1).

$$v' = \frac{v - \min(v)}{\max(v) - \min(v)} (\max_{\text{novo}} - \min_{\text{novo}}) + \min_{\text{novo}} \quad (1)$$

Em que: $\min(v)$: menor valor da variável no banco de dados de treino; $\max(v)$: maior valor da variável no banco de dados treino; \max_{novo} : maior valor transformado; \min_{novo} : menor valor transformado; v' : valor da observação transformada; e v : valor da observação não transformada.

Após a modelagem com cada técnica, os valores projetados foram retransformados para a escala original, possibilitando o cálculo das estatísticas de qualidade do ajuste.

2.5 Treinamento das redes neurais artificiais, regressões vetoriais de suporte e *random forests*

As redes neurais artificiais (RNA) utilizadas foram do tipo *Perceptron* de múltiplas camadas, também conhecidas como *Multilayer Perceptron* (MLP), com uma única camada oculta (HAYKIN, 2009). O intervalo do número de neurônios nesta camada foi estabelecido pelo método de Fletcher-Gloss (SILVA et al., 2010): $2 \times n^{0,5} + n_2 \leq n_1 \leq 2 \times n + 1$; em que n representa o número de entradas da rede; n_1 é a quantidade de neurônios na camada oculta; e n_2 é o número de neurônios na camada de saída. As funções de ativação avaliadas incluíram a exponencial, logística e tangente hiperbólica. Os algoritmos de treinamento examinados foram o *Resilient Propagation* (Rprop) e o *Scaled Conjugate Gradient* (SCG). Os pesos iniciais das RNAs foram gerados de maneira aleatória, sendo posteriormente ajustados com base na função do erro. O número máximo de ciclos foi estabelecido em 1000. As RNA foram treinadas utilizando a função *mlp* do pacote “*RSNNS*” no programa R (BERGMEIR; BENITEZ, 2012).

Nos treinamentos das regressões vetoriais de suporte (RVS) foram utilizadas quatro configurações, representadas por quatro funções kernel, em que a seguinte função de erro foi otimizada:

$$\text{minimize } \frac{1}{2} \|w\|^2 + \frac{C}{m} \sum_{i=1}^m (\xi_i + \xi_i^*)$$

$$\text{sujeito a } (\langle \Phi(x_i), w \rangle + b) - y_i \leq \epsilon - \xi_i^*$$

$$y_i - (\langle \Phi(x_i), w \rangle + b) \leq \epsilon - \xi_i$$

$$\xi_i \geq 0 \quad (i = 1, \dots, m)$$

em que: w = vetor de coeficientes; C = parâmetro de penalidade do erro; ξ_i, ξ_i^* = variáveis de folga que caracterizam o erro acima e abaixo do ϵ -tubo, respectivamente; i = casos de treinamento; m = número total de casos de treinamento; $\Phi(x_i)$ = kernel utilizado; b = bias; y_i = dados de saída e; ϵ = erro máximo admitido.

As funções kernel avaliadas foram as do tipo linear, polinomial, base radial e sigmóide (Tabela 1).

Tabela 1 - Funções kernel testadas nos treinamentos das regressões vetoriais de suporte

Tipo de kernel	Função
Linear	$u' \times v$
Polinomial	$(\gamma \times u' \times v + coef0)^d$
Base radial	$e^{(-\gamma \times u-v ^2)}$
Sigmoidal	$\tanh(\gamma \times u' \times v + coef0)$

Em que: γ = parâmetro de forma necessário para todos os kernels, exceto linear [padrão: $1/(\text{dimensão de dados})$]; d = grau do polinômio; $coef0$ = parâmetro de penalidade do erro; \tanh = tangente hiperbólica.

A função utilizada para os treinamentos das RVS foi a *svm* do pacote “*e1071*” no R (MEYER et al., 2018).

No treinamento das *random forests* (RF), três parâmetros foram configurados (BREIMAN, 2001): o número de árvores de regressão aleatória (*n tree*), no qual foram testadas 500, 1000, 1500, 2000, 2500 e 3000 árvores de regressão; o número de variáveis de divisão (*m try*), utilizado para determinar o número de variáveis disponíveis para dividir em cada nó da árvore, sendo que o número padrão de 1/3 das

variáveis independentes foi utilizado; e o tamanho mínimo de nós terminais (*nodesize*), cujo valor foi 5. A função *randomForest* do pacote “*randomForest*” no *R* foi utilizada para os treinamentos das RF (LIAW; WIENER, 2002).

2.6 Avaliação das técnicas de aprendizagem de máquina

A avaliação de cada técnica foi com base na validação cruzada do tipo *leave-one-out* (CAWLEY; TALBOT, 2003). Dessa forma, a cada iteração uma parcela foi destinada para validação, enquanto as demais parcelas foram utilizadas no treinamento até completar a validação com todas as parcelas.

Os valores projetados na validação de cada parcela foram compilados e associados a cada configuração. Após isso, esses valores foram comparados com os observados por meio dos seguintes critérios para verificar a qualidade do ajuste: correlação entre os valores observados e projetados – r_{yy} (Equação 2); raiz quadrada do erro quadrático médio, em porcentagem – *RMSE*(%) (Equação 3); erro absoluto médio percentual – *MAPE* (Equação 4) e; gráficos de dispersão dos valores observados e projetados; e gráficos de dispersão dos resíduos padronizados – *SR* (Equação 5).

$$r_{yy} = \frac{\text{cov}(y, \hat{y})}{\sqrt{s^2(y) \times s^2(\hat{y})}} \quad (2)$$

$$\text{RMSE}(\%) = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \times \frac{100}{\bar{y}} \quad (3)$$

$$\text{MAPE} = \frac{1}{n} \times \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \quad (4)$$

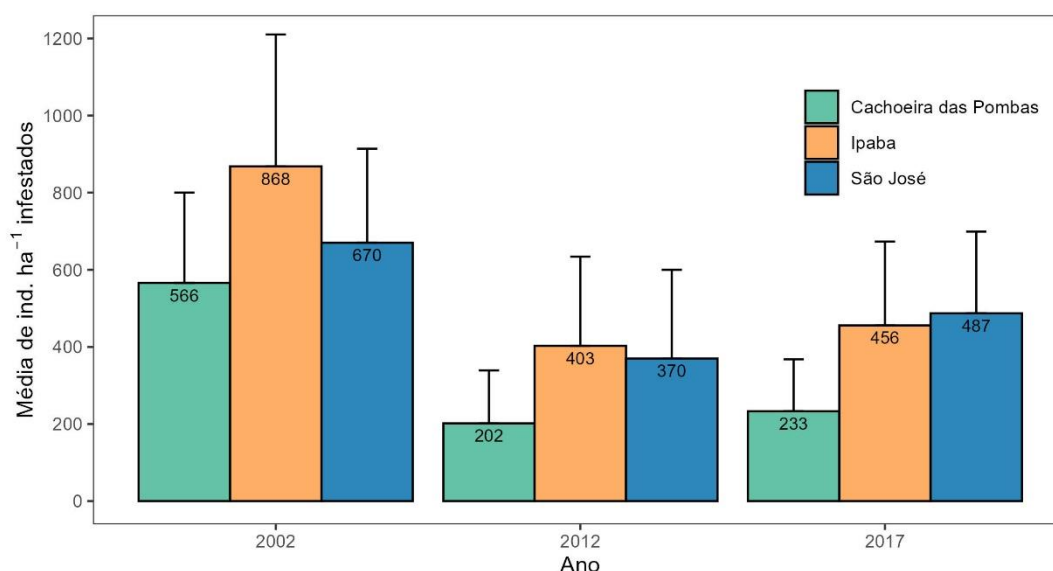
$$\text{SR} = \frac{\hat{y}_i - y_i}{\sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-1}}} \quad (5)$$

Em que: s^2 = variância; cov = covariância; y_i = valor observado da i -ésima parcela, em ind. ha^{-1} ; e \hat{y}_i = valor previsto da i -ésima parcela, em ind. ha^{-1} ; \bar{y} = média aritmética dos valores observados; n = número total de observações; e_i = diferença entre o valor observado e previsto da i -ésima parcela; \bar{e} = diferença média aritmética entre o valor observado e projetado.

3 RESULTADOS

Analisando a média de ind. ha⁻¹ infestados por lianas nos três fragmentos, a quantidade de ind. ha⁻¹ no ano de 2002 para 2012 reduziu 64, 51 e 45% nos fragmentos Cachoeira das Pombas, Ipaba e São José, respectivamente (Figura 2). Por outro lado, do ano de 2012 para 2017 a quantidade de ind. ha⁻¹ infestados aumentou 15, 8 e 33% em Cachoeira das Pombas, Ipaba e São José, respectivamente.

Figura 2 - Média de ind. ha⁻¹ por fragmento nos anos de 2002, 2012 e 2017.



Fonte: O autor (2024).

3.1 Variáveis selecionadas

Na seleção das variáveis através da abordagem do modelo completo, 19 variáveis foram selecionadas para projetar o número de ind. ha⁻¹ nos três períodos (Tabela 2). Na abordagem com modelo reduzido, 3, 4 e 3 variáveis foram selecionadas para projetar o número de ind. ha⁻¹ nos períodos 2002-2012, 2002-2017 e 2012-2017, respectivamente. As correlações das variáveis quantitativas selecionadas em cada período para a abordagem com modelo completo podem ser observadas nas Figuras B1 e B2 (APÊNDICE B). As etapas de seleção das variáveis mais importantes pela RF na abordagem com modelo reduzido estão apresentadas nas Figuras B3 a B5.

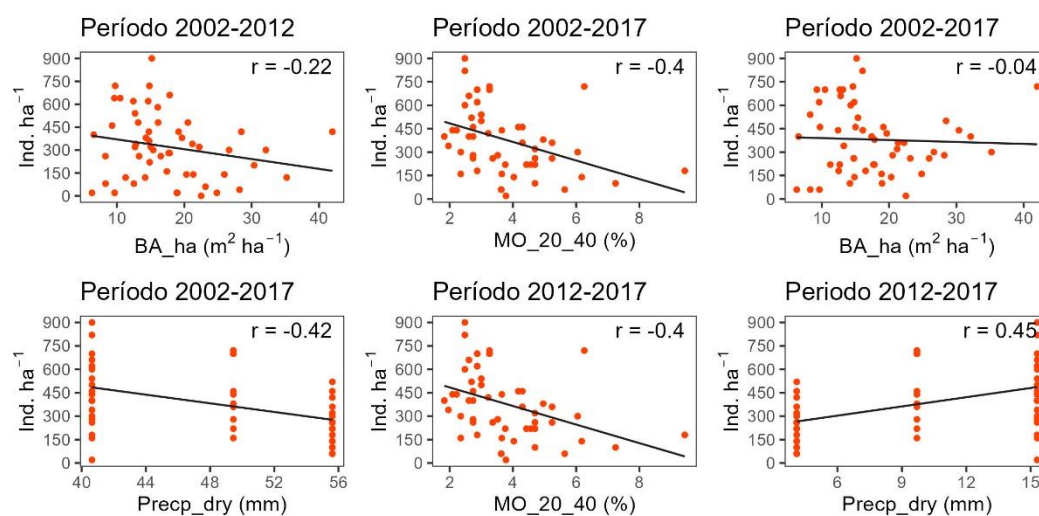
Tabela 2 - Variáveis selecionadas por abordagem (modelo completo e reduzido) para os três períodos avaliados

Abordagem	Período	Variáveis
Completo	2002-2012; 2002-2017; 2012-2017	Al3_20_40
		BA_ha
		Changes
		CWD
		Declivity_perc
		Edge_distance
		Fragmentos
		Ind_ha_02
		LUH
		m_20_40
		MO_20_40
		P_20_40
		pH_H2O_20_40
		Precp_avg
		Precp_dry
		Slope_Angle
		Species
Stems_ha		
t_20_40		
Reduzido	2002-2012	Ind_ha_02
		BA_ha
	2002-2017	Fragmentos
		MO_20_40
		Ind_ha_02
		BA_ha
	2012-2017	Precp_dry
		Ind_ha_12
2012-2017	MO_20_40	
	Precp_dry	

iEm que: Al3_20_40 = Alumínio trocável (cmolc/dm³) 20cm - 40 cm; Changes = Mudanças no uso da terra; CWD = Déficit hídrico climático (ano de medição); Declivity_perc = Declividade (%); Edge_distance = Distância da borda; LUH = Histórico de uso da terra; m_20_40 = Saturação por alumínio (%) de 20 cm - 40 cm; MO_20_40 = Matéria orgânica (dag/kg) de 20 cm - 40 cm; P_20_40 = Fósforo (cmolc/dm³) de 20 cm - 40 cm; pH_H2O_20_40 = pH em água de 20 cm - 40 cm; Precp_avg = Precipitação média anual (de 1989 até 1 ano antes da medição); Precp_dry = Precipitação total dos três meses mais secos no ano de medição; Slope_Angle = Ângulo de inclinação (°); Species = Número de espécies; t_20_40 = CTC efetiva (cmolc/dm³) de 20 cm - 40 cm; Ind_ha_2002 = número de indivíduos infestados por lianas por hectare no ano de 2002; Ind_ha_12 = número de indivíduos infestados por lianas por hectare no ano de 2012; GE = Grupo ecológico; P_Rem_0_20 = Fósforo remanescente (mg/L) de 0 cm - 20 cm; Stems_ha = número de fustes por hectare; BA_ha = área basal por hectare.

No período 2002-2012, a área basal por hectare (BA_ha) apresentou correlação negativa com as ind. ha⁻¹ ($r = -0.22$) (Figura 3). Para o período 2002-2017, a correlação entre BA_ha e as ind. ha⁻¹ também foi negativa, mas com redução do valor ($r = -0.04$). A correlação da matéria orgânica de 20-40 cm de profundidade (MO_20_40) durante este período foi negativa ($r = -0.4$). A precipitação total dos três meses mais secos (Prec_dry) apresentou correlação negativa com as ind. ha⁻¹ ($r = -0.42$). No período 2012-2017, a MO_20_40 manteve uma correlação negativa com a abundância de indivíduos infestados por lianas ($r = -0.4$). Já a relação entre Prec_dry e a abundância de indivíduos infestados por lianas apresentou uma correlação positiva ($r = 0.45$).

Figura 3 - Relação da abundância de indivíduos arbóreos infestados (Ind. ha⁻¹) com as variáveis quantitativas selecionadas na abordagem de modelo reduzido em todos os períodos. Em que: BA_ha = área basal por hectare (m² ha⁻¹); MO_20_40 = matéria orgânica de 20-40 cm de profundidade (%); Prec_dry = Precipitação total dos três meses mais secos no ano de medição (mm).



Fonte: O autor (2024).

3.2 Avaliação das técnicas de aprendizagem de máquina na projeção de ind. ha⁻¹ infestados por lianas

Avaliando as estatísticas de qualidade do ajuste no treinamento para todos os períodos e abordagens, a rede neural artificial (RNA) apresentou os maiores valores de correlação entre os valores projetados e observados de ind. ha⁻¹ (entre 0,86 e 0,99) e os menores valores de RMSE% (entre 0,88 e 34%) e MAPE (entre 0,54 e 53%) (Tabela 3). A regressão vetorial de suporte (RVS) apresentou menores correlações

(entre 0,79 e 0,91), e maiores RMSE% (entre 25 e 35%) e MAPE (entre 20 e 69%) do que a RNA em todos os períodos e abordagens. Já a *random forest* (RF) apresentou os menores valores de correlação (entre 0,58 e 0,79) e maiores RMSE% (entre 33 e 55%) e MAPE (entre 40 e 168%) em relação as outras técnicas para todos os períodos e abordagens.

Tabela 3 - Estatísticas de correlação (r_{yy}), raiz quadrada do erro quadrático médio, em porcentagem (RMSE%) e erro absoluto médio percentual (MAPE) das melhores configurações de cada técnica no treinamento nos períodos e abordagens avaliados

Período	Abordagem	Modelo	r_{yy}	RMSE%	MAPE	Kernel ¹	Algoritmo ²	Função ³	Neurônios ⁴	Árvores ⁵
2002-2012	Completo	RVS	0,8948	31,92	92,58	Polinomial				3000
		RF	0,5861	55,10	168,08					
		RNA	0,9999	0,98	0,94					
	Reduzido	RVS	0,7920	35,94	69,21	Radial				1500
		RF	0,7142	48,13	138,85					
		RNA	0,8619	34,20	53,45					
2002-2017	Completo	RVS	0,8993	26,09	22,24	Radial				2000
		RF	0,7259	41,54	76,74					
		RNA	0,9998	1,19	1,68					
	Reduzido	RVS	0,8754	31,10	44,19	Radial				3000
		RF	0,7760	34,99	47,00					
		RNA	0,9328	21,22	20,91					
2012-2017	Completo	RVS	0,9129	25,56	20,87	Radial				500
		RF	0,7230	39,15	47,03					
		RNA	0,9999	0,88	0,54					
	Reduzido	RVS	0,8562	29,50	29,52	Radial				500
		RF	0,7995	33,18	40,82					
		RNA	0,9036	24,22	26,50					

Em que: Rprop = *resilient propagation*; SCG = *scaled conjugate gradient*; RNA = rede neural artificial; RVS = regressão por vetores de suporte; RF = *random forest*. ¹ Função Kernel utilizada nas RVS. ² Algoritmo na camada oculta da RNA. ³ Função ativação na camada oculta da RNA. ⁴ Número de neurônios na camada oculta da RNA. ⁵ Número de árvores de regressão na RF.

Avaliando as estatísticas de qualidade do ajuste na validação para todos os períodos e abordagens, a RNA apresentou os maiores valores de correlação (entre 0,69 e 0,79) e menores valores de RMSE% (entre 0,69 e 0,79) e MAPE (entre 0,69 e 0,79), com exceção da abordagem com modelo reduzido no período 2002-2017 e completo no período 2012-2017 (Tabela 4). Na abordagem com modelo reduzido do período 2002-2017, a RF apresentou o maior valor de correlação (0,72) e menor RMSE% (39%) em relação as outras técnicas. Já na abordagem com modelo completo do período 2012-2017, a RVS apresentou a maior correlação (0,77) e menor RMSE% (39%) e MAPE (44%).

Comparando as abordagens de modelo completo com modelo reduzido nos períodos 2002-2012 e 2002-2017, a abordagem com modelo reduzido apresentou menores valores de correlação, com exceção da técnica RF. Por outro lado, no período de 2012-2017 a tendência foi de aumento da precisão na abordagem com modelo reduzido, com maiores correlações e menores RMSE% e MAPE.

Avaliando a dispersão dos valores observados e projetados de indivíduos infestados por lianas por hectare e os resíduos padronizados para o período de 2002-2012 nos modelos completos e reduzidos, a RNA e RVS confirmaram a maior precisão, com maior proximidade dos pontos à linha de 1:1 e do eixo zero dos resíduos padronizados (Figura 4). A técnica RF apresentou tendência de subestimação dos maiores valores de ind. ha⁻¹, tanto na abordagem com modelo completo quanto reduzido.

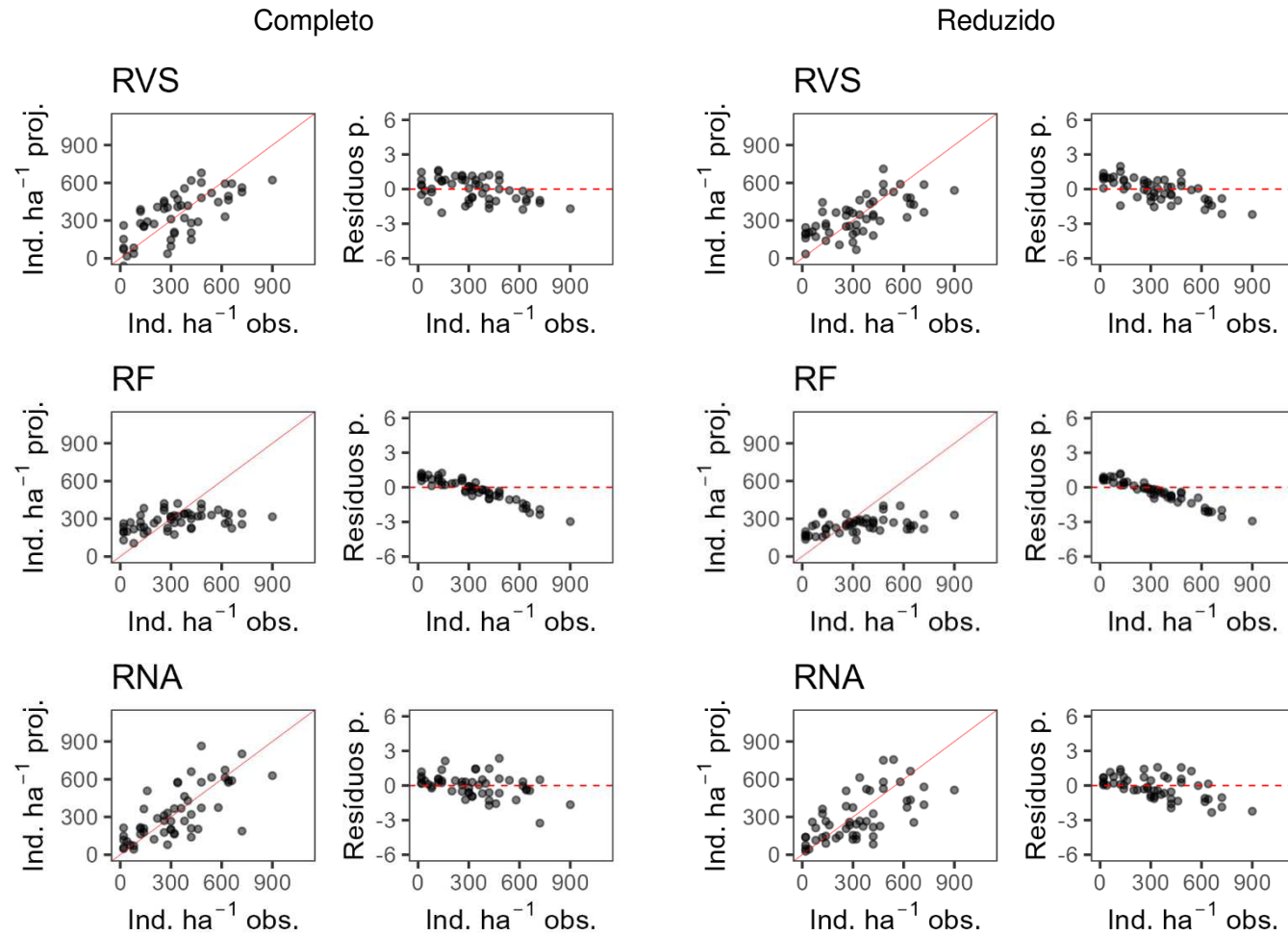
Nos períodos 2002-2017 e 2002-2017 a tendência foi a mesma de 2002-2012, com melhor distribuição dos valores observados e projetados para as RNA e RVS (Figuras 5 e 6). A RF continuou com a tendência de subestimação nos maiores valores de ind. ha⁻¹, mesmo na abordagem com modelo reduzido em 2002-2017, em que obteve as melhores estatísticas de qualidade do ajuste.

Tabela 4 - Estatísticas de correlação (r_{yy}), raiz quadrada do erro quadrático médio, em porcentagem (RMSE%) e erro absoluto médio percentual (MAPE) das melhores configurações de cada técnica na validação nos períodos e abordagens avaliados

Período	Abordagem	Modelo	r_{yy}	RMSE%	MAPE	Kernel	Algoritmo	Função	Neurônios	Árvores
2002-2012	Completo	RVS	0,6993	49,56	106,68	Linear	Rprop	Tangente Hiperbólica	55	1500
		RF	0,4191	60,85	140,25					
		RNA	0,7522	46,61	63,45					
	Reduzido	RVS	0,6384	50,40	107,64	Linear	SCG	Tangente Hiperbólica	5	1000
		RF	0,4569	63,65	121,62					
		RNA	0,6942	47,67	92,90					
2002-2017	Completo	RVS	0,6480	45,79	65,45	Linear	Rprop	Logística	21	500
		RF	0,6203	46,58	81,52					
		RNA	0,7151	43,74	51,77					
	Reduzido	RVS	0,6290	44,24	76,18	Radial	SCG	Exponencial	6	2500
		RF	0,7294	39,85	74,56					
		RNA	0,7089	42,72	59,72					
2012-2017	Completo	RVS	0,7444	39,55	44,28	Linear	SCG	Logística	54	1000
		RF	0,6819	42,39	47,03					
		RNA	0,6923	45,04	39,79					
	Reduzido	RVS	0,7786	34,52	43,70	Linear	Rprop	Exponencial	5	2000
		RF	0,7539	37,18	47,98					
		RNA	0,7903	34,05	37,37					

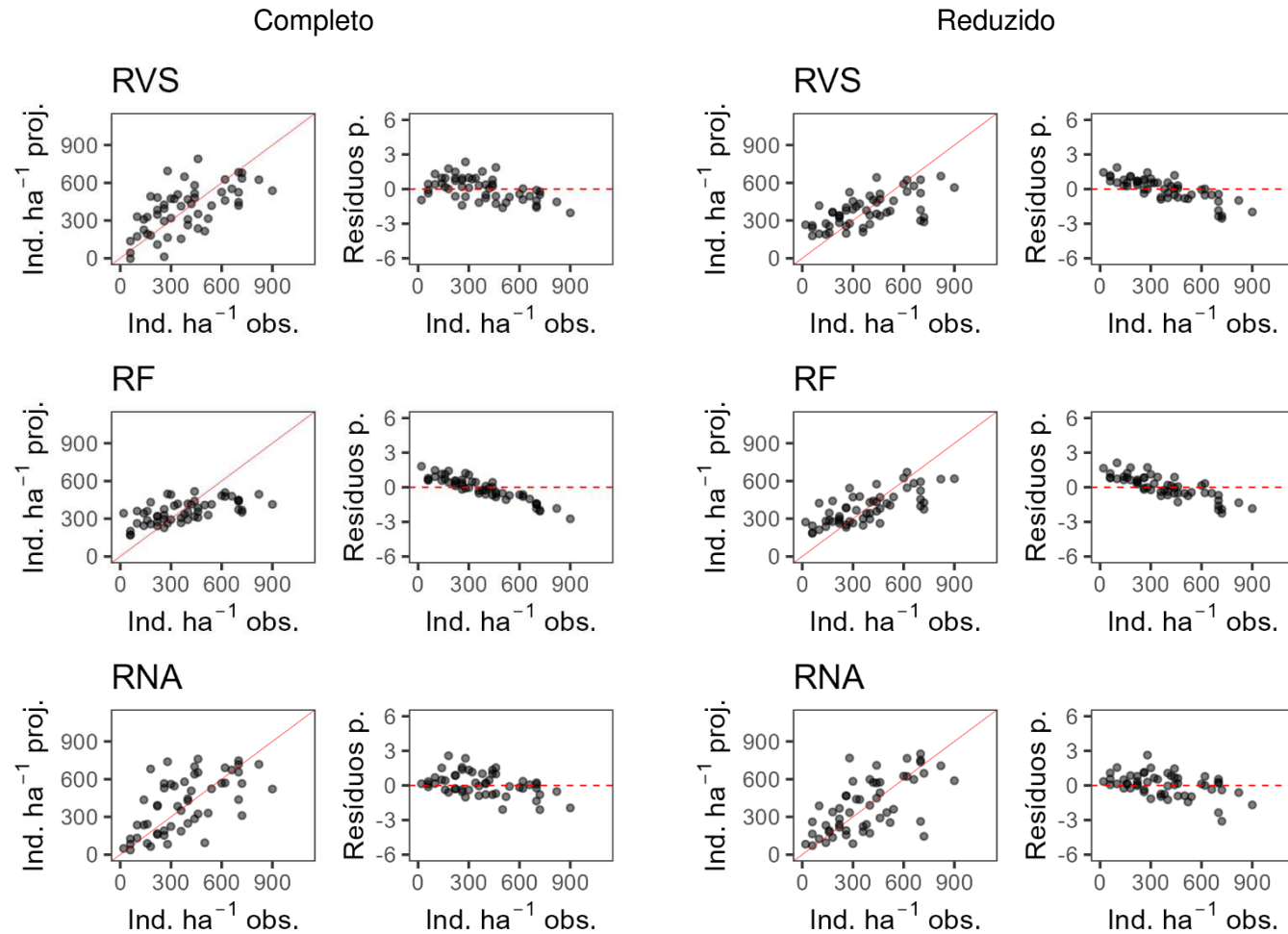
Em que: Rprop = *resilient propagation*; SCG = *scaled conjugate gradient*; RNA = rede neural artificial; RVS = regressão por vetores de suporte; RF = *random forest*. ¹ Função Kernel utilizada nas RVS. ² Algoritmo na camada oculta da RNA. ³ Função ativação na camada oculta da RNA. ⁴ Número de neurônios na camada oculta da RNA. ⁵ Número de árvores de regressão na RF.

Figura 4 - Dispersão dos valores observados e projetados de ind. ha⁻¹ e dos resíduos padronizados no período 2002-2012 nas abordagens com modelos completos e modelos reduzidos.



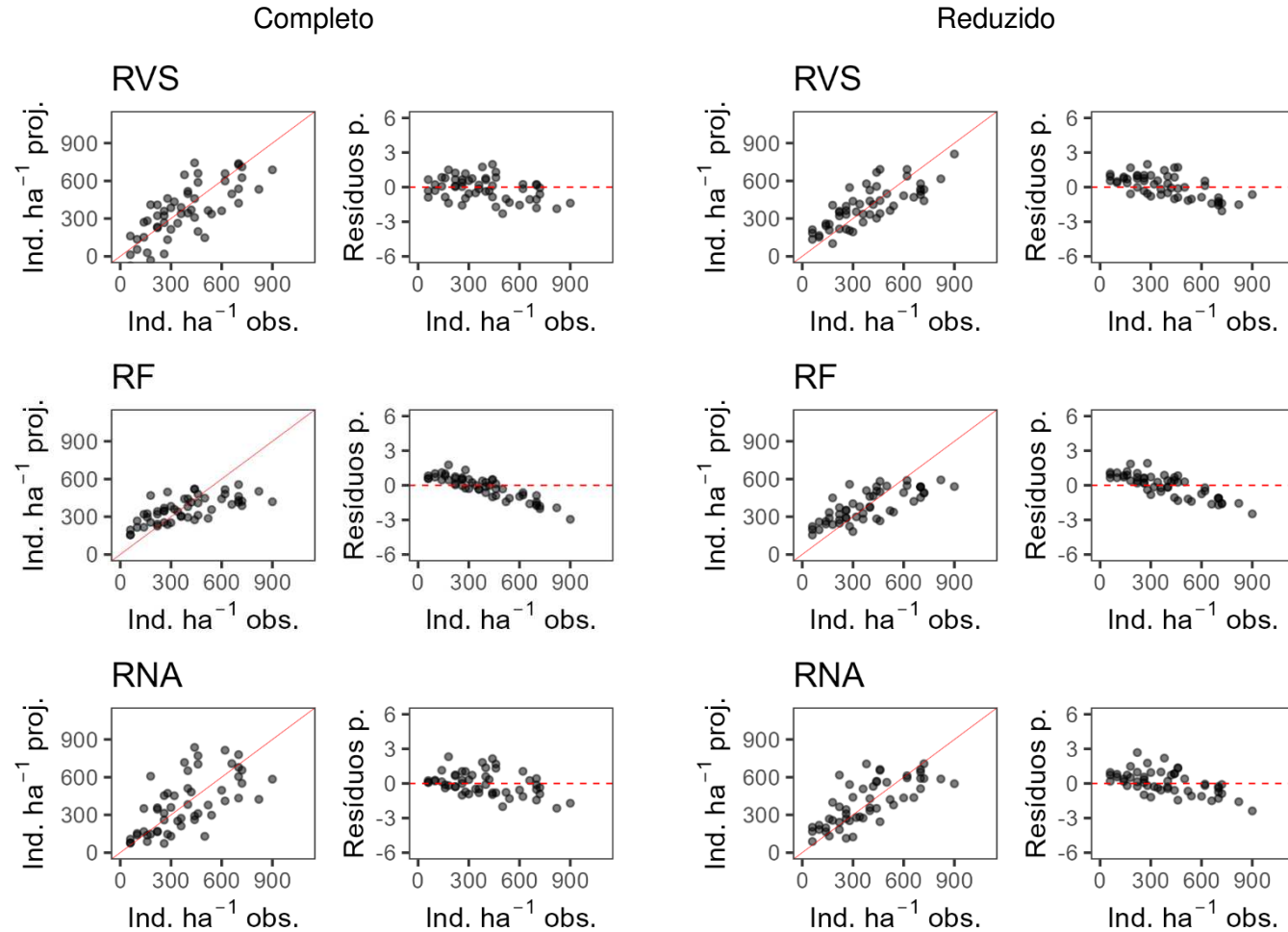
Fonte: O autor (2024).

Figura 5 - Dispersão dos valores observados e projetados de ind. ha⁻¹ e dos resíduos padronizados no período 2002-2017 nas abordagens com modelos completos e modelos reduzidos.



Fonte: O autor (2024).

Figura 6 - Dispersão dos valores observados e projetados de ind. ha⁻¹ e dos resíduos padronizados no período 2012-2017 nas abordagens com modelos completos e modelos reduzidos.



Fonte: O Autor (2024).

4 DISCUSSÃO

Neste estudo, avaliamos de forma inédita a modelagem da abundância de indivíduos infestados por lianas em fragmentos na Mata Atlântica utilizando técnicas de aprendizagem de máquina, com base em variáveis ambientais e antropogênicas. Verificamos que variáveis relacionadas ao clima, solo e atributos da floresta influenciam a abundância de indivíduos arbóreos infestados por lianas ao longo dos períodos.

Analisando a média de ind. ha⁻¹ em cada fragmento, a redução do número de indivíduos infestados entre 2002 e 2012 pode ser atribuída à sucessão florestal e ao fechamento do dossel, em que o crescimento e a regeneração das árvores proporciona um dossel mais fechado, limitando a luz disponível para as lianas e diminuindo sua abundância (ESTRADA-VILLEGAS et al., 2022; MARSHALL et al., 2020). Além disso, a sucessão ecológica nos fragmentos florestais pode ter aumentado a dominância de espécies arbóreas em relação as lianas (BECKNELL et al., 2022; TORRES et al., 2023).

O aumento na quantidade de ind. ha⁻¹ infestados entre os anos de 2012 e 2017 pode estar relacionado ao fenômeno climático do El Niño que ocorreu nos anos 2015/2016, em que esse fenômeno é conhecido por causar secas severas em várias regiões tropicais, incluindo partes do Brasil (GATEAU-REY et al., 2018; PFANNERSTILL et al., 2018). As secas severas podem causar mortalidade de árvores, abrindo o dossel e permitindo maior entrada de luz, o que favorece o crescimento das lianas (ANDEREGG; KANE; ANDEREGG, 2013; HAMMOND et al., 2022). Além disso, tempestades e ventos fortes associados a essas condições climáticas extremas podem danificar as árvores, criando mais oportunidades para as lianas se estabelecerem (MAGNAGO et al., 2017; MITCHELL, 2013).

4.1 Importância das variáveis selecionadas

Ao avaliar as variáveis mais importantes selecionadas para projeção de ind. ha⁻¹ na abordagem com modelo reduzido, a correlação negativa entre a área basal por hectare (BA_ha) e a abundância de indivíduos infestados por lianas ($r = -0.22$) nos períodos 2002-2012 e 2002-2017 sugere que áreas com maior área basal, indicativas de florestas mais maduras e densas, tendem a ter menor infestação de lianas. As

árvores em florestas maduras podem fornecer menos oportunidades para o estabelecimento de lianas devido à maior competição por luz e espaço (FINLAYSON et al., 2024).

A correlação negativa da matéria orgânica de 20-40 cm de profundidade (MO_20_40) com a abundância de indivíduos infestados por lianas para os períodos 2002-2017 e 2012-2017 ($r = -0,4$) indica que solos mais ricos em matéria orgânica podem ser menos favoráveis ao crescimento de lianas, devido a melhores condições de crescimento para árvores que competem com lianas, favorecendo o fechamento do dossel e reduzindo a quantidade de luz e nutrientes para as lianas (HOU et al., 2020).

A precipitação total dos três meses mais secos (Prec_dry) apresentou uma correlação negativa com a abundância de indivíduos infestados por lianas no período de 2002-2017 ($r = -0,42$). Esse resultado indica que o estresse hídrico limitou o crescimento das árvores, ocasionando a morte de alguns indivíduos e o aumento da infestação por lianas (DEURWAERDER et al., 2018). Por outro lado, no período de 2012-2017, a relação entre Prec_dry e a abundância de indivíduos infestados por lianas foi oposta, com uma correlação positiva ($r = 0,45$). Esse resultado pode ser explicado devido os valores de Prec_dry de 2012 (entre 4,1 e 15,3 mm) serem menores que do ano de 2002 (entre 40,6 e 55,6 mm), e desse modo, o valor máximo de Prec_dry de 2012 foi o de maior abundância de indivíduos infestados por lianas. Isso indica que as lianas se adaptaram melhor as condições de seca até um determinado limiar, onde elas podem ter desenvolvido estratégias mais eficientes para utilização da água do que as árvores (BECKNELL et al., 2022).

4.2 Desempenho das técnicas de aprendizagem de máquina

As melhores estatísticas para maioria das abordagens e períodos com as redes neurais artificiais (RNA) podem ser atribuídas a capacidade das RNA de realizar o processamento com grandes volumes de dados e lidar com muitas variáveis ao mesmo tempo, principalmente na abordagem com modelo completo, em que uma maior quantidade de variáveis foram inseridas na entrada em relação a abordagem com modelo reduzido (ALZUBAIDI et al., 2021; NAJAFABADI et al., 2015).

A regressão vetorial de suporte (RVS) também apresentou bons resultados nos ajustes, especialmente no período de 2012-2017 com o modelo completo. A RVS é

eficiente em lidar com dados de alta dimensionalidade e pode capturar relações complexas entre variáveis. Essa técnica utiliza a função de kernel para transformar dados em um espaço de características de maior dimensão, onde um hiperplano ótimo pode ser encontrado para a regressão (BASAK; PAL; PATRANABIS, 2007; WANG et al., 2016). Além disso, a RVS é eficaz na minimização de erros de previsão ao considerar margens de tolerância, o que pode explicar os bons resultados nas estatísticas de qualidade do ajuste (LANGHAMMER; ČESÁK, 2016; LÓPEZ-SERRANO et al., 2020).

A Random Forest (RF), apesar do bom desempenho em algumas abordagens, apresentou tendência de subestimação dos maiores valores de indivíduos infestados por lianas por hectare. A abordagem de média das projeções das árvores dentro da RF pode levar à suavização das estimativas, particularmente em dados com valores extremos ou altamente variáveis (WOLFENBERGER et al., 2021). Essa suavização explica a subestimação dos maiores valores de ind. ha⁻¹, pois a RF tende a diminuir os valores extremos devido a agregação dos valores projetados em um valor médio.

Os valores próximos das estatísticas de qualidade do ajuste da abordagem com modelo reduzido ao completo nos períodos 2002-2012 e 2002-2017, e maior à abordagem com modelo completo no período de 2012-2017 sugere que a seleção de variáveis mais relevantes em um modelo reduzido pode minimizar o ruído nos dados e melhorar a precisão das previsões (CHEN et al., 2020; NOURANI; BAGHANAM; GOKCEKUS, 2018). Além disso, a inclusão de menos variáveis pode reduzir a complexidade do modelo, facilitando a captura de relações fundamentais entre as variáveis selecionadas e a abundância de indivíduos infestados por lianas (GENUER; POGGI; TULEAU-MALOT, 2010; LI et al., 2019). Já o bom desempenho das técnicas no período 2012-2017 com a abordagem com modelo reduzido sugere que os dados podem ter uma estrutura que favorece a eficácia dos modelos reduzidos, em que mudanças nas condições ambientais ou nos padrões de infestação de lianas foram captadas pelas variáveis mais importantes (CAI et al., 2018; CHEN et al., 2020; LEE; LEE, 2006).

De modo geral, os resultados deste estudo demonstram que a aplicação de técnicas de aprendizagem de máquina pode melhorar a compreensão dos fatores que afetam a dinâmica dos ecossistemas florestais. Desse modo, a modelagem da abundância de lianas com a utilização de aprendizagem de máquina representa um avanço na ecologia de florestas tropicais e pode auxiliar na conservação e manejo

sustentável de fragmentos da Mata Atlântica.

5 CONCLUSÃO

As redes neurais artificiais (RNA) e regressão vetorial de suporte (RVS) são indicadas para projetar o número de indivíduos infestados por lianas por hectare (ind. ha⁻¹).

A abordagem de modelo completo é a melhor para projeção de ind. ha⁻¹ no período de 10 (2002-2012) e 15 anos (2002-2017). Para o período de 5 anos, a melhor abordagem é a de modelo reduzido.

As variáveis matéria orgânica de 20-40 cm de profundidade, área basal por hectare e precipitação total nos três meses mais secos são as mais eficientes para projetar a abundância de indivíduos infestados por lianas em fragmentos de Mata Atlântica.

REFERÊNCIAS

ÁLVAREZ-CANSINO, L.; SCHNITZER, S. A.; REID, J. P.; POWERS, J. S. Liana competition with tropical trees varies seasonally but not with tree species identity. **Ecology**, v. 96, n. 1, p. 39–45, 2015.

BASTIN, J. F.; BARBIER, N.; COUTERON, P.; ADAMS, B.; SHAPIRO, A.; BOGAERT, J.; DE CANNIÈRE, C. Aboveground biomass mapping of African forest mosaics using canopy texture analysis: Toward a regional approach. **Ecological Applications**, v. 24, n. 8, p. 1984–2001, 2014.

BERGMEIR, C.; BENITEZ, J. M. Neural networks in R using the Stuttgart Neural Network Simulator: RSNNS. **Journal of Statistical Software**, v. 46, n. 7, p. 1–26, 2012.

BERINI, J. L.; BROCKMAN, S. A.; HEGEMAN, A. D.; REICH, P. B.; MUTHUKRISHNAN, R.; MONTGOMERY, R. A.; FORESTER, J. D. Combinations of abiotic factors differentially alter production of plant secondary metabolites in five woody plant species in the boreal-temperate transition zone. **Frontiers in Plant Science**, v. 9, n. September, 2018.

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001.

CLARK, A. T.; DETTO, M.; MULLER-LANDAU, H. C.; SCHNITZER, S. A.; WRIGHT, S. J.; CONDIT, R.; HUBBELL, S. P. Functional traits of tropical trees and lianas explain spatial structure across multiple scales. **Journal of Ecology**, v. 106, n. 2, p. 795–806, 2018.

COUTERON, P.; BARBIER, N.; GAUTIER, D. Textural ordination based on fourier spectral decomposition: A method to analyze and compare landscape patterns. **Landscape Ecology**, v. 21, n. 4, p. 555–567, 2006.

DI PORCIA E BRUGNERA, M.; MEUNIER, F.; LONGO, M.; KRISHNA MOORTHY, S. M.; DE DEURWAERDER, H.; SCHNITZER, S. A.; BONAL, D.; FAYBISHENKO, B.; VERBEECK, H. Modeling the impact of liana infestation on the demography and carbon cycle of tropical forests. **Global Change Biology**, v. 25, n. 11, p. 3767–3780, 2019.

ECKERT, S. Improved forest biomass and carbon estimations using texture measures from worldView-2 satellite data. **Remote Sensing**, v. 4, n. 4, p. 810–829, 2012.

ESTRADA-VILLEGAS, S.; SCHNITZER, S. A. A comprehensive synthesis of liana removal experiments in tropical forests. **Biotropica**, v. 50, n. 5, p. 729–739, 2018.

FARRIS, J. S. On the Cophenetic Correlation Coefficient. **Systematic Zoology**, v. 18, n. 3, p. 279–285, 1969.

FOSTER, J. R.; TOWNSEND, P. A.; ZGANJAR, C. E. Spatial and temporal patterns of gap dominance by low-canopy lianas detected using EO-1 Hyperion and Landsat Thematic Mapper. **Remote Sensing of Environment**, v. 112, n. 5, p. 2104–2117, 2008.

GUARIGUATA, M. R.; OSTERTAG, R. Neotropical secondary forest succession: Changes in structural and functional characteristics. **Forest Ecology and Management**, v. 148, n. 1–3, p. 185–206, 2001.

GUSTAFSON, E. J.; SHVIDENKO, A. Z.; STURTEVANT, B. R.; SCHELLER, R. M. Predicting global change effects on forest biomass and composition in south-central Siberia. **Ecological Applications**, v. 20, n. 3, p. 700–715, 2010.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 3, p. 610–621, 1973.

HAYKIN, S. **Neural networks and learning machines**. New Jersey: Prentice Hall, 2009.

HOGAN, J. A.; MAYORQUÍN, S.; RICE, K.; THOMPSON, J.; ZIMMERMAN, J. K.; BROKAW, N. Liana dynamics reflect land-use history and hurricane response in a Puerto Rican forest. **Journal of Tropical Ecology**, v. 33, n. 2, p. 155–164, 2017.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Mapa de vegetação do Brasil**. Disponível em: <<https://www.ibge.gov.br/geociencias/informacoes-ambientais/vegetacao/10872-vegetacao.html?=&t=o-que-e>>.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Câmeras Imageadoras CBERS 04A**. Disponível em:

<<http://www.cbears.inpe.br/sobre/cameras/cbears04a.php>>.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE. **Guidelines for national greenhouse gas inventories: Agriculture, forestry and other land use.** [s.l.: s.n.]. Disponível em: <https://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/4_Volume4/V4_06_Ch6_Grassland.pdf>.

KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. **Psychometrika**, v. 23, n. 3, p. 187–200, 1958.

KAUFFMAN, J. B.; HUGHES, R. F.; HEIDER, C. Carbon pool and biomass dynamics associated with deforestation, land use, and agricultural abandonment in the neotropics. **Ecological Applications**, v. 19, n. 5, p. 1211–1222, 2009.

KAYITAKIRE, F.; HAMEL, C.; DEFOURNY, P. Retrieving forest structure variables based on image texture analysis and IKONOS-2 imagery. **Remote Sensing of Environment**, v. 102, n. 3–4, p. 390–401, 2006.

LAI, H. R.; HALL, J. S.; TURNER, B. L.; VAN BREUGEL, M. Liana effects on biomass dynamics strengthen during secondary forest succession. **Ecology**, v. 98, n. 4, p. 1062–1070, 2017.

LAURANCE, W. F.; ANDRADE, A. S.; MAGRACH, A.; CAMARGO, J. L. C.; VALSKO, J. J.; CAMPBELL, M.; FEARNSIDE, P. M.; EDWARDS, W.; LOVEJOY, T. E.; LAURANCE, S. G. Long-term changes in liana abundance and forest dynamics in undisturbed Amazonian forests. **Ecology**, v. 95, n. 6, p. 1604–1611, 2014.

LEDO, A.; SCHNITZER, S. A. Disturbance and clonal reproduction determine liana distribution and maintain liana diversity in a tropical forest. **Ecology**, v. 95, n. 8, p. 2169–2178, 2014.

LU, D.; LI, G.; MORAN, E.; KUANG, W. A comparative analysis of approaches for successional vegetation classification in the Brazilian Amazon. **GIScience and Remote Sensing**, v. 51, n. 6, p. 695–709, 2014.

MESQUITA, D. P. P.; GOMES, J. P. P.; SOUZA JUNIOR, A. H.; NOBRE, J. S. Euclidean distance estimation in incomplete datasets. **Neurocomputing**, v. 248, p. 11–18, 2017.

MEYER, L.; DINIZ-FILHO, J. A. F.; LOHMANN, L. G.; HORTAL, J.; BARRETO, E.; RANGEL, T.; KISSLING, W. D. Canopy height explains species richness in the largest clade of Neotropical lianas. **Global Ecology and Biogeography**, v. 29, n. 1, p. 26–37, 2020.

MOORTHY, S. M. K.; CALDERS, K.; E BRUGNERA, M. di P.; SCHNITZER, S. A.; VERBEECK, H. Terrestrial laser scanning to detect liana impact on forest structure. **Remote Sensing**, v. 10, n. 6, p. 1–19, 2018.

OBIN, M. I. F. T.; RIGHT, A. L. J. W.; ANGAN, S. C. A. M.; AL, T. E. T. Ind. have a greater competitive effect than trees of similar biomass on tropical canopy trees. v. 3,

n. February, p. 1–11, 2012.

PARTHASARATHY, N.; VIVEK, P.; MUTHUMPERUMAL, C.; MUTHURAMKUMAR, S.; AYYAPPAN, N. **Biodiversity of Lianas**. 2015.

PHILLIPS, O. L.; VÉSQUEZ MARTÍNEZ, R.; ARROYO, L.; BAKER, T. R.; KILLEEN, T.; LEWIS, S. L.; MALHI, Y.; MONTEAGUDO MENDOZA, A.; NEILL, D.; NÚÑEZ VARGAS, P.; ALEXIADES, M.; CERÓN, C.; DI FLORE, A.; ERWIN, T.; JARDIM, A.; PALACIOS, W.; SALDIAS, M.; VINCETI, B. Increasing dominance of large lianas in Amazonian forests. **Nature**, v. 418, n. 6899, p. 770–774, 2002.

PIELOU, E. C. **An introduction to mathematical ecology**. New York: Wiley, 1969.

POWELL, S. L.; COHEN, W. B.; HEALEY, S. P.; KENNEDY, R. E.; MOISEN, G. G.; PIERCE, K. B.; OHMANN, J. L. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. **Remote Sensing of Environment**, v. 114, n. 5, p. 1053–1068, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2009.12.018>>.

R CORE TEAM. **R: A language and environment for statistical computing** Vienna R Foundation for Statistical Computing, , 2020. . Disponível em: <<https://www.r-project.org/>>.

RESENDE, M. D. V. de; SILVA, F. F.; AZEVEDO, C. F. **Estatística matemática, biométrica e computacional: Modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTL-GWAS, estatística espacial e temporal, competição, sobrevivência**. Viçosa: Suprema gráfica e Editora Ltda, 2014.

RIPLEY, B. D. Modelling Spatial Patterns. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 39, n. 2, p. 172–192, 1977.

ROZENDAAL, D. M. A.; CHAZDON, R. L.; ARREOLA-VILLA, F.; BALVANERA, P.; BENTOS, T. V.; DUPUY, J. M.; HERNÁNDEZ-STEFANONI, J. L.; JAKOVAC, C. C.; LEBRIJA-TREJOS, E. E.; LOHBECK, M.; MARTÍNEZ-RAMOS, M.; MASSOCA, P. E. S.; MEAVE, J. A.; MESQUITA, R. C. G.; MORA, F.; PÉREZ-GARCÍA, E. A.; ROMERO-PÉREZ, I. E.; SAENZ-PEDROZA, I.; VAN BREUGEL, M.; WILLIAMSON, G. B.; BONGERS, F. Demographic Drivers of Aboveground Biomass Dynamics During Secondary Succession in Neotropical Dry and Wet Forests. **Ecosystems**, v. 20, n. 2, p. 340–353, 2017.

SARKER, L. R.; NICHOL, J. E. Improved forest biomass estimates using ALOS AVNIR-2 texture indices. **Remote Sensing of Environment**, v. 115, n. 4, p. 968–977, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2010.11.010>>.

SCHNITZER, S. A.; BONGERS, F. The ecology of lianas and their role in forests. **Trends in Ecology and Evolution**, v. 17, n. 5, p. 223–230, 2002.

SCHNITZER, S. A.; BONGERS, F. Increasing liana abundance and biomass in tropical forests: Emerging patterns and putative mechanisms. **Ecology Letters**, v. 14, n. 4, p.

397–406, 2011.

SCHNITZER, S. A.; CARSON, W. P. Lianas suppress tree regeneration and diversity in treefall gaps. **Ecology Letters**, v. 13, n. 7, p. 849–857, 2010.

SCHNITZER, S. A.; VAN DER HEIJDEN, G. M. F. Ind. have a seasonal growth advantage over co-occurring trees. **Ecology**, v. 100, n. 5, p. 1–12, 2019.

SILVA, I. N.; SPATTI, D. H.; FLAUZIN, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. São Paulo: Artliber, 2010.

SOCIEDADE DE INVESTIGAÇÕES FLORESTAIS. **Monitoramento da flora em áreas preservadas pela CELULOSE NIPO-BRASILEIRA S.A. – CENIBRA**. [s.l.: s.n.].

SPOTO, F.; MARTIMORT, P.; DRUSCH, M. **Sentinel - 2: ESA's optical high-resolution mission for GMES operational services**. [s.l.: s.n.]v. 707 SP

TAVARES JÚNIOR, I. da S.; TORRES, C. M. M. E.; LEITE, H. G.; CASTRO, N. L. M. de; SOARES, C. P. B.; CASTRO, R. V. O.; FARIAS, A. A. Machine learning: Modeling increment in diameter of individual trees on Atlantic Forest fragments. **Ecological Indicators**, v. 117, p. 106685, out. 2020.

VAN DER HEIJDEN, G. M. F.; POWERS, J. S.; SCHNITZER, S. A. Lianas reduce carbon accumulation and storage in tropical forests. **Proceedings of the National Academy of Sciences of the United States of America**, v. 112, n. 43, p. 13267–13271, 2015.

VAN DER SANDE, M. T.; PEÑA-CLAROS, M.; ASCARRUNZ, N.; ARETS, E. J. M. M.; LICONA, J. C.; TOLEDO, M.; POORTER, L. Abiotic and biotic drivers of biomass change in a Neotropical forest. **Journal of Ecology**, v. 105, n. 5, p. 1223–1234, 2017.

VELOSO, H. P.; RANGEL FILHO, A. L. R.; LIMA, J. C. A. **Classificação da vegetação brasileira, adaptada a um sistema universal**. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE), 1991.

Capítulo III: Modelagem da biomassa do fuste de árvores em fragmentos de Mata Atlântica integrando descritores texturais GLCM e aprendizagem de máquina

RESUMO

O sensoriamento remoto é uma importante ferramenta para o monitoramento da biomassa florestal, pois permite monitorar áreas de difícil acesso com precisão em diferentes períodos. Além disso, o sensoriamento remoto fornece dados essenciais para modelar a biomassa, contribuindo para a gestão sustentável e conservação da Mata Atlântica. O objetivo deste capítulo foi prever a biomassa do fuste de árvores em fragmentos da Mata Atlântica utilizando descritores texturais integrados com técnicas de aprendizagem de máquina, considerando a influência dos indivíduos arbóreos infestados por lianas. Os fragmentos estudados localizam-se nos municípios de Guanhões (Cachoeira das Pombas), Ipaba (Ipaba) e Coronel Fabriciano (São José). Imagens do sensor ETM+ do satélite Landsat 7 foram adquiridas em 2002 nos períodos seco e chuvoso. A análise textural das imagens foi realizada por meio da matriz de coocorrência em níveis de cinza (GLCM), com extração de oito descritores texturais: segundo momento angular, contraste, variância, homogeneidade, correlação, entropia, dissimilaridade e média. A biomassa do fuste foi calculada e extrapolada para hectare em cada parcela dos fragmentos. As técnicas de aprendizagem de máquina regressão vetorial de suporte (RVS) e redes neurais artificiais (RNA) e *random forest* (RF) foram treinadas e validadas utilizando descritores texturais e atributos florestais, como área basal por hectare e número de indivíduos infestados por lianas por hectare. A análise de componentes principais (PCA) foi utilizada para reduzir a dimensionalidade das variáveis. As componentes principais (CP) que representaram de forma acumulada pelo menos 80% da variação dos dados foram integradas ao conjunto de dados de entrada nos modelos. Os descritores texturais extraídos no período chuvoso no fragmento Cachoeira das Pombas apresentaram aumento nos valores de média, variância, dissimilaridade e segundo momento angular, e redução nos valores de contraste em relação ao período seco. No fragmento Ipaba, o período chuvoso teve aumento de contraste, dissimilaridade e correlação, com redução de variância e segundo momento angular. No fragmento São José, a variância aumentou no período chuvoso em relação ao período seco, enquanto homogeneidade, dissimilaridade e segundo momento angular diminuíram. A RVS com abordagem completa apresentou melhor desempenho na validação durante o período seco, com correlação de 0,8596, RMSE% de 36,72% e MAPE de 25,35%. No período chuvoso, a RVS também apresentou melhores estatísticas que a RNA e RF. A técnica de RVS mostrou robustez na validação, principalmente no período seco. A inclusão de variáveis texturais e a consideração dos indivíduos infestados por lianas melhoram a precisão dos modelos de estimativa de biomassa em fragmentos de Mata Atlântica. Esses resultados destacam a importância de adaptar os modelos às condições sazonais e estruturais da vegetação, promovendo uma compreensão detalhada e um manejo sustentável dos fragmentos florestais.

Palavras-chave: Análise textural; Floresta tropical; Modelagem da biomassa; Sensoriamento remoto.

1 INTRODUÇÃO

A Mata Atlântica é um dos biomas mais biodiversos do mundo, abrigando várias espécies endêmicas de flora e fauna (DELICIELLOS et al., 2022; PORTILLO et al., 2023). Desse modo, a conservação desse bioma é vital para manutenção de processos ecológicos e biológicos essenciais, como a regulação do clima, a ciclagem de nutrientes e o fornecimento de habitat para inúmeras espécies (VIERA et al., 2016; RODRIGUES et al., 2021). Apesar dessa importância da Mata Atlântica, atualmente restam apenas cerca de 12% da sua área original devido à intensa degradação e fragmentação florestal (RIBEIRO et al., 2011).

A fragmentação florestal na Mata Atlântica tem impactos negativos na biomassa florestal, reduzindo a densidade e diversidade de árvores e afetando diretamente a capacidade da floresta armazenar carbono (WILSON et al., 2016; RAJ et al., 2022). Além disso, a fragmentação leva ao aumento da abundância de lianas, plantas trepadeiras lenhosas que competem com as árvores por recursos essenciais como luz, água e nutrientes (VARGAS; GROMBONE-GUARATINI; MORELLATO, 2021). As lianas também podem reduzir a eficiência fotossintética das árvores, aumentando a vulnerabilidade a danos estruturais e diminuem o crescimento e a sobrevivência dos indivíduos arbóreos, resultando em uma diminuição da biomassa (ESTRADA-VILLEGAS et al., 2022). Dada a importância da biomassa das árvores para a elaboração de estratégias eficazes de conservação e manejo, bem como para a modelagem de cenários futuros de mitigação dos efeitos climáticos, a sua estimativa de forma precisa torna-se fundamental (ZHANG; LIANG; YANG, 2019; ALI HUSSIN, 2022).

Nesse contexto, as tecnologias de sensoriamento remoto surgem como uma grande aliada na quantificação da biomassa florestal, permitindo a análise de grandes áreas de forma eficiente (TIAN et al., 2023; BORSAH; NAZEER; WONG, 2023). Sensores multiespectrais e hiperespectrais são amplamente utilizados no sensoriamento remoto para estimar a biomassa em ecossistemas florestais (EDIRIWEERA et al., 2014; MASSEY et al., 2023; HEMATI et al., 2024; TIAN et al., 2024; SAINUDDIN et al., 2024; SANAM et al., 2024). No entanto, a presença de lianas nas copas das árvores apresenta um desafio na estimativa precisa da biomassa arbórea (MEUNIER et al., 2021), sendo que essas plantas podem interferir nas respostas espectrais, dificultando a distinção entre a reflectância das árvores e das

lianas (SÁNCHEZ-AZOFEIFA et al., 2009).

Desse modo, a aplicação de descritores texturais na análise de imagens de satélite para capturar a complexidade estrutural das florestas surge como uma alternativa, em que esses descritores quantificam a variação textural das imagens (PLOTON et al., 2013; SA; FAN, 2024). A textura refere-se ao padrão espacial das variações de tons ou cores que ocorrem dentro de uma imagem (HUNG; SONG; LAN, 2019). A textura descreve a distribuição espacial das variações de intensidade, representando características como rugosidade, suavidade, granularidade, e repetição de padrões, que pode ser usada para identificar características estruturais da vegetação (ECK et al., 2016; DISTANTE; DISTANTE, 2020). Estudos anteriores demonstraram a eficácia dos descritores texturais na classificação de espécies arbóreas (FERREIRA et al., 2019), estimativa da biomassa florestal (KUPLICH; CURRAN; ATKINSON, 2005), modelagem de características funcionais do dossel (AGUIRRE-GUTIÉRREZ et al., 2021), mapeamento de distúrbios florestais (BALLING; HEROLD; REICHE, 2023) e estimativa da altura de copa (CSILLIK; KUMAR; ASNER, 2020). Contudo, estudos que estimam a biomassa de árvores utilizando descritores texturais em modelos de aprendizagem de máquina e incluindo a abundância dos indivíduos infestados por lianas como entrada nesses modelos ainda são raros, principalmente para Mata Atlântica.

As técnicas de aprendizagem de máquina são capazes de processar quantidades massivas de dados, como os de sensoriamento remoto, sendo que modelos utilizando essas técnicas podem ser continuamente ajustadas e aprimoradas com a adição de novos dados, aumentando a acurácia das estimativas ao longo do tempo (SONG et al., 2024). Essa adaptabilidade torna-se importante em ecossistemas complexos como a Mata Atlântica, onde a inclusão de variáveis adicionais, como a abundância de indivíduos infestados por lianas, pode melhorar a precisão dos modelos de estimativa de biomassa (ZHANG et al., 2020; SU et al., 2020). Ainda, a estimativa da biomassa de árvores integrando descritores texturais com aprendizado de máquina é uma abordagem ainda pouco explorada. Nesse contexto, o objetivo deste capítulo foi prever a biomassa do fuste de árvores em fragmentos de Mata Atlântica utilizando descritores texturais integrados com técnicas de aprendizagem de máquina, considerando a influência dos indivíduos infestados por lianas.

2 MATERIAL E MÉTODOS

2.1 Caracterização das áreas de estudo

Os fragmentos avaliados neste capítulo foram os mesmos do Capítulo I, estando localizados nos municípios de Guanhões (Cachoeira das Pombas), Ipaba (Ipaba) e Coronel Fabriciano (São José). As características dos fragmentos podem ser consultadas seção 2.1 do Capítulo I.

2.2 Coleta dos dados

As coletas dos dados em campo foram realizadas em parcelas permanentes retangulares de área fixa de 10 × 50 m (0,05 ha). Cada parcela foi subdividida em 20 subparcelas de 5 × 5 m, e as árvores com circunferência a 1,30 m de altura do solo (CAP; cm) ≥ 15 cm foram mapeadas. O nome científico das espécies, a CAP, a altura total – Ht (m) e o grau de infestação de lianas em árvores com CAP ≥ 15 cm foram registrados nos anos de 2002 e 2012, em cada fragmento. Posteriormente, a CAP foi convertida em diâmetro à 1,30 m de altura (DAP; cm). A classificação da infestação de lianas nas árvores foi feita da seguinte forma: 1 - Nenhuma liana; 2 - Lianas apenas no tronco; 3 - Lianas apenas na copa; e 4 - Lianas no tronco e na copa.

O número de indivíduos infestados por lianas por hectare no ano de 2002 (Ind_ha_2002) foi determinado para cada parcela de todos os fragmentos (Cachoeira das Pombas, Ipaba e São José). Para isso, as árvores que possuíam a classe de infestação igual ou superior a 2, registradas em campo, foram filtradas e contadas em cada parcela dos fragmentos. Posteriormente, os valores de cada parcela foram extrapolados para hectare, resultando na abundância de ind. ha⁻¹. A área basal por hectare (BA_ha) também foi calculada para cada parcela dos fragmentos.

2.3 Dados de satélite

Os dados do sensor óptico passivo foram obtidos por meio do *Enhanced Thematic Mapper Plus* (ETM+) a bordo do satélite Landsat 7. As características do sensor estão descritas na Tabela 1. A banda 8 pancromático foi a selecionada para obtenção dos descritores de textura devido a sua resolução espacial de 15 m,

possibilitando a extração de maior quantidade de valores dos descritores texturais dentro das parcelas com 500 m² cada. As imagens possuem resolução temporal de 16 dias, resolução radiométrica de 8 bits e área imageada de 183 km. Os dados estão disponíveis no site do *Earth Explorer* administrado pelo Serviço Geológico dos Estados Unidos (USGS) no seguinte endereço eletrônico: “<https://earthexplorer.usgs.gov/>”.

Tabela 1 - Características do sensor ETM⁺ acoplado ao satélite Landsat 7

Bandas Espectrais	Res. Espectral	Res. espacial
(B1) Azul	0,45 - 0,52 μm	30 m
(B2) Verde	0,52 - 0,60 μm	30 m
(B3) Vermelho	0,63 - 0,69 μm	30 m
(B4) Infravermelho próximo	0,77 - 0,90 μm	30 m
(B5) Infravermelho de ondas curtas	1,55 - 1,75 μm	30 m
(B6) Infravermelho termal	10,4 - 12,5 μm	60 m
(B7) Infravermelho médio	2,08 - 2,35 μm	30 m
(B8) Pancromático	0,52 - 0,90 μm	15 m

As imagens foram adquiridas no ano de 2002 nas datas de 13 de setembro (período seco) e 02 de dezembro (período chuvoso) para analisar se ocorre diferenças nos descritores texturais no período seco e chuvoso.

2.4 Análise textural das imagens

A análise textural da imagem em cada fragmento foi realizada por meio da matriz de coocorrência em níveis de cinza (GLCM) proposta por Haralick et al. (1973).

A GLCM é uma técnica de extração de características de textura baseada nas relações de adjacência dos tons de cinza em uma imagem. A GLCM considera a frequência de ocorrência de pares de tons de cinza, separados por uma distância d e em uma direção específica, para definir a informação de textura da imagem. Os valores da GLCM f_{ij} representam a frequência de ocorrência dos pares de níveis de cinza. Esta matriz pode ser calculada para várias direções, geralmente correspondendo aos ângulos de 0°, 45°, 90° e 135° (MATHER; TSO, 2009). Desse modo, os descritores texturais foram calculadas com base nessa matriz para fornecer informações sobre a textura da imagem.

Haralick et al., (1973) definiram 14 descritores texturais que foram derivadas de

matrizes de coocorrência. Oito desses descritores foram analisados nas imagens: segundo momento angular, contraste, variância, homogeneidade, correlação, entropia, dissimilaridade e média (Tabela 2).

Tabela 2 - Descritores texturais analisados extraídos da matriz de coocorrência em níveis de cinza (GLCM)

Descritores	Fórmula
Média (glcm_mean)	$ME = \sum_{i,j=0}^{N-1} i p_{i,j}$
Entropia (glcm_entropy)	$ENT = - \sum_i \sum_j p(i,j) \log p(i,j)$
Contraste (glcm_contrast)	$CON = \sum_{n=0}^{N_G-1} n^2 \times \left[\sum_i \sum_{\substack{j \\ i-j =n}} p(i,j) \right]$
Segundo momento angular (glcm_second_moment)	$SMA = \sum_i \sum_j [p(i,j)]^2$
Média (glcm_mean)	$ME = \sum_{i,j=0}^{N-1} i p_{i,j}$
Variância (glcm_variance)	$VA = \sum_{i,j=0}^{N-1} i p_{i,j} (i-ME)^2$
Correlação (glcm_correlation)	$CC = \sum_{i,j=0}^{N-1} p_{i,j} \left[\frac{(i-ME)(j-ME)}{\sqrt{VA_i VA_j}} \right]$
Dissimilaridade (glcm_dissimilarity)	$DI = \sum_{i,j=0}^{N-1} i p_{i,j} i-j $

Em que: $p_{i,j}$ = matriz de coocorrência normalizada de modo que $\sum(i, j = 0)(p(i, j)) = 1$.

Os descritores foram extraídos utilizando a função `glcm` do pacote “`glcm`” no programa R (ZVOLEFF, 2020).

2.4.1 Tamanho da janela

A identificação de texturas adequadas também envolve a seleção do tamanho da janela. Se o tamanho da janela for muito pequeno em relação à estrutura da textura, a característica textural extraída não refletirá com precisão a propriedade real da textura, enquanto uma janela grande pode não extrair informações de textura de forma

eficiente devido à suavização excessiva das variações de textura (MATHER; TSO, 2009).

Desse modo, o tamanho da janela para cada fragmento foi obtido a partir das seguintes etapas: 1) aplicação do GLCM e extração do descritor média com três tamanhos de janelas (3×3 , 5×5 e 7×7) e em quatro direções (0° , 45° , 90° e 135°); 2) construção de semivariogramas para cada uma das combinações de tamanho de janela e direção obtidos na etapa anterior e; 3) avaliação dos semivariogramas e seleção do tamanho de janela que propiciou a maior semivariância em cada fragmento.

2.5 Estimativa e predição da biomassa no fuste

O volume do fuste com casca das árvores em todos os fragmentos foi obtido por meio da equação ajustada por Amaro (2013), em uma Floresta Estacional Semidecidual, no Município de Viçosa – MG (Equação 1; R^2 (%) = 97,04 e $S_{x,y}$ (%) = 17,4). A biomassa do fuste foi calculada através da Equação 2 e posteriormente, a biomassa de cada parcela foi extrapolada para $Mg\ ha^{-1}$. A densidade básica foi determinada a partir de informações extraídas de diversas fontes bibliográficas (AMARO et al., 2013; SOUZA et al., 2012; TORRES et al., 2017; ZANNE et al., 2009).

$$VFCC_j = 0,000070 \times dap_j^{2,204301} \times h_j^{0,563181} \quad (1)$$

$$B_i = VFCC_j \times d \quad (2)$$

Em que: $VFCC_j$ = volume do fuste com casca do j -ésimo indivíduo, em m^3 ; dap_j = diâmetro a 1,30 m de altura do j -ésimo indivíduo, em cm; h_j = altura total do j -ésimo indivíduo, em m; B_i = biomassa do j -ésimo indivíduo, em Kg; d = densidade básica da madeira, em $Kg\ m^{-3}$.

A predição da biomassa de cada parcela foi predita utilizando atributos da floresta (área basal por hectare e número de indivíduos infestados por lianas por hectare), descritores texturais extraídos da GLCM e componentes principais construídas a partir de todas as variáveis (atributos da floresta + descritores texturais). O período de predição foi 2002-2012, ou seja, as variáveis de entrada nos modelos foram referentes ao ano de 2002 para predizer os valores de biomassa no ano de 2012.

2.6 Análise de componentes principais

A análise de componentes principais (PCA) foi utilizada para reduzir a dimensionalidade das variáveis de entrada dos modelos. Os valores das variáveis foram padronizados antes de submeter a PCA, de modo a reduzir o efeito da magnitude das variáveis: $Z_i = (x_i - \bar{x})/\sigma$; em que: Z_i = valor padronizado da i -ésima observação; x_i = valor da i -ésima observação; \bar{x} = média dos valores observados; σ = desvio padrão.

A dispersão das variáveis dentro das duas primeiras componentes principais (CP) foi plotada e agrupadas em função dos fragmentos. As CP que representaram 80% ou mais da variação total dos dados foram selecionadas como variáveis de entrada nos modelos.

As PCA foram construídas por meio da função PCA do pacote “FactoMineR” no programa R (LÊ; JOSSE; HUSSON, 2008).

2.7 Pré-processamento dos dados

As variáveis quantitativas, com exceção das CP, foram escalonadas de forma a transformar a amplitude e magnitude das variáveis de entrada e saída para o intervalo de 0 a 1 (Equação 3).

$$v' = \frac{v - \min(v)}{\max(v) - \min(v)} (\max_{\text{novo}} - \min_{\text{novo}}) + \min_{\text{novo}} \quad (3)$$

Em que: $\min(v)$: menor valor da variável no banco de dados de treino; $\max(v)$: maior valor da variável no banco de dados treino; \max_{novo} : maior valor da normalização; \min_{novo} : menor valor da normalização; v' : valor da observação escalonada; e v : valor da observação não escalonada.

Após a modelagem com cada técnica, os valores preditos foram retransformados para a escala original, possibilitando o cálculo das estatísticas de qualidade do ajuste.

2.8 Treinamento da regressão vetorial de suporte (RVS) e das redes neurais artificiais (RNA)

Nos treinamentos das RVS foram utilizadas quatro configurações, representadas por quatro funções kernel, em que a seguinte função de erro foi otimizada:

$$\text{minimize } \frac{1}{2} \|w\|^2 + \frac{C}{m} \sum_{i=1}^m (\xi_i + \xi_i^*)$$

$$\text{sujeito a } (\langle \Phi(x_i), w \rangle + b) - y_i \leq \epsilon - \xi_i^*$$

$$y_i - (\langle \Phi(x_i), w \rangle + b) \leq \epsilon - \xi_i$$

$$\xi_i \geq 0 \quad (i = 1, \dots, m)$$

em que: w = vetor de coeficientes; C = parâmetro de penalidade do erro; ξ_i, ξ_i^* = variáveis de folga que caracterizam o erro acima e abaixo do ϵ -tubo, respectivamente; i = casos de treinamento; m = número total de casos de treinamento; $\Phi(x_i)$ = kernel utilizado; b = bias; y_i = dados de saída e; ϵ = erro máximo admitido.

As funções kernel avaliadas foram as do tipo linear, polinomial, base radial e sigmóide (Tabela 2).

Tabela 2 - Funções kernel testadas nos treinamentos das regressões vetoriais de suporte

Tipo de kernel	Função
Linear	$u' \times v$
Polinomial	$(\gamma \times u' \times v + coef0)^d$
Base radial	$e^{(-\gamma \times u-v ^2)}$
Sigmoidal	$\tanh(\gamma \times u' \times v + coef0)$

Em que: γ = parâmetro de forma necessário para todos os kernels, exceto linear [padrão: $1/(\text{dimensão de dados})$]; d = grau do polinômio; $coef0$ = parâmetro de penalidade do erro; \tanh = tangente hiperbólica.

A função utilizada para os treinamentos das RVS foi a *svm* do pacote “*e1071*” no R (MEYER et al., 2018).

Já para o treinamento das RNA, a arquitetura de múltiplas camadas ou *Multilayer Perceptron* (MLP) foi utilizada com uma única camada oculta (HAYKIN,

2009). O intervalo do número de neurônios nesta camada foi estabelecido pelo método de Fletcher-Gloss (SILVA et al., 2010): $2 \times n^{0.5} + n_2 \leq n_1 \leq 2 \times n + 1$; em que n representa o número de entradas da rede; n_1 é a quantidade de neurônios na camada oculta; e n_2 é o número de neurônios na camada de saída.

As funções de ativação avaliadas incluíram a exponencial, logística e tangente hiperbólica. Os algoritmos de treinamento examinados foram o *Resilient Propagation* (Rprop) e o *Scaled Conjugate Gradient* (SCG). Os pesos iniciais das RNAs foram gerados de maneira aleatória, sendo posteriormente ajustados com base na função do erro. O número máximo de ciclos foi estabelecido em 1000. As RNAs foram implementadas utilizando a função *mlp* do pacote “*RSNNS*” no software R (BERGMEIR; BENITEZ, 2012).

No treinamento das *random forests* (RF), três parâmetros foram configurados (BREIMAN, 2001): o número de árvores de regressão aleatória (*ntree*), no qual foram testadas 500, 1000, 1500, 2000, 2500 e 3000 árvores de regressão; o número de variáveis de divisão (*mtry*), utilizado para determinar o número de variáveis disponíveis para dividir em cada nó da árvore, sendo que o número padrão de 1/3 das variáveis independentes foi utilizado; e o tamanho mínimo de nós terminais (*nodesize*), cujo valor foi 5. A função *randomForest* do pacote “*randomForest*” no R foi utilizada para os treinamentos das RF (LIAW; WIENER, 2002).

As variáveis de entrada foram testadas em duas abordagens: modelo completo e reduzido. No modelo completo, as variáveis de entrada foram os atributos da floresta número de indivíduos infestados por lianas por hectare, área basal por hectare, descritores texturais e as CP selecionadas. No modelo reduzido apenas as CP foram utilizadas como entrada.

2.9 Análise estatística

A avaliação de cada técnica foi com base na validação cruzada do tipo *leave-one-out* (CAWLEY; TALBOT, 2003). Dessa forma, a cada iteração o registro de uma parcela foi destinada para validação, enquanto os registros das demais parcelas foram utilizadas no treinamento até completar a validação com todas as parcelas.

Os critérios para a verificação da qualidade do ajuste foram: correlação linear de Pearson entre os valores observados e preditos – r_{yy} (Equação 4); a raiz quadrada do erro quadrático médio, em porcentagem – RMSE% (Equação 5); erro absoluto

médio percentual – MAPE (Equação 6) e; gráficos de dispersão dos resíduos padronizados – SR (Equação 7).

$$r_{yy} = \frac{\text{cov}(y, \hat{y})}{\sqrt{s^2(y) \times s^2(\hat{y})}} \quad (4)$$

$$\text{RMSE}\% = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \times \frac{100}{\bar{y}} \quad (5)$$

$$\text{MAPE} = \frac{1}{n} \times \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i} \times 100 \quad (6)$$

$$\text{SR} = \frac{\hat{y}_i - y_i}{\sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n-1}}} \quad (7)$$

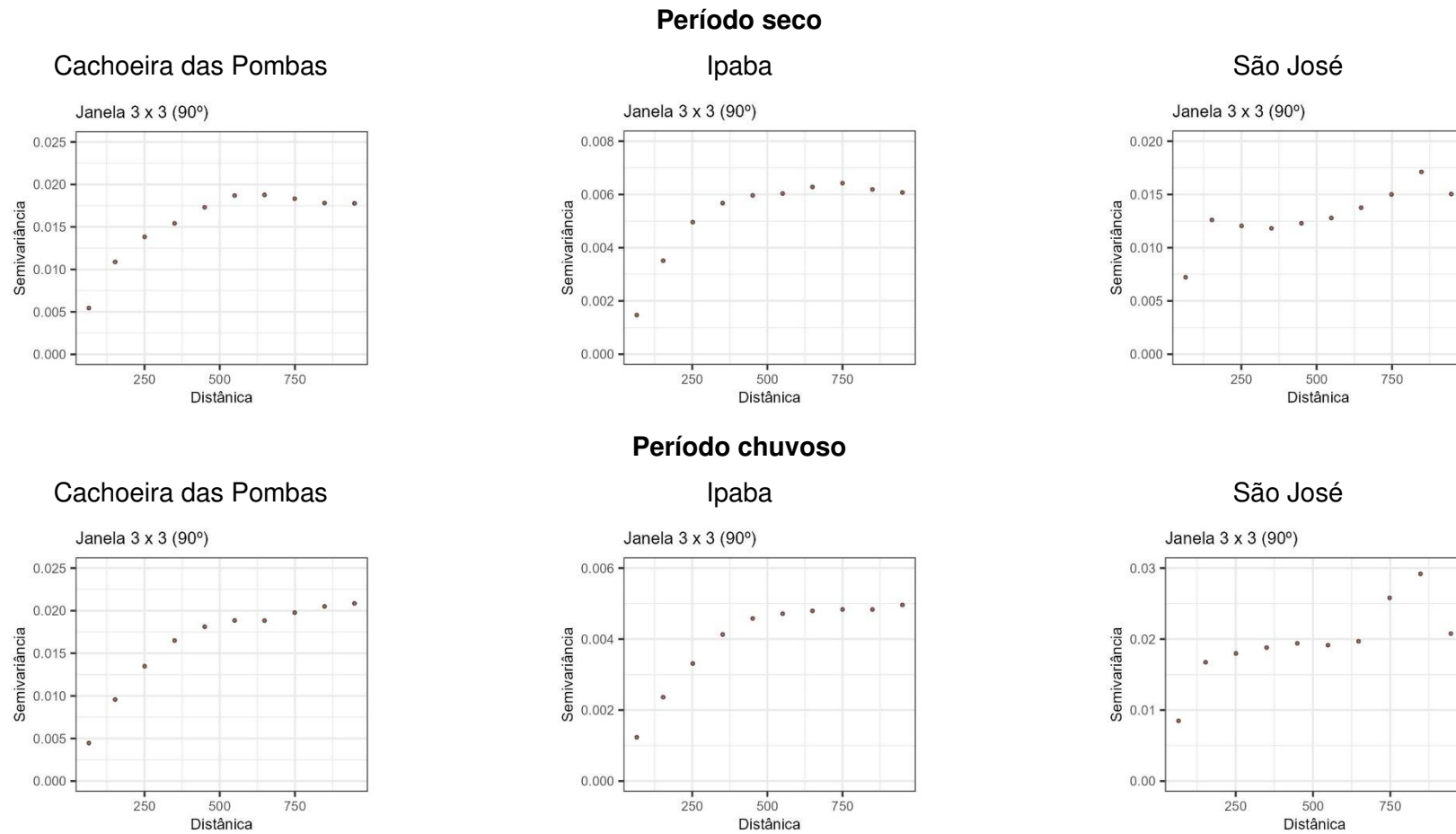
Em que: s^2 = variância; cov = covariância; y_i = valor observado da i -ésima parcela, em Mg ha⁻¹; e \hat{y}_i = valor previsto da i -ésima parcela, em Mg ha⁻¹; \bar{y} = média aritmética dos valores observados; n = número total de observações; e_i = diferença entre o valor observado e previsto da i -ésima parcela; \bar{e} = diferença média aritmética entre o valor observado e predito.

3 RESULTADOS

3.1 Tamanho da janela para extração dos descritores texturais

Analisando o resultado da seleção do tamanho da janela no fragmento Cachoeira das Pombas, a semivariância apresentou valores próximos em ambos os períodos, apesar de um leve aumento na semivariância no período chuvoso em distâncias menores (Figura 1). No fragmento de Ipaba, a semivariância foi maior no período seco e atingiu um maior patamar em relação ao período chuvoso. Por outro lado, em São José a maior semivariância foi no período chuvoso. Com base nos resultados a janela de 3 × 3 com a direção de 90° foi escolhido para a construção das GLCM em todos os fragmentos.

Figura 1 - Semivariogramas dos tamanhos de janela e direções selecionados a partir do descritor textural média para cada fragmento.



Fonte: O autor (2024).

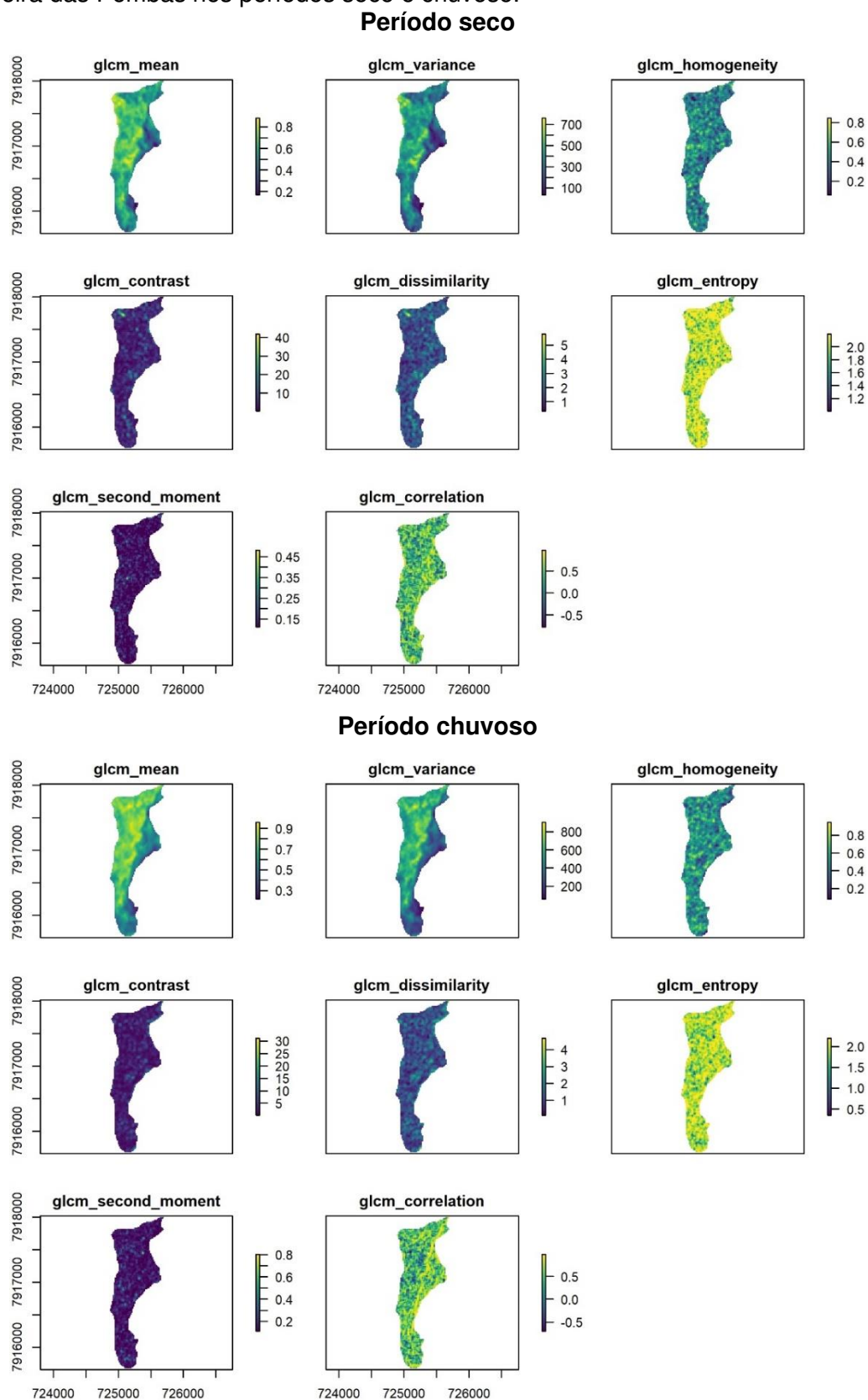
3.2 Descritores texturais extraídos da GLCM para cada fragmento

Analisando os descritores texturais extraídos da GLCM durante o período chuvoso no fragmento Cachoeira das Pombas, os valores para os descritores média (glcm_mean), variância (glcm_variance), dissimilaridade (glcm_dissimilarity) e segundo momento angular (glcm_second_moment) aumentaram em relação ao período seco (Figura 2). Por outro lado, os descritores contraste (glcm_contrast) e dissimilaridade (glcm_dissimilarity) reduziram.

No fragmento Ipaba durante o período chuvoso, os valores dos descritores contraste, dissimilaridade (glcm_dissimilarity) e correlação (glcm_correlation) aumentaram em relação ao período seco, enquanto os descritores variância (glcm_variance) e segundo momento angular (glcm_second_moment) diminuíram (Figura 3).

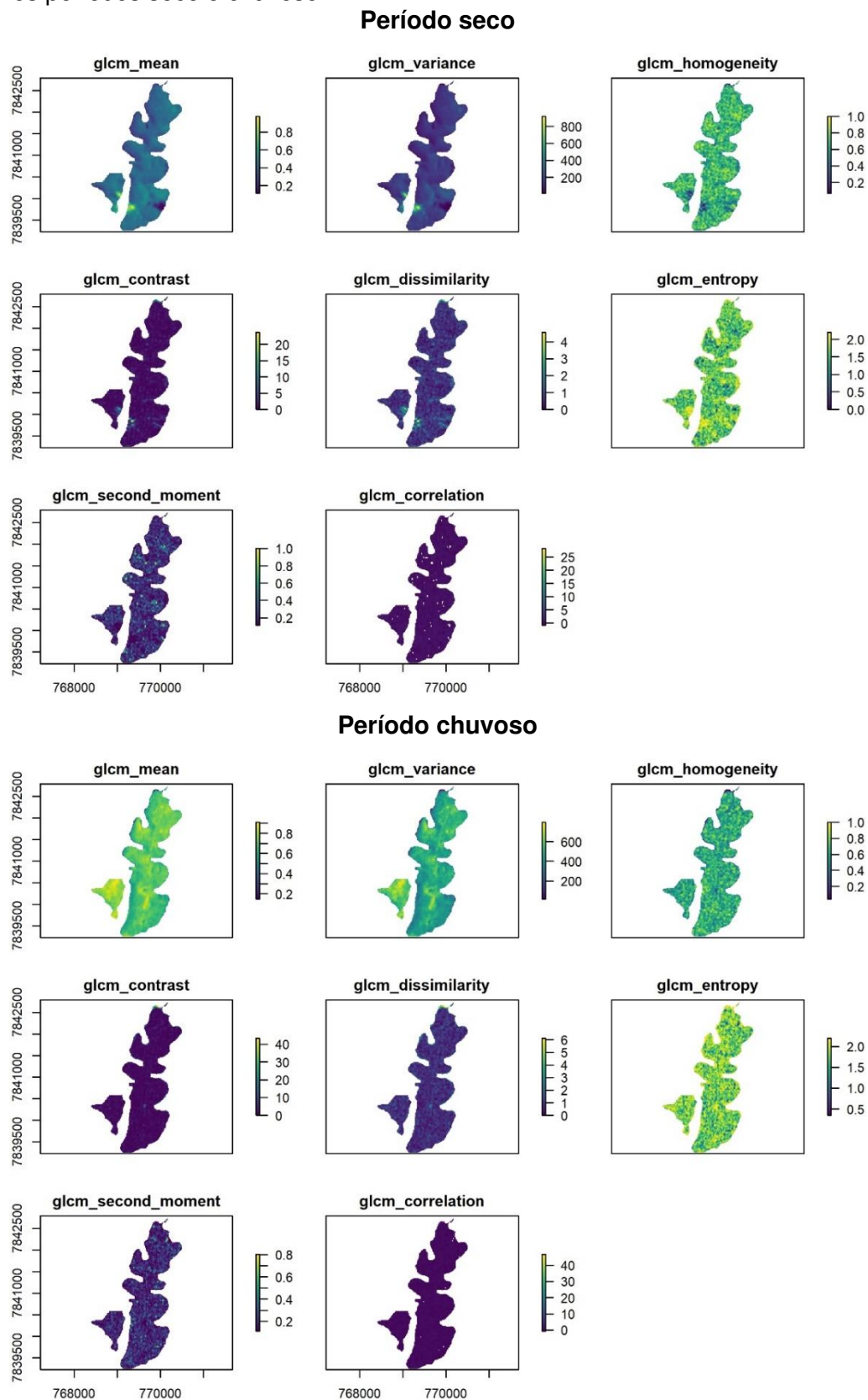
Para o fragmento São José, no período chuvoso, o descritor variância (glcm_variance) foi o único que aumentou os valores em relação ao período seco (Figura 4). Já os descritores homogeneidade, dissimilaridade e segundo momento angular reduziram em relação ao período seco.

Figura 2 - Descritores texturais calculados com tamanho da janela 3×3 e direção de 90° derivados da matriz de coocorrência em níveis de cinza (GLCM) para a imagem do fragmento Cachoeira das Pombas nos períodos seco e chuvoso.



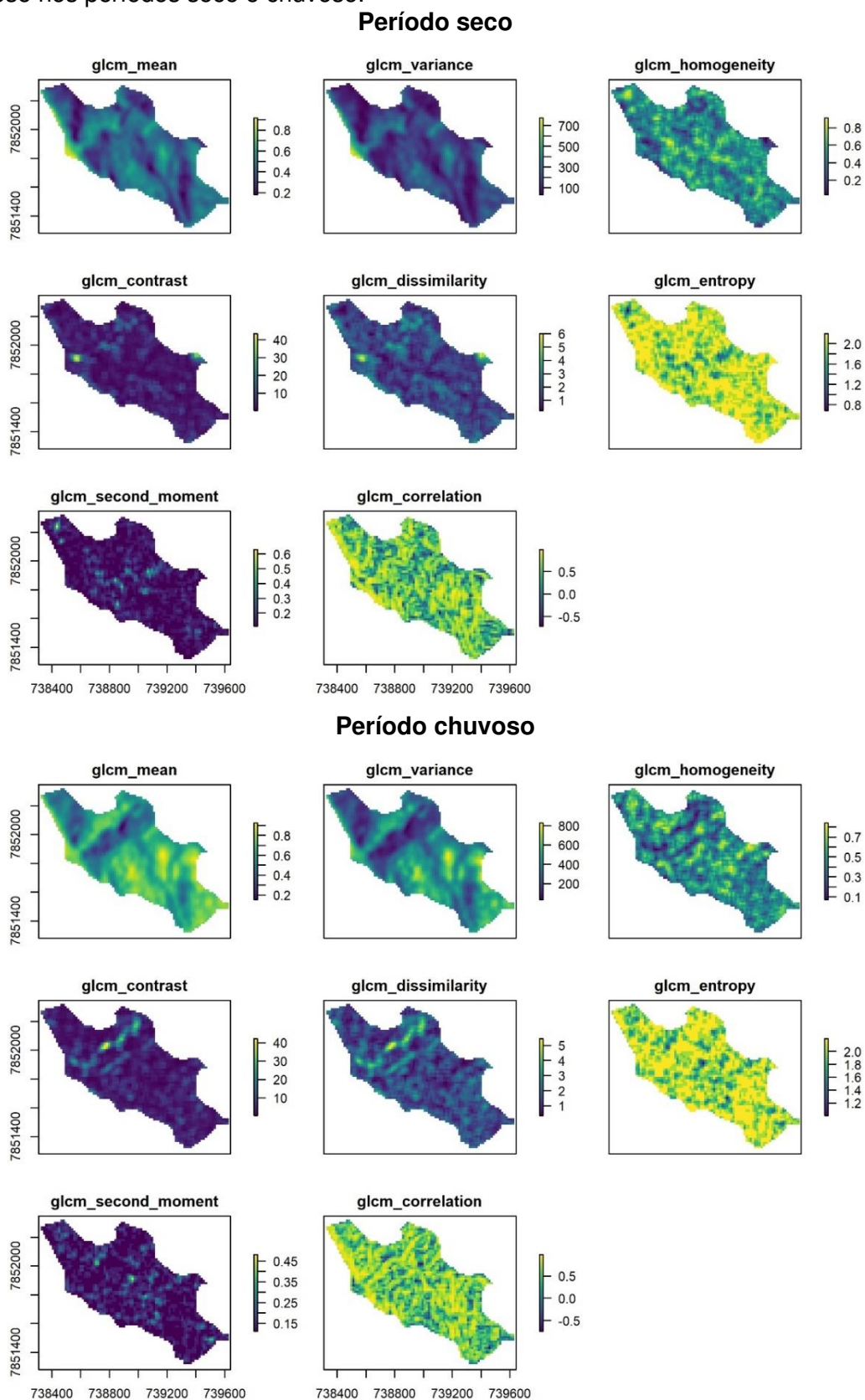
Fonte: O autor (2024).

Figura 3 - Descritores texturais calculados com tamanho da janela 3×3 e direção de 90° derivados da matriz de coocorrência em níveis de cinza (GLCM) para a imagem do fragmento Ipaba nos períodos seco e chuvoso.



Fonte: O autor (2024).

Figura 4 - Descritores texturais calculados com tamanho da janela 3×3 e direção de 90° derivados da matriz de coocorrência em níveis de cinza (GLCM) para a imagem do fragmento São José nos períodos seco e chuvoso.



Fonte: O autor (2024).

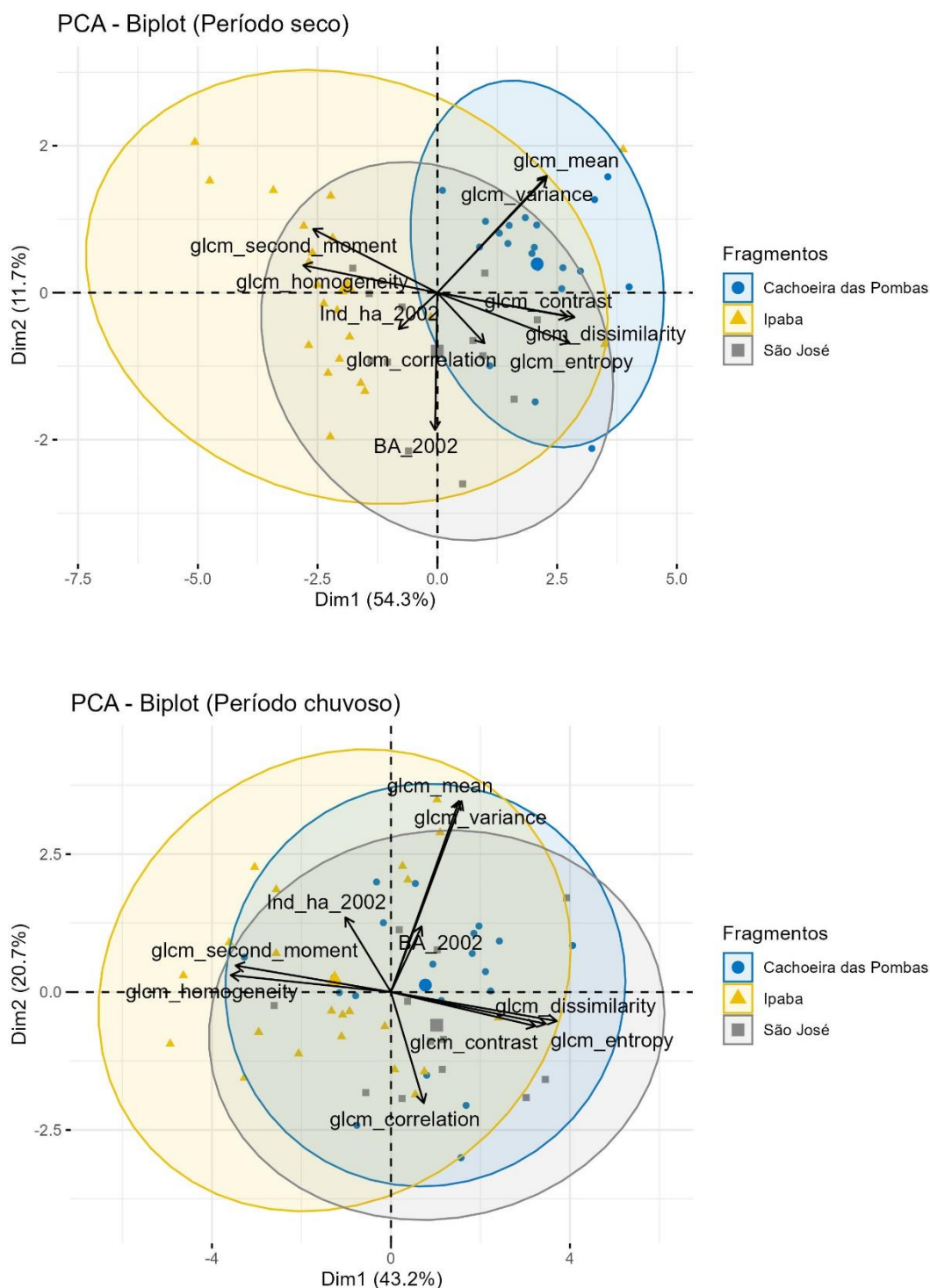
3.3 Resultado da análise de componentes principais

Na análise de componentes principais (PCA) com os descritores texturais no período seco, os primeiros componentes principais (CP1 e CP2) explicaram 54,3% e 11,7% da variância total, respectivamente (Figura 5). Os descritores `glcm_mean` e `glcm_variance` foram mais associadas ao fragmento Cachoeira das Pombas, indicado pela elipse e círculos em azul. O fragmento Ipaba, representado pela elipse e triângulos em amarelo, apresentou maior dispersão e associação com os descritores `glcm_second_moment` e `glcm_homogeneity`. Já o fragmento São José, indicado pela elipse e quadrados cinza, apresentou menor dispersão e associação com os descritores `glcm_contrast` e `glcm_dissimilarity`.

No período chuvoso, a CP1 e CP2 explicaram 43,2% e 20,7% da variância total, respectivamente. Neste período, ocorreu maior sobreposição entre os fragmentos Cachoeira das Pombas e São José. O fragmento Ipaba continuou com maior associação os descritores `glcm_second_moment` e `glcm_homogeneity`, mas com uma dispersão ligeiramente menor comparada ao período seco. O fragmento São José manteve a maior associação com os descritores `glcm_contrast` e `glcm_dissimilarity`. Os indivíduos infestados por lianas por hectare (`Ind_ha_2002`) e a área basal por hectare no ano de 2002 (`BA_2002`) apresentaram uma maior associação com os fragmentos São José e Ipaba nos períodos seco e chuvoso. Os boxplots dos atributos da floresta, descritores texturais e componentes principais selecionadas nos períodos seco e chuvoso podem ser visualizados nas Figuras C1 e C2 (APÊNDICE C).

As variâncias acumuladas das PCA nos períodos seco e chuvoso podem ser visualizadas nas Figuras C3 e C4, respectivamente. Para ambos os períodos, as 4 primeiras componentes representaram pelo menos 80% da variância acumulada e foram utilizadas como variáveis de entrada nos treinamentos das técnicas de aprendizado de máquina.

Figura 5 - Análise de Componentes Principais (PCA) das variáveis de entrada (descritores texturais, indivíduos infestados por lianas por hectare e área basal por hectare) nos fragmentos Cachoeira das Pombas, Ipaba e São José nos períodos seco e chuvoso.



Fonte: O autor (2024).

3.4 Avaliação das técnicas de aprendizagem de máquina

No treino das técnicas para previsão da biomassa do fuste, as RNA apresentaram os maiores valores de coeficiente de correlação (r_{yy}) próximos de 1, variando entre 0,9974 e 0,9999 (Tabela 3). Os valores de RMSE% e MAPE foram menores que o da RVS, variando entre 0,21 e 6,03%. As RVS também obtiveram altos valores de correlação no treino, variando de 0,9119 a 0,9847. No entanto, apresentaram maiores valores de RMSE (12,64 a 27,99%) e MAPE (10,34 a 17,74%) em comparação com as RNA. A RF apresentou os menores valores de correlação e maiores de RMSE% e MAPE em todas as abordagens dos períodos seco e chuvoso.

Na validação, a RVS na abordagem com modelo completo apresentou as melhores estatísticas no período seco, com uma correlação de 0,8596, RMSE% de 36,06% e MAPE de 22,83%. As RNA com abordagem modelo completo apresentou menor precisão, com correlação menor valor de correlação e maiores valores de RMSE% e MAPE.

No período chuvoso, a RVS apresentou a melhor estatística em relação a RNA na abordagem com modelo completo, com correlação de 0,8543, RMSE de 36,06% e MAPE de 22,83%. Assim como no treinamento, a RF apresentou os menores valores de correlação e maiores de RMSE% e MAPE em todas as abordagens nos períodos seco e chuvoso.

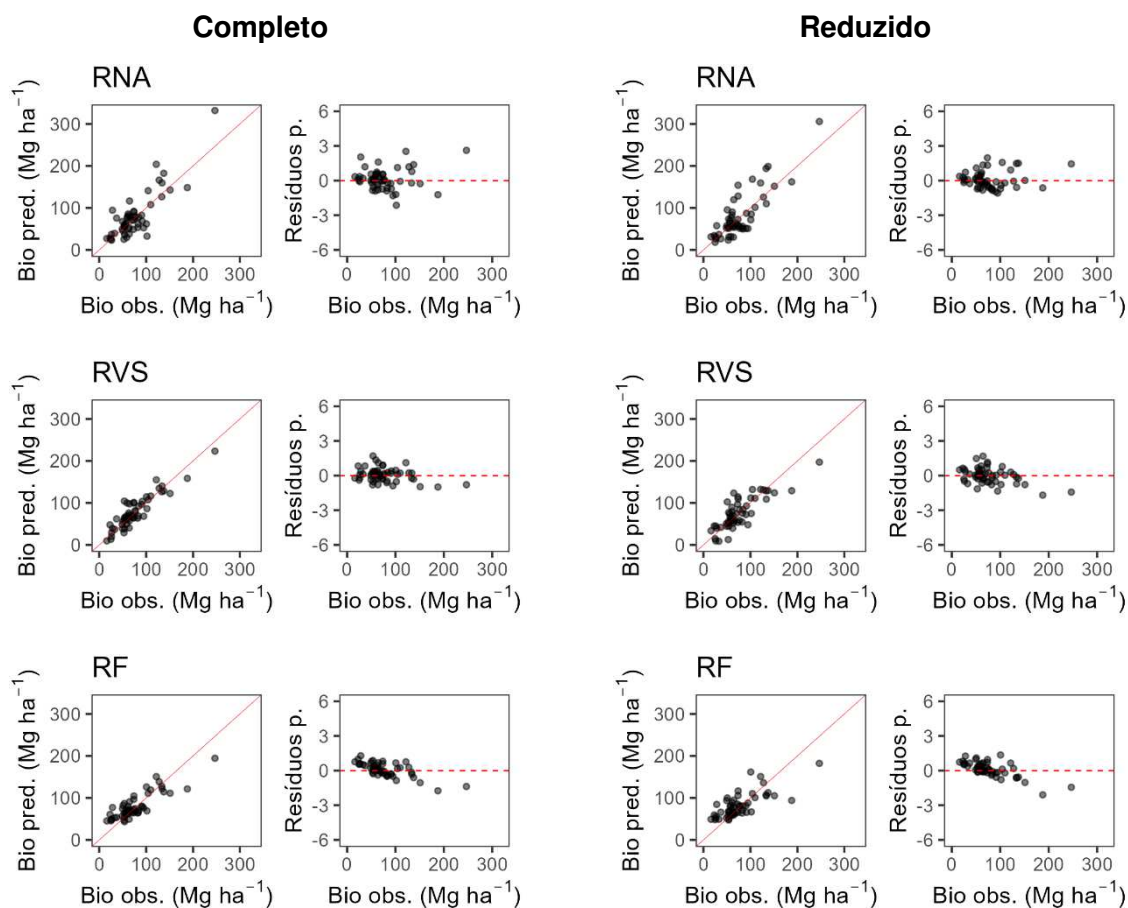
Avaliando a dispersão dos valores preditos e observados na validação de biomassa no fuste das árvores no período seco, a RVS apresentou a dispersão mais próxima da linha 1:1, tanto na abordagem com modelo completo como no reduzido (Figura 6). Os resíduos padronizados dessa técnica apresentaram a dispersão mais próxima do eixo zero em relação a RNA. No período chuvoso, a RNA e RVS apresentaram dispersão dos valores preditos próxima dos valores observados na abordagem com modelo completo (Figura 7). Na abordagem com modelo reduzido, a RNA foi a única técnica que não subestimou os maiores valores de biomassa.

Tabela 3 - Estatísticas de correlação (r_{yy}), raiz quadrada do erro quadrático médio, em porcentagem (RMSE%) e erro absoluto médio percentual (MAPE) das melhores configurações de cada técnica nos conjuntos de treino e validação com as imagens no período seco e chuvoso.

Conjunto	Período ¹	Abordagem	Modelo	r_{yy}	RMSE%	MAPE	Algoritmo ²	Função ³	Neurônios ⁴	Kernel ⁵	Árvores ⁶
Treino	Seco	Completo	RNA	0,9999	0,23	0,21	Rprop	Exponencial	25	-	-
			RVS	0,9847	12,64	13,36	-	-	-	Polinomial	-
			RF	0,8186	35,67	30,57	-	-	-	-	1000
		Reduzido	RNA	0,9981	4,61	5,40	SCG	Exponencial	8	-	-
			RVS	0,9119	25,57	15,67	-	-	-	Radial	-
			RF	0,7386	41,18	36,44	-	-	-	-	1000
	Chuvoso	Completo	RNA	0,9999	0,28	0,31	Rprop	Exponencial	26	-	-
			RVS	0,9420	27,99	10,34	-	-	-	Polinomial	-
			RF	0,8259	35,76	31,20	-	-	-	-	500
		Reduzido	RNA	0,9979	4,99	6,03	SCG	Exponencial	9	-	-
			RVS	0,9300	24,20	17,74	-	-	-	Radial	-
			RF	0,6971	43,56	39,98	-	-	-	-	2000
Validação	Seco	Completo	RNA	0,8471	39,17	30,54	Rprop	Tangente hiperbólica	24	-	-
			RVS	0,8596	36,06	22,83	-	-	-	Linear	-
			RF	0,7786	46,03	33,29	-	-	-	-	3000
		Reduzido	RNA	0,7399	49,12	33,10	Rprop	Logística	8	-	-
			RVS	0,7970	42,39	30,24	-	-	-	Linear	-
			RF	0,6302	54,14	39,31	-	-	-	-	2000
	Chuvoso	Completo	RNA	0,7957	43,73	31,30	Rprop	Logística	8	-	-
			RVS	0,8543	36,72	25,35	-	-	-	Linear	-
			RF	0,7859	47,64	34,59	-	-	-	-	2000
		Reduzido	RNA	0,6736	53,88	41,32	Rprop	Tangente hiperbólica	7	-	-
			RVS	0,7048	50,88	31,78	-	-	-	Linear	-
			RF	0,6462	54,88	39,79	-	-	-	-	1500

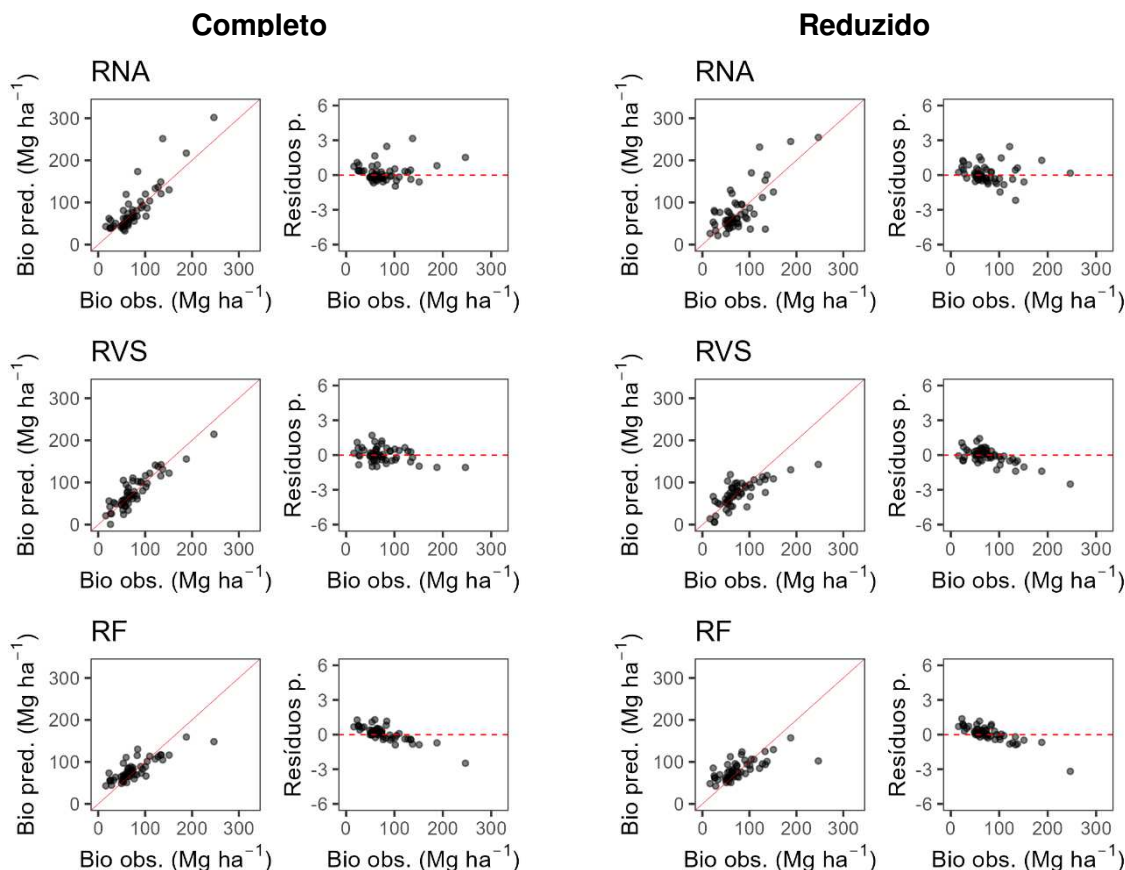
Em que: Rprop = *resilient propagation*; SCG = *scaled conjugate gradient*; RNA = rede neural artificial; RVS = regressão por vetores de suporte; RF = *random forest*. ¹ Período de aquisição da imagem. ² Algoritmo na camada oculta da RNA. ³ Função ativação na camada oculta da RNA. ⁴ Número de neurônios na camada oculta da RNA. ⁵ Função Kernel utilizada nas RVS. ⁶ Número de árvores de regressão na RF.

Figura 6 - Dispersão dos valores preditos e observados de biomassa do fuste e dos resíduos padronizados em função da biomassa observada utilizando as redes neurais artificiais (RNA) e regressão vetorial de suporte (RVS) nas abordagens com modelo completo e reduzido durante o período seco.



Fonte: O autor (2024).

Figura 7 - Dispersão dos valores preditos e observados de biomassa do fuste e dos resíduos padronizados em função da biomassa observada utilizando as redes neurais artificiais (RNA) e regressão vetorial de suporte (RVS) nas abordagens com modelo completo e reduzido durante o período chuvoso.



Fonte: O autor (2024).

4 DISCUSSÃO

No presente estudo, a biomassa do fuste de árvores em fragmentos da Mata Atlântica foi predita utilizando descritores texturais derivados da matriz de coocorrência em níveis de cinza (GLCM) e atributos da floresta como a área basal por hectare e número de indivíduos infestados por lianas por hectare em diferentes períodos sazonais (período seco e chuvoso).

Analisando os resultados da seleção do tamanho da janela no fragmento Cachoeira das Pombas, a semivariância apresentou valores semelhantes em ambos os períodos, seco e chuvoso, com um leve aumento durante o período chuvoso. Esse aumento pode ser atribuído à maior heterogeneidade da cobertura vegetal devido no período chuvoso o crescimento das árvores ser mais favorável, resultando em uma maior variabilidade de arquiteturas de copas de diferentes espécies (TRIPATHI et al.,

2020; OZKAN; DEMIREL, 2021). No fragmento de Ipaba, a semivariância foi maior no período seco, atingindo um patamar mais elevado em comparação ao período chuvoso, o que sugere uma maior homogeneidade estrutural durante a estação seca (SANTOS et al., 2023; ZHANG et al., 2023). Por outro lado, no fragmento de São José, a semivariância atingiu seu maior valor no período chuvoso, indicando uma maior variabilidade estrutural que pode ser atribuído ao crescimento vigoroso da vegetação durante este período (HUNTER et al., 2015; SONG et al., 2024).

4.1 Descritores texturais derivados da GLCM

Analisando os descritores texturais extraídos da GLCM durante o período chuvoso no fragmento Cachoeira das Pombas, valores de média (*glcm_mean*), variância (*glcm_variance*), dissimilaridade (*glcm_dissimilarity*) e segundo momento angular (*glcm_second_moment*) aumentaram em comparação ao período seco. Esse aumento pode ser atribuído à maior heterogeneidade e complexidade da estrutura florestal durante a estação chuvosa, quando a disponibilidade de água favorece o crescimento tanto das árvores quanto das lianas (MOHAMMADPOUR; VIEGAS; VIEGAS, 2022). O aumento da presença de lianas e o crescimento das árvores podem resultar em uma textura mais variada e menos uniforme, refletida nos altos valores desses descritores (SCHNITZER, 2015; ESTRADA-VILLEGAS et al., 2022). O aumento do segundo momento angular, que mede a uniformidade da textura, indica uma maior repetição de padrões de pares de pixels, sugerindo áreas com vegetação densa e homogênea (JI et al., 2019; IQBAL et al., 2021).

No fragmento Ipaba, durante o período chuvoso, o aumento dos valores de contraste, dissimilaridade e correlação (*glcm_correlation*) indicam que, apesar do crescimento da vegetação durante a estação chuvosa, a textura da imagem se tornou mais heterogênea e complexa (ZUBAIR; ALO, 2024). O aumento do contraste e da dissimilaridade pode ser interpretado como um reflexo das diferenças entre as áreas cobertas por árvores e aquelas dominadas por lianas, que competem por recursos e alteram a estrutura florestal (JONES et al., 2017). A redução da variância e do segundo momento angular sugere uma menor uniformidade e repetição de padrões texturais, indicando uma estrutura florestal mais complexa e menos homogênea (LAN; LIU, 2018; ANAND et al., 2023).

Para o fragmento São José, durante o período chuvoso, a variância foi o único

descriptor que aumentou em relação ao período seco. Esse aumento indica uma maior variação nos níveis de cinza, possivelmente devido à diversidade estrutural da vegetação durante a estação chuvosa (NUMBISI; VAN COILLIE; DE WULF, 2019). A redução da homogeneidade pode ser atribuída à menor presença de manchas homogêneas, enquanto a redução do segundo momento angular indica uma menor repetição de padrões texturais (HU; GHAMISI; ZHU, 2018; JI et al., 2019).

Considerando que as florestas estacionais semidecíduais da Mata Atlântica passam por mudanças entre os períodos seco e chuvoso, é importante levar em conta a influência das lianas nesses padrões texturais. Durante o período seco, algumas árvores perdem suas folhas, o que pode aumentar a abundância de lianas que aproveitam a maior exposição à luz solar (SCHNITZER, 2015). Isso resulta em uma maior competição por recursos e, conseqüentemente, em uma textura mais complexa e variável.

4.2 Análise de componentes principais

Na análise de componentes principais (PCA) durante o período seco, os descritores `glcm_mean` e `glcm_variance` foram mais associadas ao fragmento Cachoeira das Pombas. Essa maior variância pode ser atribuída à presença de grandes manchas vegetais e à diversidade estrutural da floresta durante o período seco, quando algumas árvores perdem suas folhas e as lianas proliferam, aumentando a heterogeneidade da cobertura vegetal (GALLARDO-CRUZ et al., 2012).

O fragmento Ipaba apresentou maior dispersão e relação com as variáveis `glcm_second_moment` e `glcm_homogeneity`. O segundo momento angular mede a uniformidade da textura e altos valores desta métrica indicam grande repetição de padrões texturais (MATHER; TSO, 2009). A associação com a homogeneidade sugere que, apesar da maior dispersão, existem grandes áreas com padrões texturais consistentes, provavelmente resultantes da predominância de certas espécies de árvores ou da distribuição das lianas (GUPTA; GONDHI; LEHANA, 2020; LI et al., 2022).

Já o fragmento São José apresentou associação com os descritores `glcm_contrast` e `glcm_dissimilarity`. O contraste e dissimilaridade indicam que ocorre maior variação entre os níveis de cinza de pixels vizinhos (MATHER; TSO, 2009). A menor dispersão

dos valores desse fragmento sugere uma textura mais uniforme, que pode ser devido a menor variação nas copas dentro do dossel (JACKSON et al., 2023).

No período chuvoso, a maior sobreposição entre os fragmentos Cachoeira das Pombas e São José sugere uma maior similaridade na textura durante a estação chuvosa. O fragmento Ipaba continuou com maior associação às variáveis *glcm_second_moment* e *glcm_homogeneity*, mas com uma dispersão ligeiramente menor comparada ao período seco, o que indica uma leve diminuição na variabilidade textural, possivelmente devido a um crescimento vegetal mais homogêneo (IQBAL et al., 2021).

Os indivíduos infestados por lianas por hectare (*Ind_ha_2002*) e a área basal por hectare no ano de 2002 (*BA_2002*) apresentaram uma maior associação com os fragmentos São José e Ipaba nos períodos seco e chuvoso. A maior presença de lianas pode ter contribuído para a complexidade textural observada, principalmente em São José, onde o contraste e a dissimilaridade foram maiores. As lianas aumentam a heterogeneidade estrutural ao competir com as árvores por luz e espaço, criando variações na textura da vegetação (MOHAMMADPOUR; VIEGAS; VIEGAS, 2022).

4.3 Avaliação das técnicas

Na avaliação das técnicas na validação, a RVS na abordagem com modelo completo apresentou as melhores estatísticas no período seco. Esses resultados destacam a robustez da RVS em estimativas no período seco, em que a variação estrutural da vegetação pode ser mais previsível. As RNA, no entanto, apresentaram menor precisão na validação, com menores valores de correlação e maiores valores de RMSE% e MAPE, o que pode ser atribuído à sua possível sobreajuste durante o treinamento, levando a um desempenho inferior ao da RVS (SANTI et al., 2020; DINOV, 2023).

No período chuvoso, a RVS continuou a apresentar a melhor estatística em comparação com as RNA na abordagem com modelo completo. Esse melhor desempenho da RVS pode ser devido a sua capacidade de lidar com a variabilidade adicional introduzida pelas condições no período chuvoso e pelo crescimento da vegetação (SU et al., 2020; LI et al., 2022). A RF apresentou os menores valores de correlação e os maiores valores de RMSE% e MAPE em todas as abordagens e períodos, indicando sua limitação na previsão precisa da biomassa do fuste com o

conjunto de variáveis utilizada.

Ao avaliar a dispersão dos valores preditos e observados na validação da biomassa do fuste das árvores no período seco, a RVS apresentou a dispersão mais próxima da linha 1:1, tanto na abordagem com modelo completo quanto no modelo reduzido. Os resíduos padronizados dessa técnica mostraram a dispersão mais próxima do eixo zero em comparação com as RNA, confirmando a maior precisão na predição da biomassa. No período chuvoso, tanto as RNA quanto as RVS apresentaram dispersão dos valores preditos próxima dos valores observados na abordagem com modelo completo. No entanto, na abordagem com modelo reduzido, as RNA foram as únicas técnicas que não subestimaram os maiores valores de biomassa, indicando sua capacidade de capturar variações extremas na biomassa durante o período chuvoso.

As diferenças observadas entre os períodos seco e chuvoso podem ser atribuídas a várias causas relacionadas à vegetação. No período seco, e à perda de folhas de algumas espécies arbóreas em florestas estacionais semidecíduais e a redução das chuvas pode causar o aumento da abundância de lianas (CAI; SCHNITZER; BONGERS, 2009; BECKNELL et al., 2022). As lianas, por sua vez, competem intensamente com as árvores por recursos como luz, água e nutrientes, afetando a estrutura da vegetação e, conseqüentemente, as estimativas de biomassa (ZHANG et al., 2020; AMITRANO et al., 2023; LUO et al., 2024). No período chuvoso, a maior disponibilidade de água favorece o crescimento tanto das árvores quanto das lianas, aumentando a complexidade estrutural da floresta e a variabilidade dos dados de entrada, o que pode explicar as diferenças de desempenho das técnicas de previsão observadas (SCHNITZER, 2015; MARSHALL et al., 2020; VARGAS; GROMBONE-GUARATINI; MORELLATO, 2021).

No geral, a escolha da técnica de previsão deve considerar as condições específicas do ecossistema e as variações sazonais. As RNA mostraram-se eficazes no treinamento, enquanto as RVS apresentaram uma melhor robustez na validação, principalmente em condições secas. A inclusão de variáveis texturais e a consideração dos indivíduos infestados por lianas podem melhorar a precisão dos modelos de estimativa de biomassa, proporcionando uma compreensão mais detalhada dos fragmentos florestais. Esses resultados destacam a importância de adaptar os modelos às condições específicas de cada período, gerando estimativas mais robustas e eficazes para a conservação e manejo sustentável dos fragmentos

de Mata Atlântica.

5 CONCLUSÃO

A abordagem com modelo completo, utilizando PCA, atributos da floresta e texturais, é indicada para a predição da biomassa do fuste de árvores para o período de 10 anos (2002-2012).

A RVS é a melhor técnica para predição de biomassa do fuste nos períodos seco e chuvoso.

O período seco é o mais indicado para ser realizado a predição de biomassa em fragmentos de Mata Atlântica utilizando descritores texturais e atributos da floresta.

REFERÊNCIAS

AGUIRRE-GUTIÉRREZ, J.; RIFAI, S.; SHENKIN, A.; OLIVERAS, I.; BENTLEY, L. P.; SVÁTEK, M.; GIRARDIN, C. A. J.; BOTH, S.; RIUTTA, T.; BERENQUER, E.; KISSLING, W. D.; BAUMAN, D.; RAAB, N.; MOORE, S.; FARFAN-RIOS, W.; FIGUEIREDO, A. E. S.; REIS, S. M.; NDONG, J. E.; ONDO, F. E.; N'SSI BENGONE, N.; MIHINDOU, V.; MORAES DE SEIXAS, M. M.; ADU-BREDU, S.; ABERNETHY, K.; ASNER, G. P.; BARLOW, J.; BURSLEM, D. F. R. P.; COOMES, D. A.; CERNUSAK, L. A.; DARGIE, G. C.; ENQUIST, B. J.; EWERS, R. M.; FERREIRA, J.; JEFFERY, K. J.; JOLY, C. A.; LEWIS, S. L.; MARIMON-JUNIOR, B. H.; MARTIN, R. E.; MORANDI, P. S.; PHILLIPS, O. L.; QUESADA, C. A.; SALINAS, N.; SCHWANTES MARIMON, B.; SILMAN, M.; TEH, Y. A.; WHITE, L. J. T.; MALHI, Y. Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data. **Remote Sensing of Environment**, v. 252, p. 112122, 1 jan. 2021.

ALI HUSSIN, Y. Assessment and Modelling of Forest Biomass and Carbon Stock and Sequestration Using Various Remote Sensing Sensor Systems. *Em*: SURATMAN, M. N. **Concepts and Applications of Remote Sensing in Forestry**. Singapore: Springer Nature, 2022. p. 75–95.

AMARO, M. A.; SOARES, C. P. B.; SOUZA, A. L. D.; LEITE, H. G.; SILVA, G. F. D. Estoque volumétrico, de biomassa e de carbono em uma floresta estacional semidecidual em viçosa, minas gerais. **Revista Árvore**, v. 37, n. 5, p. 849–857, out. 2013.

AMITRANO, D.; GIACCO, G.; MARRONE, S.; PASCARELLA, A. E.; RIGIROLI, M.; SANSONE, C. Forest Aboveground Biomass Estimation Using Machine Learning Ensembles: Active Learning Strategies for Model Transfer and Field Sampling Reduction. **Remote Sensing**, v. 15, n. 21, p. 5138, jan. 2023.

ANAND, R.; SHANTHI, T.; SABEENIAN, R. S.; VENI, S. GLCM Feature-Based Texture Image Classification Using Machine Learning Algorithms. *Em*: KUMAR, B. V.; SIVAKUMAR, P.; SURENDIRAN, B.; DING, J. **Smart Computer Vision**. Cham: Springer International Publishing, 2023. p. 103–125.

BALLING, J.; HEROLD, M.; REICHE, J. How textural features can improve SAR-based tropical forest disturbance mapping. **International Journal of Applied Earth Observation and Geoinformation**, v. 124, p. 103492, 1 nov. 2023.

BECKNELL, J. M.; VARGAS G., G.; WRIGHT, L. A.; WOODS, N.-F.; MEDVIGY, D.; POWERS, J. S. Increasing Liana Abundance and Associated Reductions in Tree Growth in Secondary Seasonally Dry Tropical Forest. **Frontiers in Forests and Global Change**, v. 5, p. 838357, 18 abr. 2022.

BERGMEIR, C.; BENÍTEZ, J. M. Neural Networks in *R* Using the Stuttgart Neural Network Simulator: **RSNNS**. **Journal of Statistical Software**, v. 46, n. 7, 2012. Disponível em: <<http://www.jstatsoft.org/v46/i07/>>.

BORSAH, A. A.; NAZEER, M.; WONG, M. S. LIDAR-Based Forest Biomass Remote Sensing: A Review of Metrics, Methods, and Assessment Criteria for the Selection of Allometric Equations. **Forests**, v. 14, n. 10, p. 2095, out. 2023.

CAI, Z.-Q.; SCHNITZER, S. A.; BONGERS, F. Seasonal Differences in Leaf-Level Physiology Give Lianas a Competitive Advantage over Trees in a Tropical Seasonal Forest. **Oecologia**, v. 161, n. 1, p. 25–33, 1 ago. 2009.

CAWLEY, G. C.; TALBOT, N. L. C. Efficient Leave-One-out Cross-Validation of Kernel Fisher Discriminant Classifiers. **Pattern Recognition**, v. 36, n. 11, p. 2585–2592, nov. 2003.

CSILLIK, O.; KUMAR, P.; ASNER, G. P. Challenges in Estimating Tropical Forest Canopy Height from Planet Dove Imagery. **Remote Sensing**, v. 12, n. 7, p. 1160, jan. 2020.

DELICIELLOS, A. C.; PREVEDELLO, J. A.; FIGUEIREDO, M. S. L.; WEBER, M. M.; LORINI, M. L. Species Richness and Endemism of Marsupials in the Atlantic Forest: Spatial Patterns and Drivers. *Em*: CÁCERES, N. C.; DICKMAN, C. R. **American and Australasian Marsupials: An Evolutionary, Biogeographical, and Ecological Approach**. Cham: Springer International Publishing, 2022. p. 1–21.

DINOV, I. D. Model Performance Assessment, Validation, and Improvement. *Em*: DINOV, I. D. **Data Science and Predictive Analytics: Biomedical and Health Applications using R**. Cham: Springer International Publishing, 2023. p. 477–531.

DISTANTE, A.; DISTANTE, C. Texture Analysis. *Em*: DISTANTE, A.; DISTANTE, C. **Handbook of Image Processing and Computer Vision: Volume 3: From Pattern to Object**. Cham: Springer International Publishing, 2020. p. 261–314.

ECK, J.; KAAS, A. L.; MULDER, J. L.; HAUSFELD, L.; KOURTZI, Z.; GOEBEL, R. The Effect of Task Instruction on Haptic Texture Processing: The Neural Underpinning of Roughness and Spatial Density Perception. **Cerebral Cortex**, v. 26, n. 1, p. 384–401, 1 jan. 2016.

EDIRIWEERA, S.; PATHIRANA, S.; DANAHER, T.; NICHOLS, D. Estimating Above-Ground Biomass by Fusion of LiDAR and Multispectral Data in Subtropical Woody Plant Communities in Topographically Complex Terrain in North-Eastern Australia. **Journal of Forestry Research**, v. 25, n. 4, p. 761–771, 1 dez. 2014.

ESTRADA-VILLEGAS, S.; PEDRAZA NARVAEZ, S. S.; SANCHEZ, A.; SCHNITZER, S. A. Lianas Significantly Reduce Tree Performance and Biomass Accumulation Across Tropical Forests: A Global Meta-Analysis. **Frontiers in Forests and Global Change**, v. 4, p. 812066, 11 fev. 2022.

FERREIRA, M. P.; WAGNER, F. H.; ARAGÃO, L. E. O. C.; SHIMABUKURO, Y. E.; DE SOUZA FILHO, C. R. Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and texture analysis. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 149, p. 119–131, 1 mar. 2019.

GALLARDO-CRUZ, J. A.; MEAVE, J. A.; GONZÁLEZ, E. J.; LEBRIJA-TREJOS, E. E.; ROMERO-ROMERO, M. A.; PÉREZ-GARCÍA, E. A.; GALLARDO-CRUZ, R.; HERNÁNDEZ-STEFANONI, J. L.; MARTORELL, C. Predicting Tropical Dry Forest Successional Attributes from Space: Is the Key Hidden in Image Texture? **PLOS ONE**, v. 7, n. 2, p. e30506, 20 fev. 2012.

GUPTA, C.; GONDHI, N. K.; LEHANA, P. K. Gray Level Co-Occurrence Matrix (GLCM) Parameters Analysis for Pyoderma Image Variants. **Journal of Computational and Theoretical Nanoscience**, v. 17, n. 1, p. 353–358, 1 jan. 2020.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features for Image Classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. SMC-3, n. 6, p. 610–621, nov. 1973.

HEMATI, M.; MAHDIANPARI, M.; SHIRI, H.; MOHAMMADIMANESH, F. Integrating SAR and Optical Data for Aboveground Biomass Estimation of Coastal Wetlands Using Machine Learning: Multi-Scale Approach. **Remote Sensing**, v. 16, n. 5, p. 831, jan. 2024.

HU, J.; GHAMISI, P.; ZHU, X. X. Feature Extraction and Selection of Sentinel-1 Dual-Pol Data for Global-Scale Local Climate Zone Classification. **ISPRS International Journal of Geo-Information**, v. 7, n. 9, p. 379, set. 2018.

HUNG, C.-C.; SONG, E.; LAN, Y. **Image Texture Analysis: Foundations, Models and Algorithms**. Cham: Springer International Publishing, 2019.

HUNTER, M. O.; KELLER, M.; MORTON, D.; COOK, B.; LEFSKY, M.; DUCEY, M.; SALESKA, S.; JR, R. C. de O.; SCHIETTI, J. Structural Dynamics of Tropical Moist Forest Gaps. **PLOS ONE**, v. 10, n. 7, p. e0132144, 13 jul. 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, I. **Manual técnico da vegetação brasileira**. Rio de Janeiro: IBGE, 2012.

IQBAL, N.; MUMTAZ, R.; SHAFI, U.; ZAIDI, S. M. H. Gray Level Co-Occurrence Matrix (GLCM) Texture Based Crop Classification Using Low Altitude Remote Sensing Platforms. **PeerJ Computer Science**, v. 7, p. e536, 19 maio 2021.

JACKSON, C. M.; ADAM, E.; ATIF, I.; MAHBOOB, M. A. Feature Extraction and Classification of Canopy Gaps Using GLCM- and MLBP-Based Rotation-Invariant Feature Descriptors Derived from WorldView-3 Imagery. **Geomatics**, v. 3, n. 1, p. 250–265, mar. 2023.

JI, M.; LIU, L.; DU, R.; BUCHROITHNER, M. F. A Comparative Study of Texture and Convolutional Neural Network Features for Detecting Collapsed Buildings After Earthquakes Using Pre- and Post-Event Satellite Imagery. **Remote Sensing**, v. 11, n. 10, p. 1202, jan. 2019.

JONES, I. L.; PERES, C. A.; BENCHIMOL, M.; BUNNEFELD, L.; DENT, D. H. Woody Lianas Increase in Dominance and Maintain Compositional Integrity across an Amazonian Dam-Induced Fragmented Landscape. **PLOS ONE**, v. 12, n. 10, p. e0185527, 17 out. 2017.

KUPLICH, T. M.; CURRAN, P. J.; ATKINSON, P. M. Relating SAR image texture to the biomass of regenerating tropical forests. **International Journal of Remote Sensing**, v. 26, n. 21, p. 4829–4854, 10 nov. 2005.

LAN, Z.; LIU, Y. Study on Multi-Scale Window Determination for GLCM Texture Description in High-Resolution Remote Sensing Image Geo-Analysis Supported by GIS and Domain Knowledge. **ISPRS International Journal of Geo-Information**, v. 7, n. 5, p. 175, maio 2018.

LÊ, S.; JOSSE, J.; HUSSON, F. **FactoMineR**: An R Package for Multivariate Analysis. **Journal of Statistical Software**, v. 25, n. 1, 2008. Disponível em: <<http://www.jstatsoft.org/v25/i01/>>.

LI, B.; ZHANG, Y.; LUAN, F.; YUAN, Z.; ALI, A.; CHU, C.; BIN, Y. Habitat Conditions and Tree Species Shape Liana Distribution in a Subtropical Forest. **Forests**, v. 13, n. 9, p. 1358, set. 2022a.

LI, N.; NING, Z.; CHEN, M.; WU, D.; HAO, C.; ZHANG, D.; BAI, R.; LIU, H.; CHEN, X.; LI, W.; ZHANG, W.; CHEN, Y.; LI, Q.; ZHANG, L. Satellite and Machine Learning Monitoring of Optically Inactive Water Quality Variability in a Tropical River. **Remote Sensing**, v. 14, n. 21, p. 5466, jan. 2022b.

LIAW, A.; WIENER, M. Classification and Regression by randomForest. v. 2, 2002.

LUO, M.; ANEES, S. A.; HUANG, Q.; QIN, X.; QIN, Z.; FAN, J.; HAN, G.; ZHANG, L.; SHAFRI, H. Z. M. Improving Forest Above-Ground Biomass Estimation by Integrating Individual Machine Learning Models. **Forests**, v. 15, n. 6, p. 975, jun. 2024.

MARINHO, N. de T.; BURNHAM, R. J.; MAEDA, E. E.; NUNES, M. H.; LAURANCE, W. F.; LAURANCE, S. G.; ANDRADE, A. C. S. de; CAMARGO, J. L. C. **The influence of forest structure on the abundance, biomass, and composition of lianas in tropical forest fragments** 20 nov. 2023. Disponível em: <<https://www.researchsquare.com/article/rs-3410725/v1>>.

MARSHALL, A. R.; PLATTS, P. J.; CHAZDON, R. L.; SEKI, H.; CAMPBELL, M. J.; PHILLIPS, O. L.; GEREAU, R. E.; MARCHANT, R.; LIANG, J.; HERBOHN, J.; MALHI, Y.; PFEIFER, M. Conceptualising the Global Forest Response to Liana Proliferation.

Frontiers in Forests and Global Change, v. 3, p. 35, 24 abr. 2020.

MASSEY, R.; BERNER, L. T.; FOSTER, A. C.; GOETZ, S. J.; VEPAKOMMA, U. Remote Sensing Tools for Monitoring Forests and Tracking Their Dynamics. *Em: GIRONA, M. M.; MORIN, H.; GAUTHIER, S.; BERGERON, Y. **Boreal Forests in the Face of Climate Change: Sustainable Management***. Cham: Springer International Publishing, 2023. p. 637–655.

MATHER, P.; TSO, B. **Classification Methods for Remotely Sensed Data**. 2. ed. Boca Raton: CRC Press, 2009. 376 p.

MEUNIER, F.; VAN DER HEIJDEN, G. M. F.; SCHNITZER, S. A.; DE DEURWAERDER, H. P. T.; VERBEECK, H. Lianas Significantly Reduce Aboveground and Belowground Carbon Storage: A Virtual Removal Experiment. **Frontiers in Forests and Global Change**, v. 4, p. 663291, 22 jun. 2021.

MOHAMMADPOUR, P.; VIEGAS, D. X.; VIEGAS, C. Vegetation Mapping with Random Forest Using Sentinel 2 and GLCM Texture Feature—A Case Study for Lousã Region, Portugal. **Remote Sensing**, v. 14, n. 18, p. 4585, jan. 2022.

NUMBISI, F. N.; VAN COILLIE, F. M. B.; DE WULF, R. Delineation of Cocoa Agroforests Using Multiseason Sentinel-1 SAR Images: A Low Grey Level Range Reduces Uncertainties in GLCM Texture-Based Mapping. **ISPRS International Journal of Geo-Information**, v. 8, n. 4, p. 179, abr. 2019.

OZKAN, U. Y.; DEMIREL, T. The influence of window size on remote sensing-based prediction of forest structural variables. **Ecological Processes**, v. 10, n. 1, p. 60, 8 set. 2021.

PLOTON, P.; PÉLISSIER, R.; BARBIER, N.; PROISY, C.; RAMESH, B. R.; COUTERON, P. Canopy Texture Analysis for Large-Scale Assessments of Tropical Forest Stand Structure and Biomass. *Em: LOWMAN, M.; DEVY, S.; GANESH, T. **Treetops at Risk: Challenges of Global Canopy Ecology and Conservation***. New York, NY: Springer, 2013. p. 237–245.

PORTILLO, J. T. da M.; AZEVEDO, J. A. R.; BARBO, F. E.; SAWAYA, R. J. Opposite latitudinal gradients for species richness and phylogenetic diversity of endemic snakes in the Atlantic Forest. **Current Zoology**, p. zoad032, 19 jul. 2023.

RAJ, A.; JHARIYA, M. K.; KHAN, N.; NEMA, S.; DEVI, A.; GHANSHYAM. Effect of Deforestation and Forest Fragmentation on Ecosystem Services. *Em: PANWAR, P.; SHUKLA, G.; BHAT, J. A.; CHAKRAVARTY, S. **Land Degradation Neutrality: Achieving SDG 15 by Forest Management***. Singapore: Springer Nature, 2022. p. 25–41.

RIBEIRO, M. C.; MARTENSEN, A. C.; METZGER, J. P.; TABARELLI, M.; SCARANO, F.; FORTIN, M.-J. The Brazilian Atlantic Forest: A Shrinking Biodiversity Hotspot. *Em: ZACHOS, F. E.; HABEL, J. C. **Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas***. Berlin, Heidelberg: Springer, 2011. p. 405–434.

RODRIGUES, A. C.; VILLA, P. M.; FERREIRA-JÚNIOR, W. G.; SCHAEFER, C. E. R. G.; NERI, A. V. Effects of topographic variability and forest attributes on fine-scale soil

fertility in late-secondary succession of Atlantic Forest. **Ecological Processes**, v. 10, n. 1, p. 62, 25 set. 2021.

SA, R.; FAN, W. Forest Structure Mapping of Boreal Coniferous Forests Using Multi-Source Remote Sensing Data. **Remote Sensing**, v. 16, n. 11, p. 1844, jan. 2024.

SAINUDDIN, F. V.; MALEK, G.; RAJWADI, A.; NAGAR, P. S.; ASOK, S. V.; REDDY, C. S. Estimating Above-Ground Biomass of the Regional Forest Landscape of Northern Western Ghats Using Machine Learning Algorithms and Multi-Sensor Remote Sensing Data. **Journal of the Indian Society of Remote Sensing**, v. 52, n. 4, p. 885–902, 1 abr. 2024.

SANAM, H.; THOMAS, A. A.; KUMAR, A. P.; LAKSHMANAN, G. Multi-Sensor Approach for the Estimation of Above-Ground Biomass of Mangroves. **Journal of the Indian Society of Remote Sensing**, v. 52, n. 4, p. 903–916, 1 abr. 2024.

SÁNCHEZ-AZOFEIFA, G. A.; CASTRO, K.; WRIGHT, S. J.; GAMON, J.; KALACSKA, M.; RIVARD, B.; SCHNITZER, S. A.; FENG, J. L. Differences in leaf traits, leaf internal structure, and spectral reflectance between two communities of lianas and trees: Implications for remote sensing in tropical environments. 2009. Disponível em: <<http://repository.si.edu/xmlui/handle/10088/12163>>.

SANTI, E.; PALOSCIA, S.; PETTINATO, S.; CUOZZO, G.; PADOVANO, A.; NOTARNICOLA, C.; ALBINET, C. Machine-Learning Applications for the Retrieval of Forest Biomass from Airborne P-Band SAR Data. **Remote Sensing**, v. 12, n. 5, p. 804, jan. 2020.

SANTOS, H. K. V.; BORGES DE LIMA, R.; FIGUEIREDO DE SOUZA, R. L.; CARDOSO, D.; MOONLIGHT, P. W.; TEIXEIRA SILVA, T.; PEREIRA DE OLIVEIRA, C.; ALVES JÚNIOR, F. T.; VEENENDAAL, E.; PAGANUCCI DE QUEIROZ, L.; RODRIGUES, P. M. S.; DOS SANTOS, R. M.; SARKINEN, T.; DE PAULA, A.; BARRETO-GARCIA, P. a. B.; PENNINGTON, T.; PHILLIPS, O. L. Spatial Distribution of Aboveground Biomass Stock in Tropical Dry Forest in Brazil. **iForest - Biogeosciences and Forestry**, v. 16, n. 2, p. 116, 2023.

SCHNITZER, S. A. The Contribution of Lianas to Forest Ecology, Diversity, and Dynamics. *Em*: PARTHASARATHY, N. **Biodiversity of Lianas**. Cham: Springer International Publishing, 2015. p. 149–160.

SONG, Y.; LI, Y.; ZOU, Y.; WANG, R.; LIANG, Y.; XU, S.; HE, Y.; YU, X.; WU, W. Synergizing Multiple Machine Learning Techniques and Remote Sensing for Advanced Landslide Susceptibility Assessment: A Case Study in the Three Gorges Reservoir Area. **Environmental Earth Sciences**, v. 83, n. 8, p. 227, 2 abr. 2024.

SOUZA, A. L. de; BOINA, A.; SOARES, C. P. B.; VITAL, B. R.; GASPAR, R. de O.; LANA, J. M. de. Estrutura fitossociológica, estoques de volume, biomassa, carbono e dióxido de carbono em Floresta Estacional Semidecidual. **Revista Árvore**, v. 36, p. 169–179, fev. 2012.

SU, H.; SHEN, W.; WANG, J.; ALI, A.; LI, M. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. **Forest Ecosystems**, v. 7, n. 1, p. 64, 26 nov. 2020.

TIAN, L.; WU, X.; TAO, Y.; LI, M.; QIAN, C.; LIAO, L.; FU, W. Review of Remote Sensing-Based Methods for Forest Aboveground Biomass Estimation: Progress, Challenges, and Prospects. **Forests**, v. 14, n. 6, p. 1086, jun. 2023.

TIAN, X.; LI, J.; ZHANG, F.; ZHANG, H.; JIANG, M. Forest Aboveground Biomass Estimation Using Multisource Remote Sensing Data and Deep Learning Algorithms: A Case Study over Hangzhou Area in China. **Remote Sensing**, v. 16, n. 6, p. 1074, jan. 2024.

TORRES, C. M. M. E.; JACOVINE, L. A. G.; OLIVEIRA NETO, S. N. de; SOUZA, A. L. de; CAMPOS, R. A.; SCHETTINI, B. L. S. Análise Fitossociológica e Valor de Importância em Carbono para uma Floresta Estacional Semidecidual. **Floresta e Ambiente**, v. 24, p. e00099714, 17 ago. 2017.

TRIPATHI, S.; BHADOURIA, R.; SRIVASTAVA, P.; DEVI, R. S.; CHATURVEDI, R.; RAGHUBANSHI, A. S. Effects of light availability on leaf attributes and seedling growth of four tree species in tropical dry forest. **Ecological Processes**, v. 9, n. 1, p. 2, 6 jan. 2020.

VARGAS, B. da C.; GROMBONE-GUARATINI, M. T.; MORELLATO, L. P. C. Lianas Research in the Neotropics: Overview, Interaction with Trees, and Future Perspectives. **Trees**, v. 35, n. 2, p. 333–345, abr. 2021.

VIERA, M.; CALDEIRA, M. V. W.; ROVANI, F. F. M.; CASTRO, K. C.; VIERA, M.; CALDEIRA, M. V. W.; ROVANI, F. F. M.; CASTRO, K. C. Ecological and Environmental Aspects of Nutrient Cycling in the Atlantic Forest, Brazil. *Em: Tropical Forests - The Challenges of Maintaining Ecosystem Services while Managing the Landscape*. [s.l.] IntechOpen, 2016.

WILSON, M. C.; CHEN, X.-Y.; CORLETT, R. T.; DIDHAM, R. K.; DING, P.; HOLT, R. D.; HOLYOAK, M.; HU, G.; HUGHES, A. C.; JIANG, L.; LAURANCE, W. F.; LIU, J.; PIMM, S. L.; ROBINSON, S. K.; RUSSO, S. E.; SI, X.; WILCOVE, D. S.; WU, J.; YU, M. Habitat Fragmentation and Biodiversity Conservation: Key Findings and Future Challenges. **Landscape Ecology**, v. 31, n. 2, p. 219–227, 1 fev. 2016.

ZANNE, A. E.; LOPEZ-GONZALEZ, G.; COOMES, D. A.; ILIC, J.; JANSEN, S.; LEWIS, S. L.; MILLER, R. B.; SWENSON, N. G.; WIEMANN, M. C.; CHAVE, J. **Data from: Towards a worldwide wood economics spectrum** Dryad, 4 fev. 2009. Disponível em: <<https://datadryad.org/stash/dataset/doi:10.5061/dryad.234>>.

ZHANG, X.; CHEN, G.; LIU, C.; FAN, Q.; LI, W.; WU, Y.; XU, H.; OU, G. Spatial Effects Analysis on Individual-Tree Aboveground Biomass in a Tropical Pinus Kesiya Var. Langbianensis Natural Forest in Yunnan, Southwestern China. **Forests**, v. 14, n. 6, p. 1177, jun. 2023.

ZHANG, Y.; LIANG, S.; YANG, L. A Review of Regional and Global Gridded Forest Biomass Datasets. **Remote Sensing**, v. 11, n. 23, p. 2744, jan. 2019.

ZHANG, Y.; MA, J.; LIANG, S.; LI, X.; LI, M. An Evaluation of Eight Machine Learning Regression Algorithms for Forest Aboveground Biomass Estimation from Multiple Satellite Data Products. **Remote Sensing**, v. 12, n. 24, p. 4015, jan. 2020.

ZUBAIR, A. R.; ALO, O. A. **Grey Level Co-occurrence Matrix (GLCM) Based Second Order Statistics for Image Texture Analysis** arXiv, 6 mar. 2024. Disponível em: <<http://arxiv.org/abs/2403.04038>>.

ZVOLEFF, A. **glcm**: Calculate Textures from Grey-Level Co-Occurrence Matrices (GLCMs). R package version 1.6.5, 2020. <<https://CRAN.R-project.org/package=glcm>>.

CONCLUSÕES GERAIS

A modelagem da ocorrência de lianas em fragmentos de Mata Atlântica utilizando técnicas de aprendizagem de máquina demonstrou a eficácia dessas técnicas na classificação de árvores com infestação de lianas. As técnicas de redes neurais artificiais (RNA), máquina de vetor de suporte (SVM) e *random forest* (RF) são as mais indicadas para classificar a infestação por lianas. A inclusão de variáveis climáticas, edáficas e antrópicas nos modelos aumentam a acurácia das previsões, evidenciando a complexidade das interações ecológicas envolvidas na infestação de lianas.

A modelagem da abundância de lianas também utilizando técnicas de aprendizagem de máquina reforça a relevância dessas metodologias na modelagem de florestas tropicais. As abordagens com RNA e regressão vetorial de suporte (RVS) permitem prever com acurácia a abundância de lianas, considerando variáveis ambientais e de uso da terra. Os fatores ambientais influenciam de forma significativa na abundância de lianas, com correlações observadas entre a precipitação, matéria orgânica do solo e a abundância de lianas.

A predição da biomassa do fuste de árvores em fragmentos de Mata Atlântica, integrando descritores texturais GLCM e técnicas de aprendizagem de máquina, mostrou-se uma abordagem inovadora e eficaz. A utilização de variáveis texturais e a inclusão de lianas nos modelos aumentam a precisão das estimativas de biomassa, principalmente em condições sazonais distintas. A RVS destacou-se como a técnica mais robusta para predições em períodos secos e chuvosos. Esses resultados reforçam a importância de adaptar os modelos às condições específicas de cada ecossistema, fornecendo estimativas mais confiáveis para a conservação e manejo sustentável da Mata Atlântica.

APÊNDICE A

Tabela A1 - Estatística descritiva das variáveis de entrada avaliadas referentes ao ano de 2002

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Climáticas						
Temp	Temperatura média anual 1 ano antes da medição	°C	18,11	21,03	23,8	±2,36
Temp_avrg	Média da temperatura média anual de 1989 até 1 ano antes da medição	°C	21,43	23,55	25,68	±1,8
Precp	Precipitação anual no ano de medição	mm	1101,27	1290,39	1491,51	±157,5
Precp_1	Precipitação anual 1 ano antes da medição	mm	864,33	1057,81	1292,17	±167,62
Precp_2	Precipitação média anual 1 e 2 anos antes da medição	mm	1009,33	1129,97	1424,08	±163,69
Precp_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	906,55	1062,25	1377,72	±181,7
Precp_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	893,16	1060,82	1296,29	±156,88
Precp_avrg	Precipitação média anual de 1989 até 1 ano antes da medição	mm	1185,53	1208,36	1272,78	±35,36
Precp_dry	Precipitação total dos três meses mais secos	mm	40,64	48,29	55,62	±6,68
CWD	Déficit hídrico climático no ano de medição	mm	-1863,6	-984,34	-422,75	±421,88
CWD_1	Déficit hídrico climático 1 ano antes da medição	mm	-2328,9	-1358,6	-537,97	±531,39
CWD_2	Déficit hídrico climático médio 1 e 2 anos antes da medição	mm	-2148,4	-1325,3	-537,76	±452,28
CWD_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	-2283,9	-1462,3	-654,64	±452,67
CWD_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	-2555	-1614,2	-695,71	±508,5
CWD_avrg	Déficit hídrico climático de 1989 até 1 ano antes da medição	mm	-2839,8	-1674,9	-640,56	±606,17
Solo						
pH_H2O_0_20	pH H ₂ O de 0 - 20 cm	-	3,62	4,2	5,67	±0,42
P_0_20	Fósforo (P) de 0 - 20 cm	cmolc/dm ³	0	0,01	0,04	±0,01
K_0_20	Potássio (K) de 0 - 20 cm	cmolc/dm ³	0,05	0,11	0,46	±0,08
Ca2_0_20	Cálcio (Ca ²⁺) de 0 cm - 20 cm	cmolc/dm ³	0,11	0,4	1,87	±0,4
Mg2_0_20	Magnésio (Mg ²⁺) de 0 - 20 cm	cmolc/dm ³	0,07	0,19	1,25	±0,18
Al3_0_20	Alumínio trocável (Al ³⁺) de 0 - 20 cm	cmolc/dm ³	0,27	1,2	2,16	±0,46
HAl3_20_40	Acidez potencial (H + Al) de 20 - 40 cm	cmolc/dm ³	4,7	8,15	16,9	±2,15
SB_0_20	Soma de bases (SB) de 0 - 20 cm	cmolc/dm ³	0,25	0,7	3,31	±0,62
t_0_20	CTC efetiva (t) de 0 - 20 cm	cmolc/dm ³	0,71	1,9	3,89	±0,57
t_20_40	CTC efetiva (t) de 20 - 40 cm	cmolc/dm ³	5,2	8,85	17,82	±2,25
V_0_20	Saturação por bases (V) de 0 - 20 cm	%	2,7	7,94	28	±6,17

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
m_0_20	Saturação por alumínio (m) de 0 - 20 cm	%	10,1	64,96	88,3	±18,91
MO_0_20	Matéria orgânica de 0 - 20 cm	dag/kg	2,61	4,72	12,1	±1,66
P_Rem_0_20	Fósforo remanescente (P-Rem) de 0 - 20 cm	mg/L	7,3	17,44	28,9	±6,14
pH_H2O_20_40	pH H2O de 0 - 20 cm	-	3,34	4,33	5,58	±0,42
P_20_40	Fósforo (P) de 20 - 40 cm	cmolc/dm ³	0	0,01	0,02	±0,01
K_20_40	Potássio (K) de 20 - 40 cm	cmolc/dm ³	0,04	0,09	0,49	±0,08
Ca2_20_40	Cálcio (Ca ²⁺) de 20 cm - 40 cm	cmolc/dm ³	0,06	0,33	1,75	±0,37
Mg2_20_40	Magnésio (Mg ²⁺) de 20 - 40 cm	cmolc/dm ³	0,06	0,16	1,27	±0,22
Al3_20_40	Alumínio trocável (Al ³⁺) de 20 - 40 cm	cmolc/dm ³	0,18	1,01	2,16	±0,47
HAl3_0_20	Acidez potencial (H + Al) de 0 - 20 cm	cmolc/dm ³	3,9	6,87	16,4	±2,25
SB_20_40	Soma de bases (SB) de 20 - 40 cm	cmolc/dm ³	0,21	0,58	3,41	±0,64
T_0_20	CTC potencial (T) de 0 - 20 cm	cmolc/dm ³	0,68	1,59	3,99	±0,66
T_20_40	CTC potencial (T) de 20 - 40 cm	cmolc/dm ³	4,21	7,45	17,12	±2,45
V_20_40	Saturação por bases (V) de 20 - 40 cm	%	3	7,52	28,9	±6,16
m_20_40	Saturação por alumínio (m) de 20 - 40 cm	%	14,2	65,64	89,6	±19,69
MO_20_40	Matéria orgânica de 20 - 40 cm	dag/kg	1,83	3,75	9,44	±1,47
P-Rem_20_40	Fósforo remanescente de 20 - 40 cm	mg/L	7,3	16,7	32,1	±6,34
Paisagem						
Elevation	Elevação	m	242	660,23	1169	±329,25
Declivity_perc	Declividade	%	2,97	30,24	65,81	±12,27
Slope_Angle	Ângulo de Inclinação	°	8,58	189,95	338,9	±80,11
Edge_distance	Distância da Borda	m	10,74	100,79	262,11	±69,69
Fragmento	Nome do fragmento (Cachoeira das Pombas, Ipaba e São José)	-	-	-	-	-
Atributos da floresta						
Species	Número de espécies	-	9	30,87	64	±11,68
Stems	Número de fustes por hectare	-	900	1770,57	2940	±516,64
BA_ha	Área basal por hectare	m ² ha ⁻¹	6,28	17,46	41,99	±7,39
DAP	Diâmetro do fuste à 1,30 m de altura	cm	4,77	9,43	73,21	±5,83
Ht	Altura total da árvore	m	1,7	8,66	35	±3,32

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Lianas_2002	Classe de infestação de lianas descrita em campo no ano de 2002 (1, 2, 3 e 4)	-	-	-	-	-
GE	Grupo ecológico (Pioneira, Secundaria inicial, Secundária tardia e Clímax)	-	-	-	-	-
Antropogênicas						
LUH	Histórico de uso da terra (Produção agrícola; desmatamento; eucalipto; extração seletiva)	-	-	-	-	-
Changes	Mudanças no uso da terra: Solo descoberto-solo descoberto Produção agrícola-Eucalipto Eucalipto-Eucalipto Solo descoberto-Floresta*** Produção agrícola-Floresta*** Floresta***-Floresta*** Floresta***-Floresta Floresta-Floresta	-	-	-	-	-

Em que: Min = Valor mínimo da variável; Max = Valor máximo da variável; Med = Valor médio da variável; Dp = Desvio padrão.

Tabela A2 - Estatística descritiva das variáveis explicativas referentes ao ano de 2012.

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Climáticas						
Temp	Temperatura média anual 1 ano antes da medição	°C	17,38	20,28	22,43	±1,99
Temp_avrg	Média da temperatura média anual de 1989 até 1 ano antes da medição	°C	19,79	22,27	24,54	±1,98
Precp	Precipitação anual no ano de medição	mm	782,06	944,84	1135,34	±144,21
Precp_1	Precipitação anual 1 ano antes da medição	mm	1120,10	1247,85	1467,10	±130,57
Precp_2	Precipitação média anual 1 e 2 anos antes da medição	mm	975,00	1167,65	1299,40	±150,48
Precp_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	1035,63	1242,43	1386,60	±160,65
Precp_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	1111,12	1261,21	1355,63	±116,25

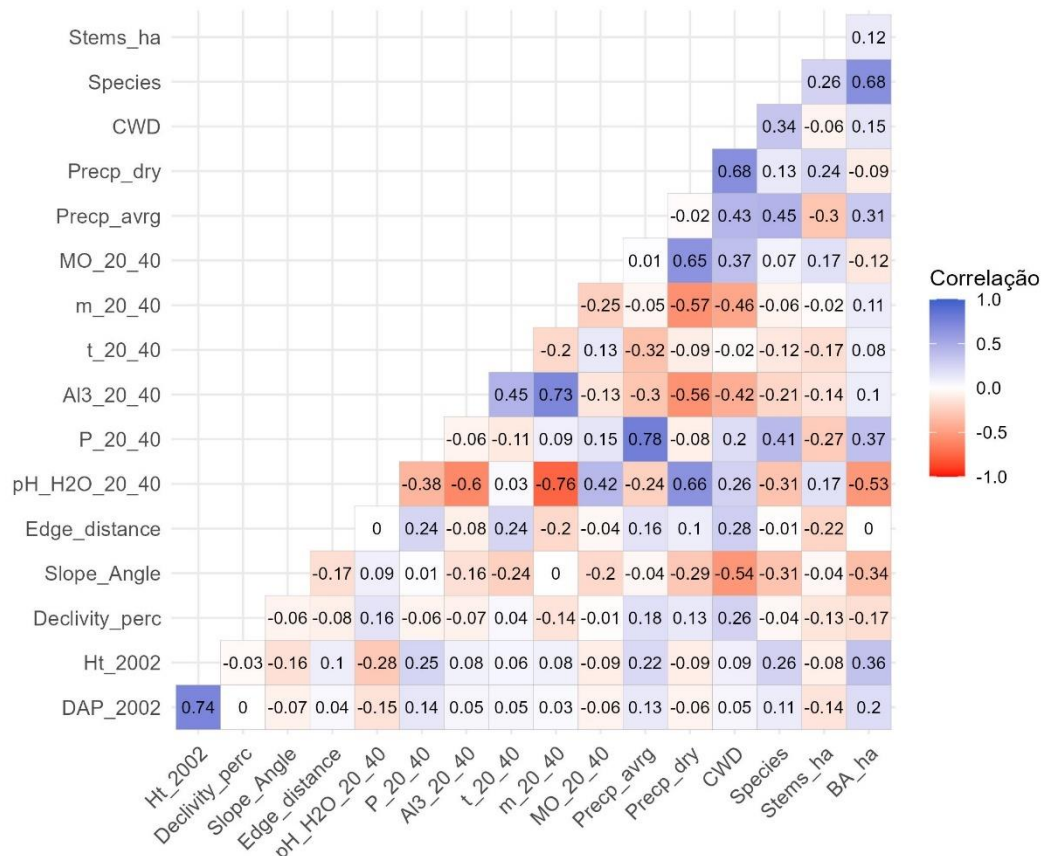
Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Precp_avrg	Precipitação média anual de 1989 até 1 ano antes da medição	mm	1157,89	1225,97	1290,82	±54,71
Precp_dry	Precipitação total dos três meses mais secos	mm	4,10	9,91	15,30	±4,98
CWD	Déficit hídrico climático no ano de medição	mm	-	-	-427,35	±577,99
CWD_1	Déficit hídrico climático 1 ano antes da medição	mm	-	-	-437,82	±408,78
CWD_2	Déficit hídrico climático médio 1 e 2 anos antes da medição	mm	-	-	-478,07	±395,72
CWD_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	1794,18	1003,65	-388,08	±421,28
CWD_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	-	-	-368,42	±422,92
CWD_avrg	Déficit hídrico climático de 1989 até 1 ano antes da medição	mm	-	-	-532,45	±542,58
Solo						
pH_H2O_0_20	pH H2O de 0 - 20 cm	-	3,62	4,20	5,67	±0,42
P_0_20	Fósforo (P) de 0 - 20 cm	cmolc/dm ³	0	0,01	0,04	±0,01
K_0_20	Potássio (K) de 0 - 20 cm	cmolc/dm ³	0,05	0,11	0,46	±0,08
Ca2_0_20	Cálcio (Ca ²⁺) de 0 cm - 20 cm	cmolc/dm ³	0,11	0,40	1,87	±0,4
Mg2_0_20	Magnésio (Mg ²⁺) de 0 - 20 cm	cmolc/dm ³	0,07	0,19	1,25	±0,18
Al3_0_20	Alumínio trocável (Al ³⁺) de 0 - 20 cm	cmolc/dm ³	0,27	1,21	2,16	±0,47
HAl3_20_40	Acidez potencial (H + Al) de 20 - 40 cm	cmolc/dm ³	4,70	8,16	16,90	±2,14
SB_0_20	Soma de bases (SB) de 0 - 20 cm	cmolc/dm ³	0,25	0,70	3,31	±0,62
t_0_20	CTC efetiva (t) de 0 - 20 cm	cmolc/dm ³	0,71	1,91	3,89	±0,57
t_20_40	CTC efetiva (t) de 20 - 40 cm	cmolc/dm ³	5,20	8,86	17,82	±2,23
V_0_20	Saturação por bases (V) de 0 - 20 cm	%	2,70	7,88	28,00	±6,13
m_0_20	Saturação por alumínio (m) de 0 - 20 cm	%	10,10	65,27	88,30	±18,87
MO_0_20	Matéria orgânica de 0 - 20 cm	dag/kg	2,61	4,72	12,10	±1,64
P_Rem_0_20	Fósforo remanescente (P-Rem) de 0 - 20 cm	mg/L	7,30	17,49	28,90	±6,1
pH_H2O_20_40	pH H2O de 0 - 20 cm	-	3,34	4,32	5,58	±0,42
P_20_40	Fósforo (P) de 20 - 40 cm	cmolc/dm ³	0,00	0,01	0,02	±0,01
K_20_40	Potássio (K) de 20 - 40 cm	cmolc/dm ³	0,04	0,09	0,49	±0,08

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Ca2_20_40	Cálcio (Ca ²⁺) de 20 cm - 40 cm	cmolc/dm ³	0,06	0,33	1,75	±0,36
Mg2_20_40	Magnésio (Mg ²⁺) de 20 - 40 cm	cmolc/dm ³	0,06	0,16	1,27	±0,22
Al3_20_40	Alumínio trocável (Al ³⁺) de 20 - 40 cm	cmolc/dm ³	0,18	1,02	2,16	±0,47
HAl3_0_20	Acidez potencial (H + Al) de 0 - 20 cm	cmolc/dm ³	3,90	6,87	16,40	±2,23
SB_20_40	Soma de bases (SB) de 20 - 40 cm	cmolc/dm ³	0,21	0,58	3,41	±0,63
T_0_20	CTC potencial (T) de 0 - 20 cm	cmolc/dm ³	0,68	1,60	3,99	±0,66
T_20_40	CTC potencial (T) de 20 - 40 cm	cmolc/dm ³	4,21	7,45	17,12	±2,43
V_20_40	Saturação por bases (V) de 20 - 40 cm	%	3,00	7,45	28,90	±6,12
m_20_40	Saturação por alumínio (m) de 20 - 40 cm	%	14,20	66,00	89,60	±19,68
MO_20_40	Matéria orgânica de 20 - 40 cm	dag/kg	1,83	3,75	9,44	±1,45
P-Rem_20_40	Fósforo remanescente de 20 - 40 cm	mg/L	7,30	16,74	32,10	±6,28
Paisagem						
Elevation	Elevação	m	242	652,72	1169	±330,76
Declivity_perc	Declividade	%	2,97	30,39	65,81	±12,21
Slope_Angle	Ângulo de Inclinação	°	8,58	190,55	338,9	±79,47
Edge_distance	Distância da Borda	m	6,98	99,05	262,11	±70,2
Fragmento	Nome do fragmento (Cachoeira das Pombas, Ipaba e São José)	-	-	-	-	-
Atributos da floresta						
Species	Número de espécies	-	15	32,8	64	±10,98
Stems	Número de fustes por hectare	-	860	1888,52	3440	±546,13
BA_ha	Área basal por hectare	m ² ha ⁻¹	6,71	21,04	44,3	±7,37
DAP	Diâmetro do fuste à 1,30 m de altura	cm	1,75	10,29	73,12	±6,14
Ht	Altura total da árvore	m	1,7	8,98	35	±3,37
Lianas_2012	Classe de infestação de lianas descrita em campo no ano de 2012 (1, 2, 3 e 4)	-	-	-	-	-
GE	Grupo ecológico (Pioneira, Secundária inicial, Secundária tardia e Climax)	-	-	-	-	-
Antropogênicas						
LUH	Histórico de uso da terra (produção agrícola; desmatamento; eucalipto; extração seletiva)	-	-	-	-	-

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Changes	Mudanças no uso da terra:					
	Solo descoberto-solo descoberto					
	Produção agrícola-Eucalipto					
	Eucalipto-Eucalipto					
	Solo descoberto-Floresta***	-	-	-	-	-
	Produção agrícola-Floresta***					
	Floresta***-Floresta***					
	Floresta***-Floresta					
	Floresta-Floresta					

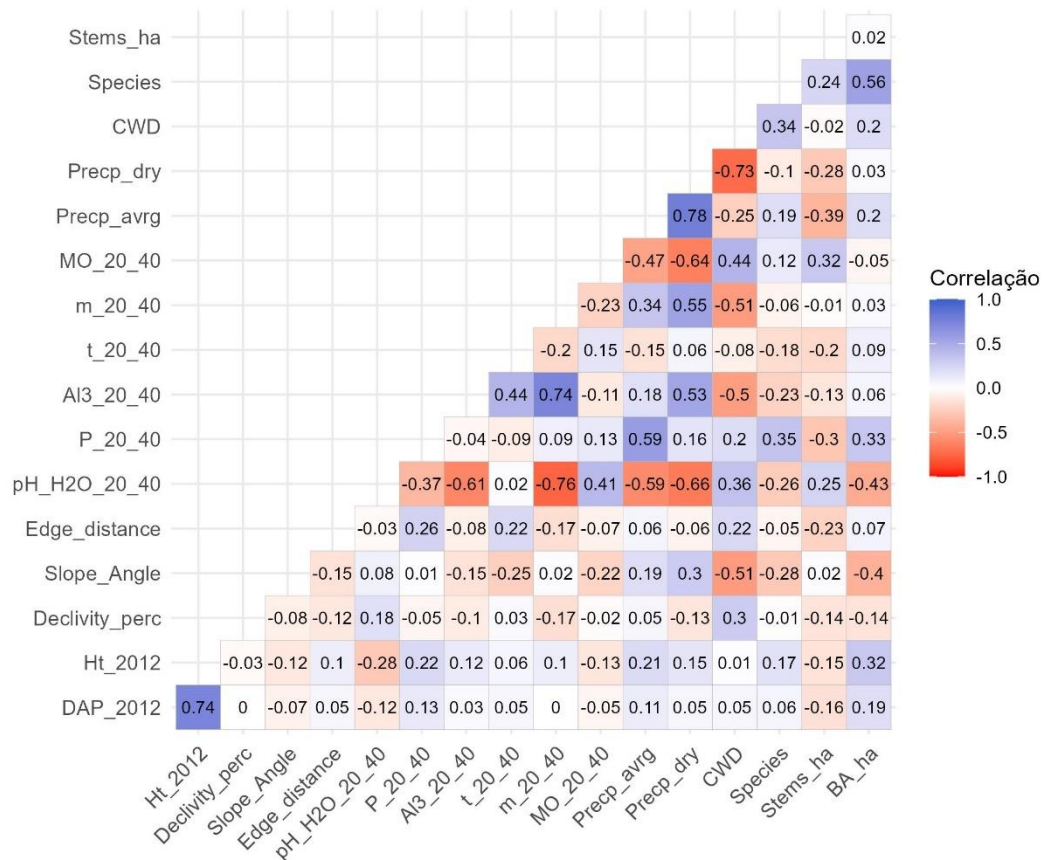
Em que: Min = Valor mínimo da variável; Max = Valor máximo da variável; Méd = Valor médio da variável; Dp = Desvio padrão.

Figura A1 - Correlação linear de Pearson entre as variáveis quantitativas pré-selecionadas para projeção da ocorrência de árvores infestadas por lianas nos períodos 2002-2012 e 2002-2017.



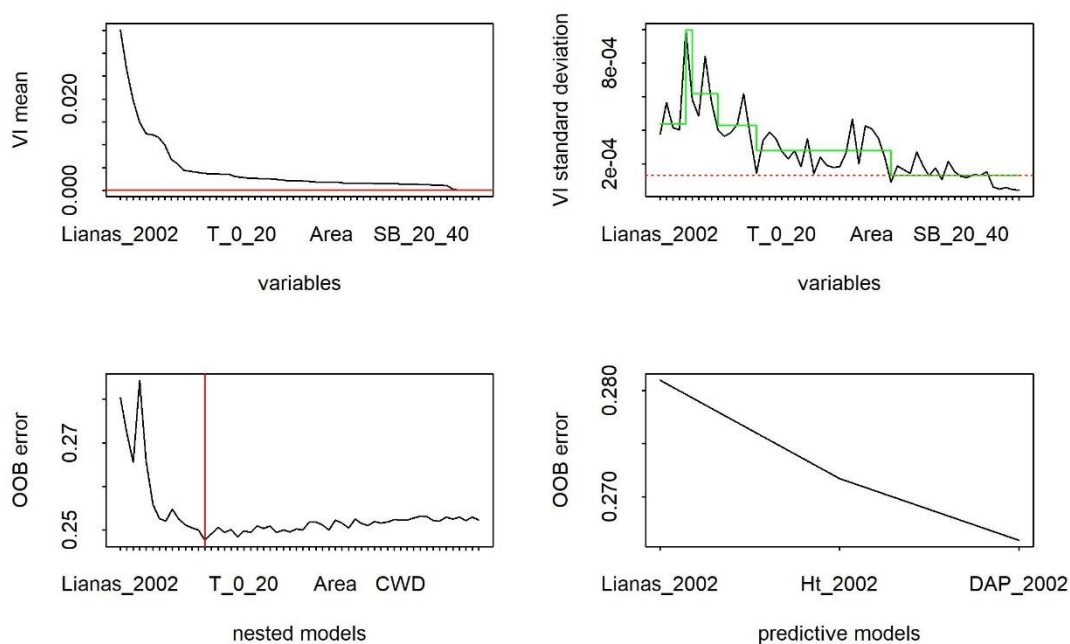
Fonte: O autor (2024).

Figura A2 - Correlação linear de Pearson entre as variáveis quantitativas pré-selecionadas para projeção da ocorrência de árvores infestadas por lianas no período 2012-2017.



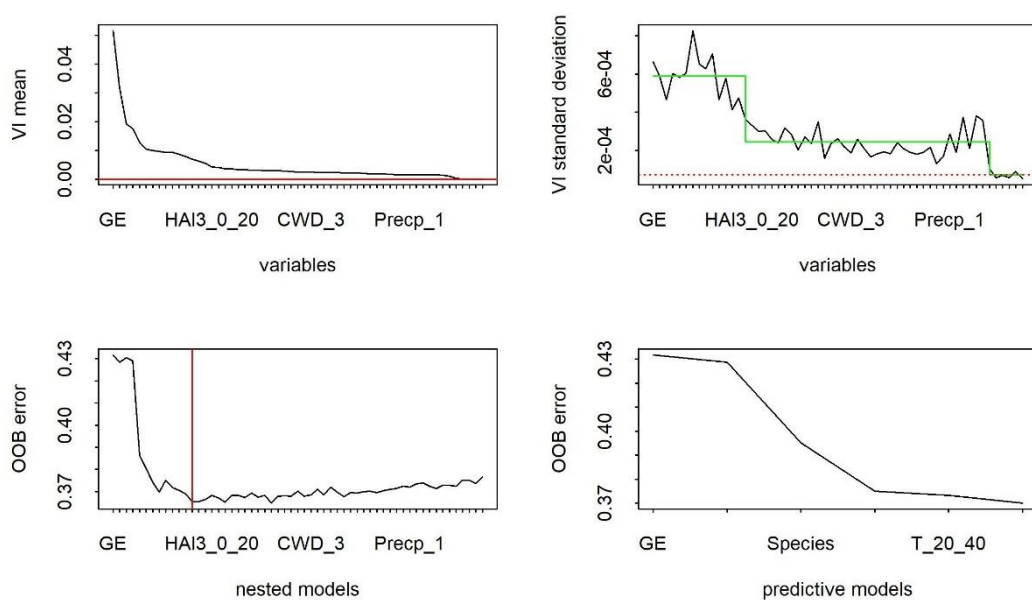
Fonte: O autor (2024).

Figura A3 - Seleção das variáveis mais importantes através do *Random Forest* para classificação da infestação por lianas no período 2002-2012.



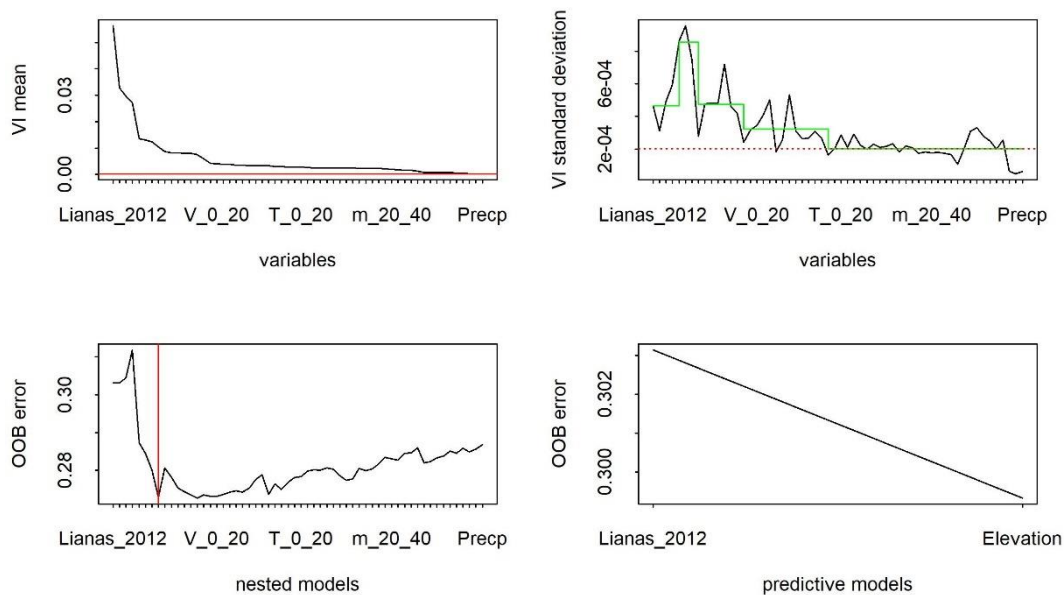
Fonte: O autor (2024).

Figura A4 - Seleção das variáveis mais importantes através do *Random Forest* para classificação da infestação por lianas no período 2002-2017.



Fonte: O autor (2024).

Figura A5 - Seleção das variáveis mais importantes através do *Random Forest* para classificação da infestação por lianas no período 2012-2017.



Fonte: O autor (2024).

APÊNDICE B

Tabela B1 - Estatística descritiva das variáveis explicativas referentes ao ano de 2002.

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Climáticas						
Temp	Temperatura média anual 1 ano antes da medição	°C	18,11	21,03	23,8	±2,36
Temp_avrg	Média da temperatura média anual de 1989 até 1 ano antes da medição	°C	21,43	23,55	25,68	±1,8
Precp	Precipitação anual no ano de medição	mm	1101,27	1290,39	1491,51	±157,5
Precp_1	Precipitação anual 1 ano antes da medição	mm	864,33	1057,81	1292,17	±167,62
Precp_2	Precipitação média anual 1 e 2 anos antes da medição	mm	1009,33	1129,97	1424,08	±163,69
Precp_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	906,55	1062,25	1377,72	±181,7
Precp_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	893,16	1060,82	1296,29	±156,88
Precp_avrg	Precipitação média anual de 1989 até 1 ano antes da medição	mm	1185,53	1208,36	1272,78	±35,36
Precp_dry	Precipitação total dos três meses mais secos	mm	40,64	48,29	55,62	±6,68
CWD	Déficit hídrico climático no ano de medição	mm	-1863,6	-984,34	-422,75	±421,88
CWD_1	Déficit hídrico climático 1 ano antes da medição	mm	-2328,9	-1358,6	-537,97	±531,39
CWD_2	Déficit hídrico climático médio 1 e 2 anos antes da medição	mm	-2148,4	-1325,3	-537,76	±452,28
CWD_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	-2283,9	-1462,3	-654,64	±452,67
CWD_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	-2555	-1614,2	-695,71	±508,5
CWD_avrg	Déficit hídrico climático de 1989 até 1 ano antes da medição	mm	-2839,8	-1674,9	-640,56	±606,17
Solo						
pH_H2O_0_20	pH H ₂ O de 0 - 20 cm	-	3,62	4,2	5,67	±0,42
P_0_20	Fósforo (P) de 0 - 20 cm	cmolc/dm ³	0	0,01	0,04	±0,01
K_0_20	Potássio (K) de 0 - 20 cm	cmolc/dm ³	0,05	0,11	0,46	±0,08
Ca2_0_20	Cálcio (Ca ²⁺) de 0 cm - 20 cm	cmolc/dm ³	0,11	0,4	1,87	±0,4
Mg2_0_20	Magnésio (Mg ²⁺) de 0 - 20 cm	cmolc/dm ³	0,07	0,19	1,25	±0,18
Al3_0_20	Alumínio trocável (Al ³⁺) de 0 - 20 cm	cmolc/dm ³	0,27	1,2	2,16	±0,46
HAl3_20_40	Acidez potencial (H + Al) de 20 - 40 cm	cmolc/dm ³	4,7	8,15	16,9	±2,15
SB_0_20	Soma de bases (SB) de 0 - 20 cm	cmolc/dm ³	0,25	0,7	3,31	±0,62
t_0_20	CTC efetiva (t) de 0 - 20 cm	cmolc/dm ³	0,71	1,9	3,89	±0,57
t_20_40	CTC efetiva (t) de 20 - 40 cm	cmolc/dm ³	5,2	8,85	17,82	±2,25

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
V_0_20	Saturação por bases (V) de 0 - 20 cm	%	2,7	7,94	28	±6,17
m_0_20	Saturação por alumínio (m) de 0 - 20 cm	%	10,1	64,96	88,3	±18,91
MO_0_20	Matéria orgânica de 0 - 20 cm	dag/kg	2,61	4,72	12,1	±1,66
P_Rem_0_20	Fósforo remanescente (P-Rem) de 0 - 20 cm	mg/L	7,3	17,44	28,9	±6,14
pH_H2O_20_40	pH H2O de 0 - 20 cm	-	3,34	4,33	5,58	±0,42
P_20_40	Fósforo (P) de 20 - 40 cm	cmolc/dm ³	0	0,01	0,02	±0,01
K_20_40	Potássio (K) de 20 - 40 cm	cmolc/dm ³	0,04	0,09	0,49	±0,08
Ca2_20_40	Cálcio (Ca ²⁺) de 20 cm - 40 cm	cmolc/dm ³	0,06	0,33	1,75	±0,37
Mg2_20_40	Magnésio (Mg ²⁺) de 20 - 40 cm	cmolc/dm ³	0,06	0,16	1,27	±0,22
Al3_20_40	Alumínio trocável (Al ³⁺) de 20 - 40 cm	cmolc/dm ³	0,18	1,01	2,16	±0,47
HAl3_0_20	Acidez potencial (H + Al) de 0 - 20 cm	cmolc/dm ³	3,9	6,87	16,4	±2,25
SB_20_40	Soma de bases (SB) de 20 - 40 cm	cmolc/dm ³	0,21	0,58	3,41	±0,64
T_0_20	CTC potencial (T) de 0 - 20 cm	cmolc/dm ³	0,68	1,59	3,99	±0,66
T_20_40	CTC potencial (T) de 20 - 40 cm	cmolc/dm ³	4,21	7,45	17,12	±2,45
V_20_40	Saturação por bases (V) de 20 - 40 cm	%	3	7,52	28,9	±6,16
m_20_40	Saturação por alumínio (m) de 20 - 40 cm	%	14,2	65,64	89,6	±19,69
MO_20_40	Matéria orgânica de 20 - 40 cm	dag/kg	1,83	3,75	9,44	±1,47
P-Rem_20_40	Fósforo remanescente de 20 - 40 cm	mg/L	7,3	16,7	32,1	±6,34
Paisagem						
Elevation	Elevação	m	242	660,23	1169	±329,25
Declivity_perc	Declividade	%	2,97	30,24	65,81	±12,27
Slope_Angle	Ângulo de Inclinação	°	8,58	189,95	338,9	±80,11
Edge_distance	Distância da Borda	m	10,74	100,79	262,11	±69,69
Fragmento	Nome do fragmento (Cachoeira das Pombas, Ipaba e São José)	-	-	-	-	-
Atributos da floresta						
Species	Número de espécies	-	9	30,87	64	±11,68
Stems	Número de fustes por hectare	-	900	1770,57	2940	±516,64
BA_ha	Área basal por hectare	m ² ha ⁻¹	6,28	17,46	41,99	±7,39
Ind_ha_2002	Número de indivíduos infestados por lianas por hectare no ano de 2002	Ind. ha ⁻¹	60	712,2	1540	±311,4

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
GE	Grupo ecológico (Pioneira, Secundária inicial, Secundária tardia e Clímax)	-	-	-	-	-
Antropogênicas						
LUH	Histórico de uso da terra (Produção agrícola; desmatamento; eucalipto; extração seletiva)	-	-	-	-	-
Changes	Mudanças no uso da terra: Solo descoberto-solo descoberto Produção agrícola-Eucalipto Eucalipto-Eucalipto Solo descoberto-Floresta*** Produção agrícola-Floresta*** Floresta***-Floresta*** Floresta***-Floresta Floresta-Floresta	-	-	-	-	-

Em que: Min = Valor mínimo da variável; Max = Valor máximo da variável; Méd = Valor médio da variável; Dp = Desvio padrão.

Tabela B2 - Estatística descritiva das variáveis explicativas referentes ao ano de 2012.

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Climáticas						
Temp	Temperatura média anual 1 ano antes da medição	°C	17,38	20,28	22,43	±1,99
Temp_avrg	Média da temperatura média anual de 1989 até 1 ano antes da medição	°C	19,79	22,27	24,54	±1,98
Precp	Precipitação anual no ano de medição	mm	782,06	944,84	1135,34	±144,21
Precp_1	Precipitação anual 1 ano antes da medição	mm	1120,10	1247,85	1467,10	±130,57
Precp_2	Precipitação média anual 1 e 2 anos antes da medição	mm	975,00	1167,65	1299,40	±150,48

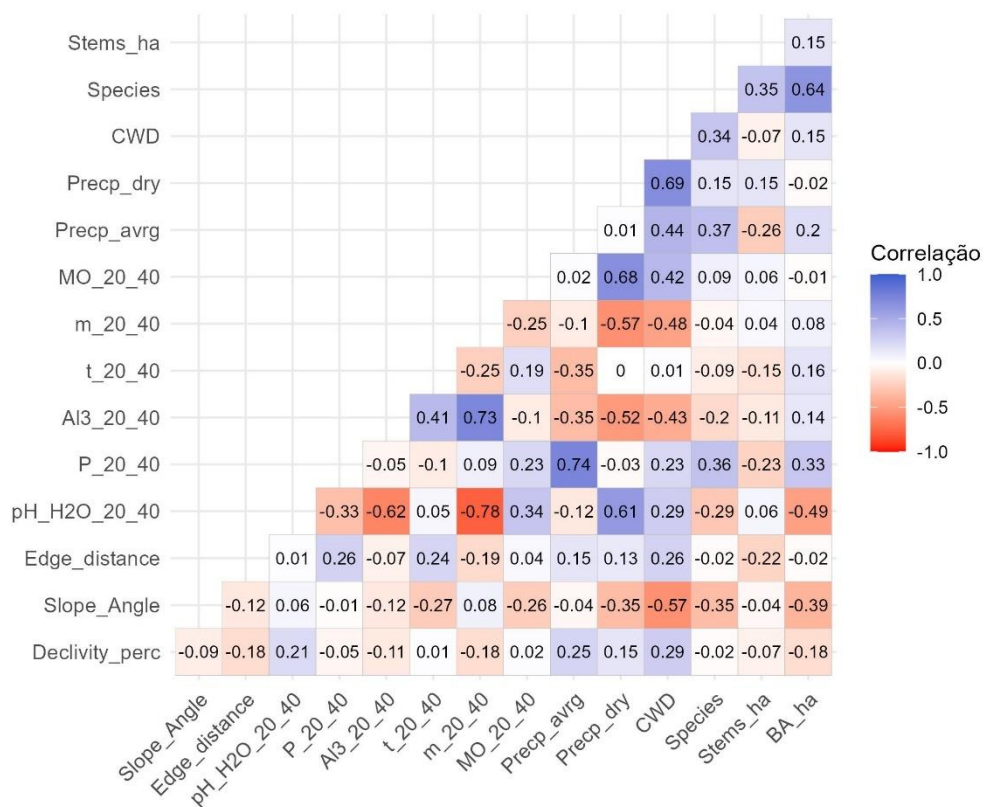
Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Precp_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	1035,63	1242,43	1386,60	±160,65
Precp_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	1111,12	1261,21	1355,63	±116,25
Precp_avrg	Precipitação média anual de 1989 até 1 ano antes da medição	mm	1157,89	1225,97	1290,82	±54,71
Precp_dry	Precipitação total dos três meses mais secos	mm	4,10	9,91	15,30	±4,98
CWD	Déficit hídrico climático no ano de medição	mm	-	-	-427,35	±577,99
CWD_1	Déficit hídrico climático 1 ano antes da medição	mm	-	-	-437,82	±408,78
CWD_2	Déficit hídrico climático médio 1 e 2 anos antes da medição	mm	-	-	-478,07	±395,72
CWD_3	Precipitação média anual 1, 2 e 3 anos antes da medição	mm	-	-	-388,08	±421,28
CWD_4	Precipitação média anual 1, 2, 3 e 4 anos antes da medição	mm	-	-	-368,42	±422,92
CWD_avrg	Déficit hídrico climático de 1989 até 1 ano antes da medição	mm	-	-	-532,45	±542,58
Solo						
pH_H2O_0_20	pH H2O de 0 - 20 cm	-	3,62	4,20	5,67	±0,42
P_0_20	Fósforo (P) de 0 - 20 cm	cmolc/dm ³	0	0,01	0,04	±0,01
K_0_20	Potássio (K) de 0 - 20 cm	cmolc/dm ³	0,05	0,11	0,46	±0,08
Ca2_0_20	Cálcio (Ca ²⁺) de 0 cm - 20 cm	cmolc/dm ³	0,11	0,40	1,87	±0,4
Mg2_0_20	Magnésio (Mg ²⁺) de 0 - 20 cm	cmolc/dm ³	0,07	0,19	1,25	±0,18
Al3_0_20	Alumínio trocável (Al ³⁺) de 0 - 20 cm	cmolc/dm ³	0,27	1,21	2,16	±0,47
HAl3_20_40	Acidez potencial (H + Al) de 20 - 40 cm	cmolc/dm ³	4,70	8,16	16,90	±2,14
SB_0_20	Soma de bases (SB) de 0 - 20 cm	cmolc/dm ³	0,25	0,70	3,31	±0,62
t_0_20	CTC efetiva (t) de 0 - 20 cm	cmolc/dm ³	0,71	1,91	3,89	±0,57
t_20_40	CTC efetiva (t) de 20 - 40 cm	cmolc/dm ³	5,20	8,86	17,82	±2,23
V_0_20	Saturação por bases (V) de 0 - 20 cm	%	2,70	7,88	28,00	±6,13
m_0_20	Saturação por alumínio (m) de 0 - 20 cm	%	10,10	65,27	88,30	±18,87
MO_0_20	Matéria orgânica de 0 - 20 cm	dag/kg	2,61	4,72	12,10	±1,64
P_Rem_0_20	Fósforo remanescente (P-Rem) de 0 - 20 cm	mg/L	7,30	17,49	28,90	±6,1
pH_H2O_20_40	pH H2O de 0 - 20 cm	-	3,34	4,32	5,58	±0,42

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
P_20_40	Fósforo (P) de 20 - 40 cm	cmolc/dm ³	0,00	0,01	0,02	±0,01
K_20_40	Potássio (K) de 20 - 40 cm	cmolc/dm ³	0,04	0,09	0,49	±0,08
Ca2_20_40	Cálcio (Ca ²⁺) de 20 cm - 40 cm	cmolc/dm ³	0,06	0,33	1,75	±0,36
Mg2_20_40	Magnésio (Mg ²⁺) de 20 - 40 cm	cmolc/dm ³	0,06	0,16	1,27	±0,22
Al3_20_40	Alumínio trocável (Al ³⁺) de 20 - 40 cm	cmolc/dm ³	0,18	1,02	2,16	±0,47
HAl3_0_20	Acidez potencial (H + Al) de 0 - 20 cm	cmolc/dm ³	3,90	6,87	16,40	±2,23
SB_20_40	Soma de bases (SB) de 20 - 40 cm	cmolc/dm ³	0,21	0,58	3,41	±0,63
T_0_20	CTC potencial (T) de 0 - 20 cm	cmolc/dm ³	0,68	1,60	3,99	±0,66
T_20_40	CTC potencial (T) de 20 - 40 cm	cmolc/dm ³	4,21	7,45	17,12	±2,43
V_20_40	Saturação por bases (V) de 20 - 40 cm	%	3,00	7,45	28,90	±6,12
m_20_40	Saturação por alumínio (m) de 20 - 40 cm	%	14,20	66,00	89,60	±19,68
MO_20_40	Matéria orgânica de 20 - 40 cm	dag/kg	1,83	3,75	9,44	±1,45
P-Rem_20_40	Fósforo remanescente de 20 - 40 cm	mg/L	7,30	16,74	32,10	±6,28
Paisagem						
Elevation	Elevação	m	242	652,72	1169	±330,76
Declivity_perc	Declividade	%	2,97	30,39	65,81	±12,21
Slope_Angle	Ângulo de Inclinação	°	8,58	190,55	338,9	±79,47
Edge_distance	Distância da Borda	m	6,98	99,05	262,11	±70,2
Fragmento	Nome do fragmento (Cachoeira das Pombas, Ipaba e São José)	-	-	-	-	-
Atributos da floresta						
Species	Número de espécies	-	15	32,8	64	±10,98
Stems	Número de fustes por hectare	-	860	1888,52	3440	±546,13
BA_ha	Área basal por hectare	m ² ha ⁻¹	6,71	21,04	44,3	±7,37
Ind_ha_2012	Número de indivíduos infestados por lianas por hectare no ano de 2012	Ind. ha ⁻¹	20	327,2	900	±215,6
GE	Grupo ecológico (Pioneira, Secundaria inicial, Secundária tardia e Clímax)	-	-	-	-	-
Antropogênicas						
LUH	Histórico de uso da terra (produção agrícola; desmatamento; eucalipto; extração seletiva)	-	-	-	-	-

Variáveis	Descrição	Unidade	Min	Méd.	Max	Dp
Changes	Mudanças no uso da terra:					
	Solo descoberto-solo descoberto					
	Produção agrícola-Eucalipto					
	Eucalipto-Eucalipto					
	Solo descoberto-Floresta***	-	-	-	-	-
	Produção agrícola-Floresta***					
	Floresta***-Floresta***					
	Floresta***-Floresta					
	Floresta-Floresta					

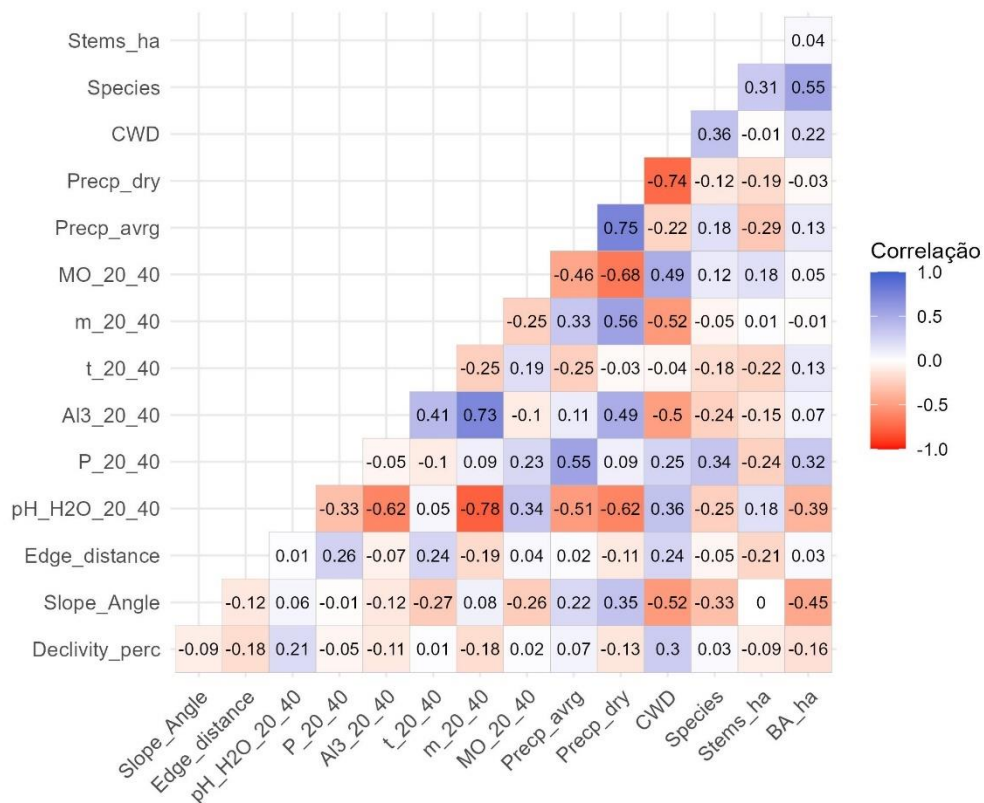
Em que: Min = Valor mínimo da variável; Max = Valor máximo da variável; Méd = Valor médio da variável; Dp = Desvio padrão.

Figura B1 - Correlação linear de Pearson entre as variáveis quantitativas pré-selecionadas para projeção da abundância de árvores infestadas por lianas por hectare nos períodos 2002-2012 e 2002-2017.



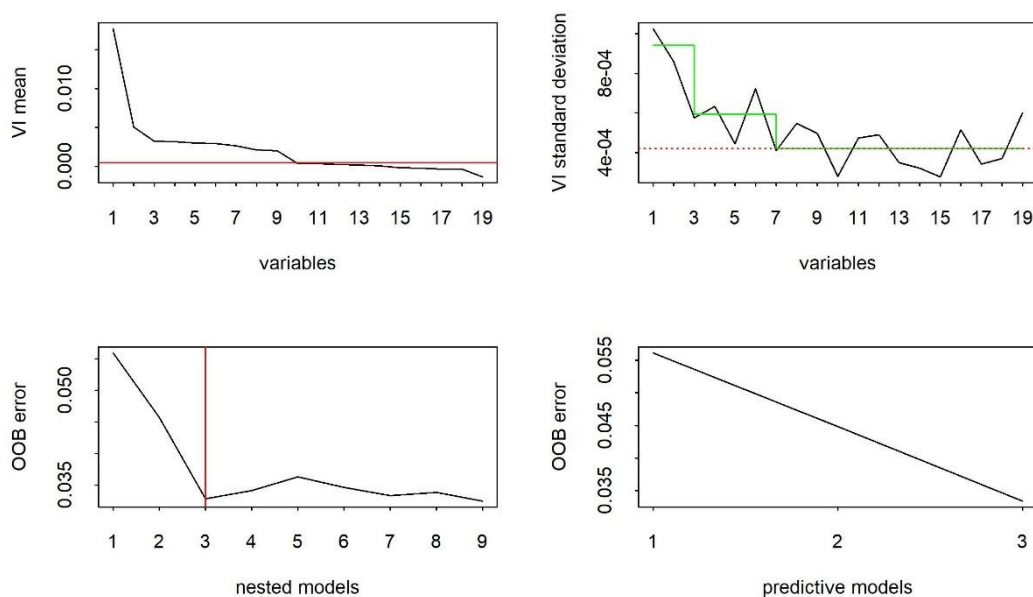
Fonte: O autor (2024).

Figura B2 - Correlação linear de Pearson entre as variáveis quantitativas pré-selecionadas para projeção da abundância de árvores infestadas por lianas por hectare nos períodos 2012-2017.



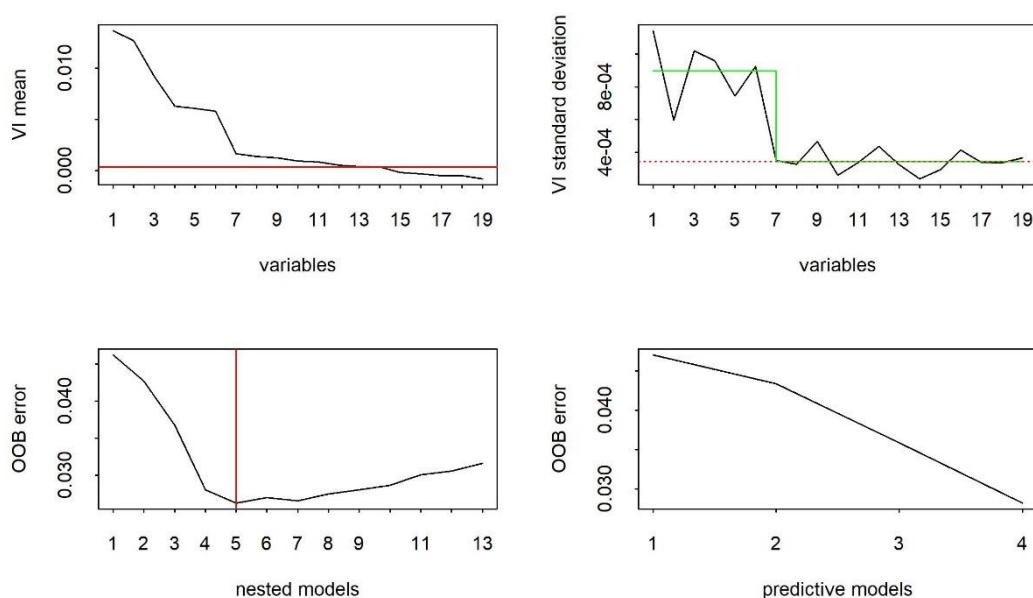
Fonte: O autor (2024).

Figura B3 - Seleção das variáveis mais importantes através do *Random Forest* para projeção no período 2002-2012. Os valores 1, 2 e 3 representam as variáveis Ind_ha_2002, BA_ha e Fragmentos, respectivamente.



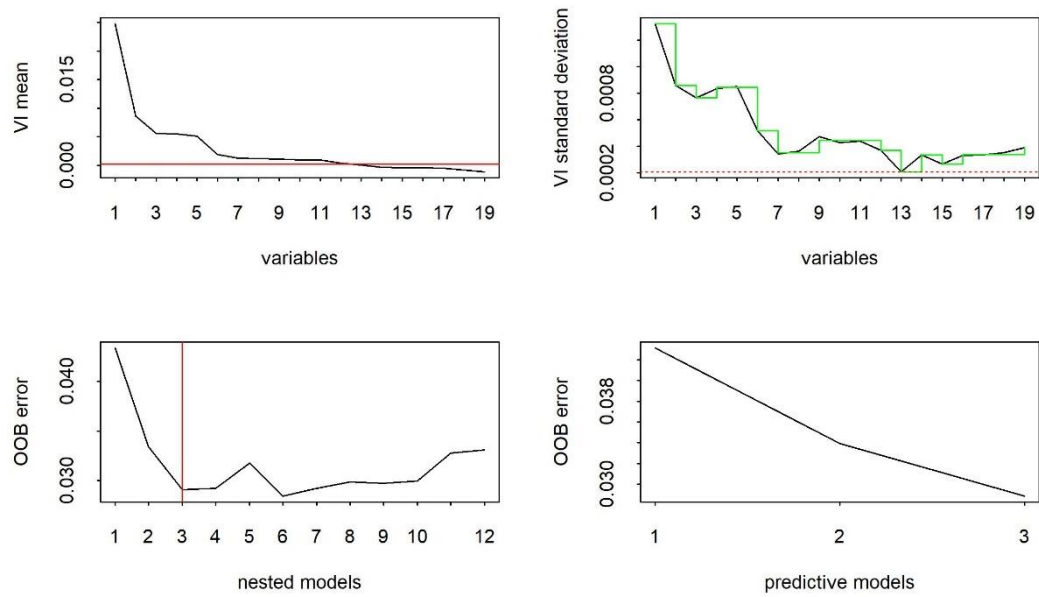
Fonte: O autor (2024).

Figura B4 - Seleção das variáveis mais importantes através do *Random Forest* para projeção no período 2002-2017. Os valores 1, 2, 3 e 4 representam as variáveis MO_20_40, Ind_ha_02, BA_ha e Precp_dry, respectivamente.



Fonte: O autor (2024).

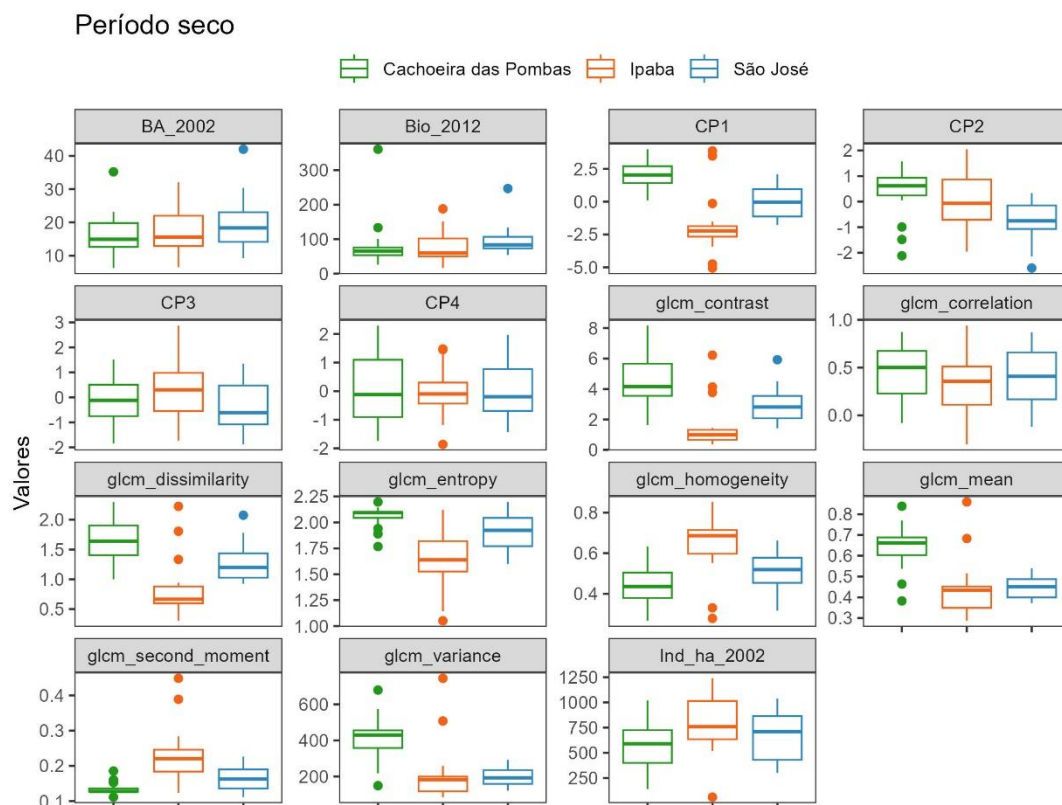
Figura B5 - Seleção das variáveis mais importantes através do *Random Forest* para projeção no período 2012-2017. Os valores 1, 2 e 3 representam as variáveis *Ind_ha_12*, *MO_20_40* e *Precp_dry*, respectivamente.



Fonte: O autor (2024).

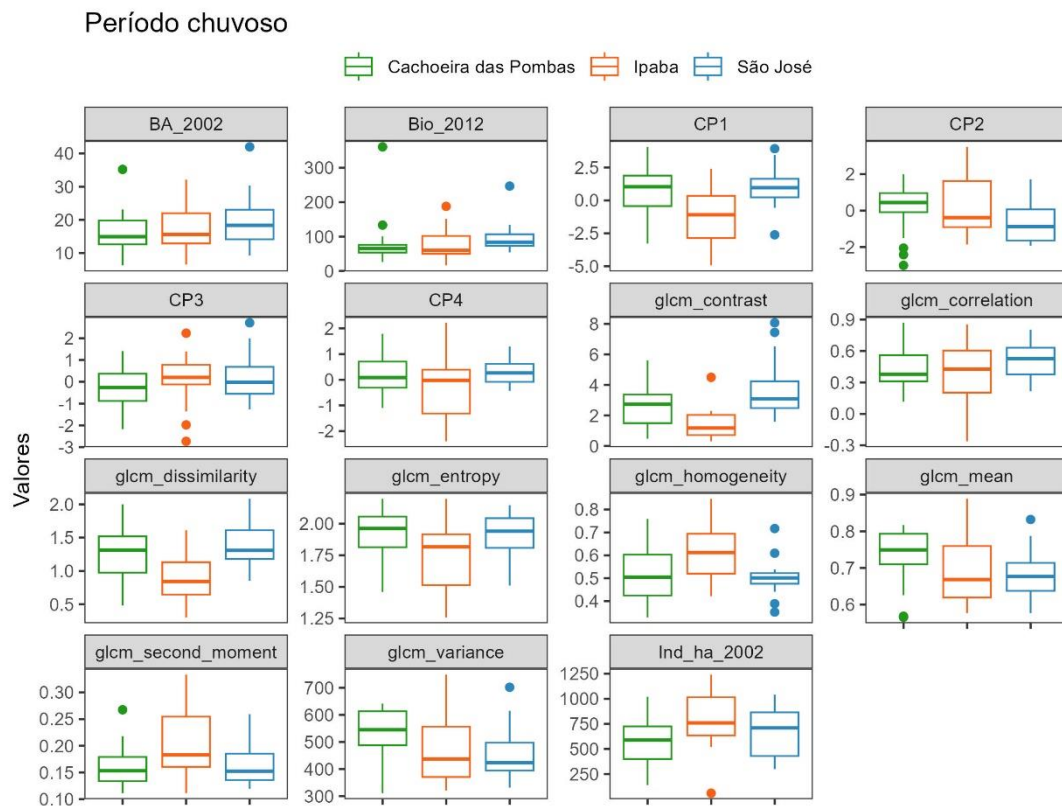
APÊNDICE C

Figura C1 - Boxplot das componentes principais (CP) e descritores texturais extraídos de imagem do satélite Landsat 7 no período seco.



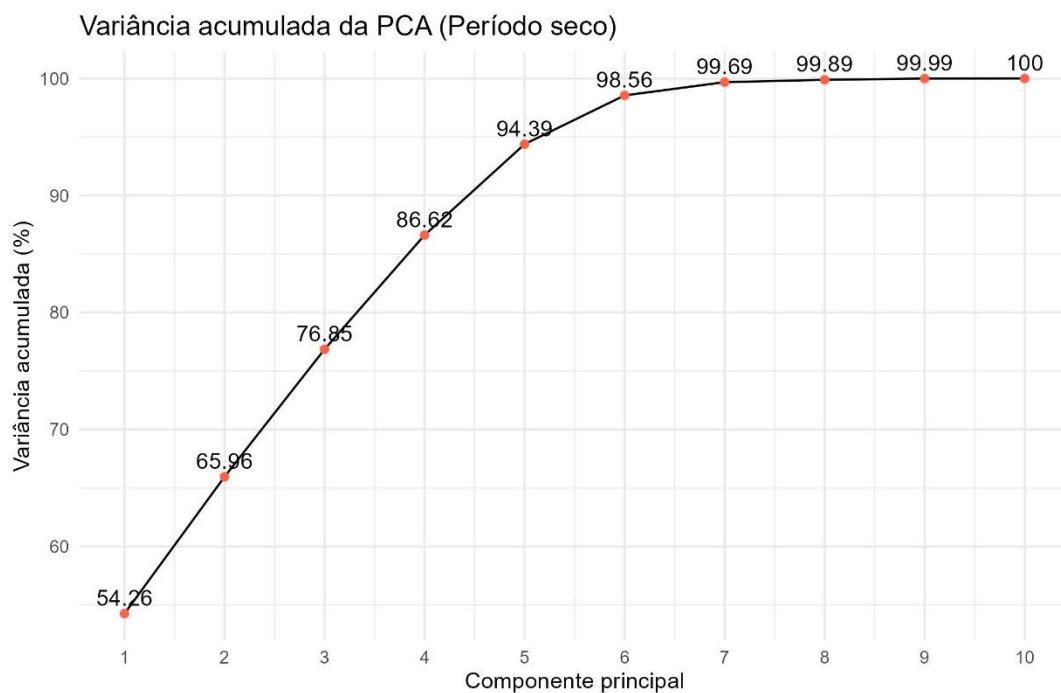
Fonte: O autor (2024).

Figura C2 - Boxblot dos componentes principais (CP) e descritores texturais extraídos de imagem do satélite Landsat 7 no período chuvoso.



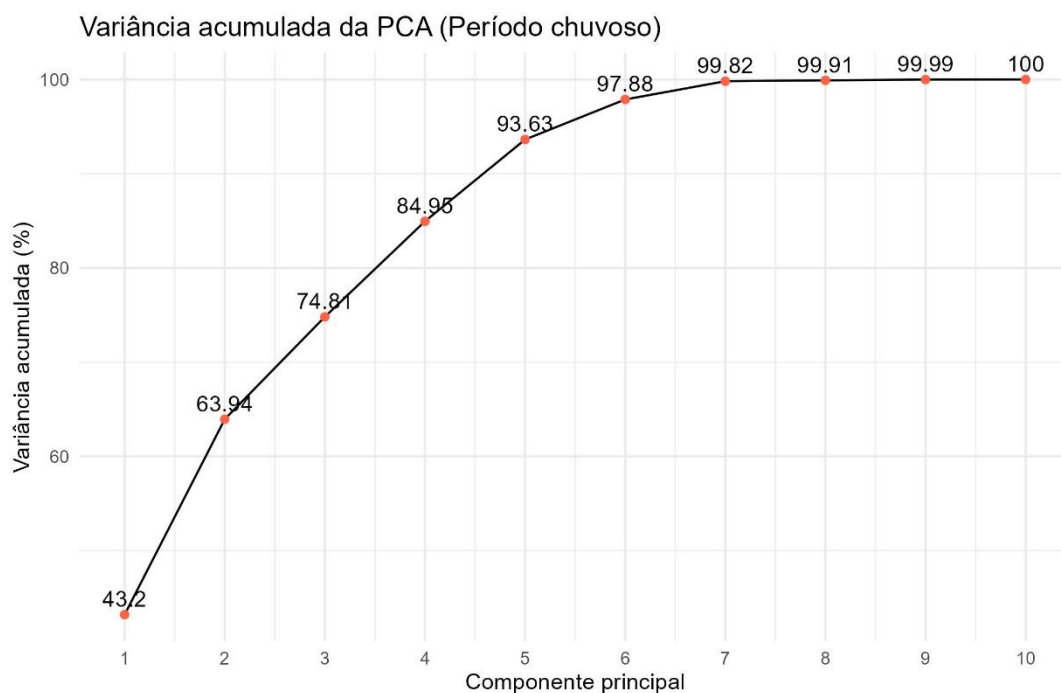
Fonte: O autor (2024).

Figura C3 - Variância acumulada da análise de componentes principais (PCA) no período seco.



Fonte: O autor (2024).

Figura C4 - Variância acumulada da análise de componentes principais (PCA) no período chuvoso.



Fonte: O autor (2024).