

**UNIVERSIDADE FEDERAL DE VIÇOSA**

**Análise de sobrevivência para avaliação dos processos demográficos de  
árvores da Mata Atlântica brasileira**

Marciel Lelis Duarte  
*Doctor Scientiae*

**VIÇOSA - MINAS GERAIS  
2025**

**MARCIEL LELIS DUARTE**

**Análise de sobrevivência para avaliação dos processos demográficos de  
árvores da Mata Atlântica brasileira**

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

Orientador: Sebastiao Martins Filho

Coorientador: Carlos M. M. Eleto Torres

**VIÇOSA - MINAS GERAIS  
2025**

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

T

D812a Duarte, Marciel Lelis, 1989-  
2025 Análise de sobrevivência para avaliação dos processos  
demográficos de árvores da Mata Atlântica brasileira / Marciel  
Lelis Duarte. – Viçosa, MG, 2025.  
1 tese eletrônica (81 f.): il. (algumas color.).

Orientador: Sebastião Martins Filho.  
Tese (doutorado) - Universidade Federal de Viçosa,  
Departamento de Estatística, 2025.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2025.177>

Modo de acesso: World Wide Web.

1. Análise de sobrevivência (Biometria). 2. Ecologia das  
florestas tropicais - Mata Atlântica. 3. Desmatamento - Mata  
Atlântica - Estatísticas. 4. Aprendizado do computador.  
I. Martins Filho, Sebastião, 1961-. II. Universidade Federal de  
Viçosa. Departamento de Estatística. Programa de  
Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.546

**MARCIEL LELIS DUARTE**

**Análise de sobrevivência para avaliação dos processos demográficos de  
árvores da Mata Atlântica brasileira**

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

APROVADA: 12 de março de 2025.

Assentimento:

---

Marciel Lelis Duarte  
Autor

---

Sebastiao Martins Filho  
Orientador

Essa tese foi assinada digitalmente pelo autor em 03/04/2025 às 21:55:39 e pelo orientador em 03/04/2025 às 22:06:33. 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 **CMQJ.JJ9C.N1FH** e clique no botão 'Validar documento'.

## **AGRADECIMENTOS**

Ao Grande Espírito e o Divino Pai Eterno, que me presenteou com a força e a coragem necessária para trilhar mais este percurso.

À minha mãe, Eunice, pelo encorajamento e amor incondicional, sem a qual nada seria possível.

Meu irmão Roniel pela ajuda, apoio, amizade em todos os momentos de minha vida.

Ao meu vô João (in memoria) pelos ensinamentos, amizade que irei levar comigo por toda a vida.

Ao professor Sebastião, expresse minha sincera gratidão pela sua orientação, ensinamentos e amizade, transmitindo-me valiosos ensinamentos ao longo do caminho.

Ao professor Carlos Torres pela coorientação, sua disponibilidade, conselhos e questionamentos.

Aos professores do Programa de Pós-Graduação em Estatística Aplicada e Biometria aplicada pelos ensinamentos ao longo do curso.

À banca examinadora por sua participação na defesa e pelas valiosas contribuições oferecidas para melhoria deste trabalho.

À toda minha família, que mesmo longe, sempre me apoiou com muito carinho.

Aos sempre amigos, que sempre me ajudaram e me escutaram sempre que precisei.

Aos funcionários do Departamento de Estatística, por serem sempre tão prestativos.

A Universidade Federal de Viçosa e ao Departamento de Estatística pela oportunidade de realizar o curso.

A Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela bolsa de estudos.

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

A todos aqueles que direta ou indiretamente, contribuíram para realização deste trabalho.

## RESUMO

DUARTE, Marciel Leis, D.Sc., Universidade Federal de Viçosa, março de 2025. **Análise de sobrevivência para avaliação dos processos demográficos de árvores da Mata Atlântica brasileira.** Orientador: Sebastiao Martins Filho. Coorientador: Carlos Moreira Miquelino Eleto Torres.

As florestas fornecem múltiplas funções e serviços ecossistêmicos, incluindo o fornecimento de bens por meio da produção primária, regulação do clima por meio do sequestro de carbono e serviços hidrológicos. Dentre as florestas tropicais brasileiras, a Mata Atlântica é considerada uma das mais importantes em termos de biodiversidade. No entanto, sofreu uma drástica redução da sua área devido a atividades antrópicas. Por tanto, identificar e entender os fatores que impulsionam a persistência dos ecossistemas florestais em meio às perturbações é essencial. O objetivo geral deste trabalho foi avaliar diferentes técnicas de análise de sobrevivência para modelar o efeito de covariáveis ambientais e antrópicas no desenvolvimento de fragmentos de Mata Atlântica. O trabalho foi dividido em três capítulos. No primeiro capítulo o objetivo foi avaliar a técnica de análise de sobrevivência intervalar para modelar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. A metodologia avaliou com eficácia o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. No segundo capítulo avaliou-se a mortalidade individual de árvores de um fragmento de Mata Atlântica, por meio da técnica de análise de sobrevivência com dados discretos. Esta metodologia apresentou uma acurácia superior a 80% na avaliação da mortalidade de árvores na Mata Atlântica. No terceiro capítulo foi utilizada a abordagem de análise de sobrevivência, por meio de métodos de aprendizado de máquina, para estudar a mortalidade de árvores da Mata Atlântica, em função de covariáveis ambientais e antrópicas. A metodologia random survival forests apresentou os maiores valores de índice de concordância e menores valores do Brier-Score, sendo eficiente em avaliar a mortalidade individual de árvores nos fragmentos de Mata Atlântica. Os métodos aplicados neste estudo podem servir como ferramentas estratégicas para a tomada de decisões por gestores florestais, contribuindo para a conservação e manejo sustentável da Mata Atlântica.

Palavras-chave: Dados censurados, Modelos paramétricos, Modelos de regressão discretos, Aprendizado de máquina, Florestas tropicais, Mortalidade, Recrutamento.

## ABSTRACT

DUARTE, Marciel Lelis, D.Sc., Universidade Federal de Viçosa, March, 2025. **Survival analysis to assess demographic processes of trees in the Brazilian Atlantic Forest.** Adviser: Sebastiao Martins Filho. Co-adviser: Carlos Moreira Miquelino Eleto Torres.

Forests provide multiple ecosystem functions and services, including the provision of goods through primary production, climate regulation through carbon sequestration, and hydrological services. Among the Brazilian tropical forests, the Atlantic Forest is considered one of the most important in biodiversity. However, its area has been drastically reduced due to human activities. It is virtually impossible to isolate or measure, at the field scale, all the biotic and abiotic components that influence the development of a forest, making it essential to identify and understand the factors that drive the persistence of forest ecosystems amid disturbances. The overall objective of this work was to evaluate different survival analysis techniques to model the effect of environmental and anthropogenic covariates on the development of Atlantic Forest fragments. The work was divided into three chapters. The first chapter aimed to evaluate the interval survival analysis technique to model the effect of environmental covariates on tree recruitment and mortality in Atlantic Forest fragments. A methodology effectively evaluated the effect of environmental covariates on tree recruitment and mortality in fragments of the Atlantic Forest. The second chapter evaluated the individual mortality of trees in an Atlantic Forest fragment using the survival analysis technique with discrete data. This methodology showed an accuracy of over 80% in assessing tree mortality in the Atlantic Forest. The third chapter used a survival analysis approach using machine learning methods to study tree mortality in the Atlantic Forest as a function of environmental and anthropogenic covariates. The random forest survival methodology presented the highest concordance index values and the lowest Brier-Score values, being efficient in assessing individual tree mortality in Atlantic Forest fragments. The methods applied in this study can serve as strategic tools for managers' decision-making, contributing to forest conservation and sustainable management of the Atlantic Forest.

Keywords: Censored data, Parametric models, Discrete regression models, Machine learning, Tropical forests, Mortality, Recruitment.

## SUMÁRIO

INTRODUÇÃO GERAL.....	7
CAPÍTULO 1 - INFLUÊNCIA DE ELEMENTOS DO CLIMA E SOLO EM GRUPOS ECOLÓGICOS DE ARVORES DE MATA ATLÂNTICA: UMA ABORDAGEM DA ANÁLISE DE SOBREVIVÊNCIA INTERVALAR. ....	12
RESUMO .....	12
INTRODUÇÃO .....	13
MATERIAL E MÉTODOS.....	15
RESULTADOS.....	20
DISCUSSÃO .....	25
CONCLUSÃO .....	27
REFERÊNCIAS.....	27
MATERIAL SUPLEMENTAR .....	34
CAPITULO 2 - ANÁLISE DE SOBREVIVÊNCIA COM DADOS GRUPADOS: MORTALIDADE DE ÁRVORES EM UM FRAGMENTO DE MATA ATLÂNTICA NO BRASIL .....	37
RESUMO .....	37
INTRODUÇÃO .....	38
MATERIAL E MÉTODOS.....	39
RESULTADOS.....	45
DISCUSSÃO .....	51
CONCLUSÃO .....	53
REFERENCIAS.....	53
CAPITULO 3 – MÉTODOS DE ANÁLISE DE SOBREVIVÊNCIA BASEADOS EM APRENDIZADO DE MÁQUINA: MODELAGEM DA MORTALIDADE DE ARVORES NA MATA ATLÂNTICA .....	58
RESUMO .....	58
INTRODUÇÃO .....	58
MATERIAL E METODOS.....	60
RESULTADOS.....	68
DISCUSSÃO .....	73
CONCLUSÃO .....	75
REFERENCIAS.....	75
CONCLUSÕES GERAIS.....	81

## INTRODUÇÃO GERAL

A Mata Atlântica é o bioma mais ameaçado do Brasil (Rocha et al., 2024), com apenas 12,4% de sua cobertura original remanescente (SOS Mata Atlântica, 2024). Esse ecossistema desempenha um papel fundamental na manutenção da biodiversidade, abrigando mais de 20.000 espécies de plantas (Campoe et al., 2014). Além disso, fornece uma ampla gama de serviços ecossistêmicos essenciais, incluindo regulação hídrica, modulação do clima, provisão de alimentos, madeira e compostos de interesse farmacológico (Bullock et al., 2011; Melo et al., 2013).

O monitoramento da dinâmica florestal em biomas ameaçados é essencial para compreender as respostas dessas florestas às mudanças no uso e cobertura da terra, bem como às alterações climáticas globais (Bustamante et al., 2016). A dinâmica florestal resulta da interação entre fatores ambientais, perturbações e processos demográficos (crescimento, mortalidade e recrutamento), os quais influenciam diretamente a biomassa e a composição de espécies (McDowell et al., 2020; Torres et al., 2023; Xu et al., 2016).

Embora seja relativamente simples e fácil enumerar os fatores que influenciam o recrutamento ou a mortalidade da floresta, a compreensão e avaliação da soma de interação desses fatores e seus efeitos no local são complexos (Freitas et al., 2020). Ainda mais em florestas tropicais brasileiras, que apresentam grande heterogeneidade, variabilidade temporal e espacial (Fien et al., 2019; Van der Sande et al., 2018). Nesse contexto, a construção de modelos de mortalidade em nível individual de árvore, que considerem covariáveis ambientais e antrópicas, torna-se essencial para entender a dinâmica nessas florestas.

Nas florestas tropicais, a modelagem da mortalidade em nível de árvore individual, geralmente tem sido realizada por meio de regressão logística (Figueredo et al., 2020; Rocha et al., 2020) ou redes neurais artificiais (Castro et al., 2015; Rocha et al., 2018; Reis et al., 2018). No entanto, estes modelos tem apresentado baixa precisão, o que pode estar ligada ao fato de a mortalidade em florestas tropicais ser um evento raro (King et al., 2000, Reis et al., 2018).

Além disso, um ponto comum nessas abordagens é que elas não levam em consideração a censura, que corresponde a uma observação parcial da resposta de interesse. Colosimo e Giolo (2024) ressaltam o fato de que todas as observações provenientes de um estudo (completas e parciais) devem ser utilizadas na análise estatística. Esses autores afirmam que, mesmo sendo incompletas, as censuras fornecem informações relevantes sobre o tempo do evento e sua omissão no cálculo das estatísticas de interesse pode acarretar conclusões viesadas. Sendo assim, são necessários métodos estatísticos que possibilitem incorporar na análise a informação contida tanto nas observações completas quanto nas censuras.

A análise de sobrevivência é um conjunto de técnicas e modelos estatísticos usados na análise de experimentos, em que a variável resposta é o tempo avaliado até a ocorrência de um evento de interesse, denominado tempo de falha (Colosimo e Giolo, 2024). Essa metodologia tem como principal característica a presença de dados censurados. A censura ocorre quando o estudo é interrompido, seja porque o estudo terminou para a análise dos dados sem a ocorrência do evento de interesse, ou o indivíduo falhou por uma causa diferente e não associada àquela estudada. Isto significa que a informação referente à resposta se resume ao conhecimento de que o tempo até o evento é superior ao observado, e este conhecimento precisa ser incorporado na análise.

A fim de quantificar e comparar a suscetibilidade de diferentes espécies ou tipos de floresta a riscos naturais, a técnica de análise de sobrevivência têm sido cada vez mais usadas nas ciências florestais (Brandl et al., 2020; Griess et al., 2012; Neuner et al., 2015; Neumann et al., 2017; Nothdurft 2013; Paul et al., 2019; Staupendahl e Möhring, 2011; Thiele et al., 2017). No entanto, todos estes estudos foram realizados em florestas de clima temperado, onde a idade das árvores na maioria das vezes é conhecida. Outro ponto é a que a modelagem utilizada nestes trabalhos foi realizada sempre considerando o tempo de falha exato, que pode nem sempre ser a solução mais apropriada, uma vez que os dados de inventário florestal, geralmente são realizados em intervalos de três, quatro ou cinco anos. Assumir que dados coletados em intervalos tenham tempo exato de falha pode conduzir a viés, bem como resultados e conclusões não confiáveis (Bogaerts et al., 2017; Colosimo e Giolo, 2024).

Desta forma são necessário pesquisas com uso de modelos de análise de sobrevivência em florestas tropicais. Estes modelos podem fornecer informações que ajudem a subsidiar práticas de manejo mais eficientes e sustentáveis. Portanto, o objetivo geral deste trabalho foi avaliar diferentes técnicas de análise de sobrevivência para modelar o efeito de covariáveis ambientais e antrópicas no desenvolvimento de fragmentos de Mata Atlântica.

Este trabalho está dividido em três capítulos. No primeiro capítulo o objetivo foi avaliar a técnica de análise de sobrevivência intervalar para modelar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. No capítulo dois avaliou-se a mortalidade individual de árvores de um fragmento de Mata Atlântica, por meio da técnica de análise de sobrevivência com dados discretos. O último capítulo teve como objetivo avaliar o uso de métodos de análise de sobrevivência baseados em aprendizado de máquina na mortalidade de árvores da Mata Atlântica.

## REFERÊNCIAS

- Bogaerts, K., Komarek, A., Lesaffre, E., 2017. *Survival Analysis with Interval-Censored Data: A Practical Approach with Examples in R, SAS, and BUGS*. Chapman and Hall/CRC, New York. <https://doi.org/10.1201/9781315116945>
- Brandl, S., Paul, C., Knoke, T., Falk, W., 2020. The influence of climate and management on survival probability for Germany's most important tree species. *Forest Ecology and Management* 458, 117652. <https://doi.org/10.1016/j.foreco.2019.117652>
- Bullock, J.M., Aronson, J., Newton, A.C., Pywell, R.F., Rey-Benayas, J.M., 2011. Restoration of ecosystem services and biodiversity: conflicts and opportunities. *Trends in Ecology & Evolution* 26, 541–549. <https://doi.org/10.1016/j.tree.2011.06.011>
- Bustamante, M.M.C., Roitman, I., Aide, T.M., Alencar, A., Anderson, L.O., Aragão, L., Asner, G.P., Barlow, J., Berenguer, E., Chambers, J., Costa, M.H., Fanin, T., Ferreira, L.G., Ferreira, J., Keller, M., Magnusson, W.E., Morales-Barquero, L., Morton, D., Ometto, J.P.H.B., Palace, M., Peres, C.A., Silvério, D., Trumbore, S., Vieira, I.C.G., 2016. Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology* 22, 92–109. <https://doi.org/10.1111/gcb.13087>
- Campoe, O.C., Iannelli, C., Stape, J.L., Cook, R.L., Mendes, J.C.T., Vivian, R., 2014. Atlantic forest tree species responses to silvicultural practices in a degraded pasture restoration plantation: From leaf physiology to survival and initial growth. *Forest Ecology and Management* 313, 233–242. <https://doi.org/10.1016/j.foreco.2013.11.016>
- Castro, R.V.O., Soares, C.P.B., Leite, H.G., de Souza, A.L., Martins, F.B., Nogueira, G.S., 2015. Artificial Neural Networks Effectiveness to Estimate Mortality in a Semi-Deciduous Seasonal Forest. *Australian Journal of Basic and Applied Sciences*.
- Colosimo, E.A., Giolo, S.R., 2024. *Análise de sobrevivência aplicada*, 2nd ed. Editora Edgard Blucher, São Paulo, SP.
- Fien, E.K.P., Fraver, S., Teets, A., Weiskittel, A.R., Hollinger, D.Y., 2019. Drivers of individual tree growth and mortality in an uneven-aged, mixed-species conifer forest. *Forest Ecology and Management* 449, 117446. <https://doi.org/10.1016/j.foreco.2019.06.043>
- Figueiredo, L.T.M.D., Soares, C.P.B., Castro, R.V.O., Leite, H.G., Souza, A.L.D., Silva, G.F.D., Torres, C.M.M.E., 2020. Modeling Mortality of Individual Trees in a Seasonal Semi-Deciduous Forest. *Floresta Ambient.* 27, e20170980. <https://doi.org/10.1590/2179-8087.098017>
- Freitas, E.C.S. de, Paiva, H.N. de, Neves, J.C.L., Marcatti, G.E., Leite, H.G., 2020. Modeling of eucalyptus productivity with artificial neural networks. *Industrial Crops and Products* 146, 112149. <https://doi.org/10.1016/j.indcrop.2020.112149>
- Griess, V.C., Acevedo, R., Härtl, F., Staupendahl, K., Knoke, T., 2012. Does mixing tree species enhance stand resistance against natural hazards? A case study for spruce. *Forest Ecology and Management* 267, 284–296. <https://doi.org/10.1016/j.foreco.2011.11.035>
- King, S.L., Bennett, K.P., List, S., 2000. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. *Computers and Electronics in Agriculture* 27, 401–406. [https://doi.org/10.1016/S0168-1699\(00\)00106-X](https://doi.org/10.1016/S0168-1699(00)00106-X)
- McDowell, N.G., Allen, C.D., Anderson-Teixeira, K., Aukema, B.H., Bond-Lamberty, B.,

- Chini, L., Clark, J.S., Dietze, M., Grossiord, C., Hanbury-Brown, A., Hurtt, G.C., Jackson, R.B., Johnson, D.J., Kueppers, L., Lichstein, J.W., Ogle, K., Poulter, B., Pugh, T.A.M., Seidl, R., Turner, M.G., Uriarte, M., Walker, A.P., Xu, C., 2020. Pervasive shifts in forest dynamics in a changing world. *Science* 368, eaaz9463. <https://doi.org/10.1126/science.aaz9463>
- Melo, D.H.A., Filgueiras, B.K.C., Iserhard, C.A., Iannuzzi, L., Freitas, A.V.L., Leal, I.R., 2019. Effect of habitat loss and fragmentation on fruit-feeding butterflies in the Brazilian Atlantic Forest. *Can. J. Zool.* 97, 588–596. <https://doi.org/10.1139/cjz-2018-0202>
- Neumann, M., Mues, V., Moreno, A., Hasenauer, H., Seidl, R., 2017. Climate variability drives recent tree mortality in Europe. *Global Change Biology* 23, 4788–4797. <https://doi.org/10.1111/gcb.13724>
- Neuner, S., Albrecht, A., Cullmann, D., Engels, F., Griess, V.C., Hahn, W.A., Hanewinkel, M., Härtl, F., Kölling, C., Staupendahl, K., Knoke, T., 2015. Survival of Norway spruce remains higher in mixed stands under a dryer and warmer climate. *Global Change Biology* 21, 935–946. <https://doi.org/10.1111/gcb.12751>
- Nothdurft, A., 2013. Spatio-temporal prediction of tree mortality based on long-term sample plots, climate change scenarios and parametric frailty modeling. *Forest Ecology and Management* 291, 43–54. <https://doi.org/10.1016/j.foreco.2012.11.028>
- Paul, C., Brandl, S., Friedrich, S., Falk, W., Härtl, F., Knoke, T., 2019. Climate change and mixed forests: how do altered survival probabilities impact economically desirable species proportions of Norway spruce and European beech? *Annals of Forest Science* 76, 1–15. <https://doi.org/10.1007/s13595-018-0793-8>
- Reis, L.P., de Souza, A.L., dos Reis, P.C.M., Mazzei, L., Soares, C.P.B., Miquelino Eleto Torres, C.M., da Silva, L.F., Ruschel, A.R., Rêgo, L.J.S., Leite, H.G., 2018. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. *Ecological Engineering* 112, 140–147. <https://doi.org/10.1016/j.ecoleng.2017.12.014>
- Rocha, S.J.S.S. da, 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. da, Reis, L.P., Zanuncio, J.C., 2018. Artificial neural networks: Modeling tree survival and mortality in the Atlantic Forest biome in Brazil. *Science of The Total Environment* 645, 655–661. <https://doi.org/10.1016/j.scitotenv.2018.07.123>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Schettini, B.L.S., Jacovine, L.A.G., Leite, H.G., Gelcer, E.M., Reis, L.P., Neves, K.M., Comini, I.B., Silva, L.F. da, 2020. Drought effects on carbon dynamics of trees in a secondary Atlantic Forest. *Forest Ecology and Management* 465, 118097. <https://doi.org/10.1016/j.foreco.2020.118097>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Tavares Júnior, I. da S., Rufino, M.P.M.X., Romero, F.M.B., Jacovine, L.A.G., de Moraes Junior, V.T.M., França, L.C. de 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., 2024. Machine learning methods: Modeling net growth in the Atlantic Forest of Brazil. *Ecological Informatics* 81, 102564. <https://doi.org/10.1016/j.ecoinf.2024.102564>
- SOS Mata Atlântica, 2024. Atlas da Mata Atlântica. <https://www.sosma.org.br/iniciativas/atlas-da-mata-atlantica>.
- Staupendahl, K., Möhring, B., 2011. Integrating natural risks into silvicultural decision models:

- A survival function approach. *Forest Policy and Economics* 13, 496–502. <https://doi.org/10.1016/j.forpol.2011.05.007>
- Thiele, J.C., Nuske, R.S., Ahrends, B., Panferov, O., Albert, M., Staupendahl, K., Junghans, U., Jansen, M., Saborowski, J., 2017. Climate change impact assessment—A simulation experiment with Norway spruce for a forest district in Central Europe. *Ecological Modelling* 346, 30–47. <https://doi.org/10.1016/j.ecolmodel.2016.11.013>
- Torres, C.M.M.E., Medina-Vega, J.A., Rocha, S.J.S.S. da, Costa, W. da S., Soares, C.P.B., Souza, A.L. de, Jacovine, L.A.G., Lana, J.M. de, Peña-Claros, M., 2023. Drivers of tree demographic processes in forest fragments of the Brazilian Atlantic forest. *Forest Ecology and Management* 534, 120893. <https://doi.org/10.1016/j.foreco.2023.120893>
- van der Sande, M.T., Arets, E.J.M.M., Peña-Claros, M., Hoosbeek, M.R., Cáceres-Siani, Y., van der Hout, P., Poorter, L., 2018. Soil fertility and species traits, but not diversity, drive productivity and biomass stocks in a Guyanese tropical rainforest. *Functional Ecology* 32, 461–474. <https://doi.org/10.1111/1365-2435.12968>
- Xu, X., Medvigy, D., Powers, J.S., Becknell, J.M., Guan, K., 2016. Diversity in plant hydraulic traits explains seasonal and inter-annual variations of vegetation dynamics in seasonally dry tropical forests. *New Phytologist* 212, 80–95. <https://doi.org/10.1111/nph.14009>

## **CAPÍTULO 1 - INFLUÊNCIA DE ELEMENTOS DO CLIMA E SOLO EM GRUPOS ECOLÓGICOS DE ÁRVORES DE MATA ATLÂNTICA: UMA ABORDAGEM DA ANÁLISE DE SOBREVIVÊNCIA INTERVALAR.**

### **RESUMO**

As florestas fornecem múltiplas funções e serviços ecossistêmicos, incluindo o fornecimento de bens por meio da produção primária, regulação do clima por meio do sequestro de carbono e serviços hidrológicos. Dentre as florestas tropicais brasileiras, a Mata Atlântica é considerada uma das mais importantes em termos de biodiversidade. No entanto, sofreu uma drástica redução da sua área devido a atividades antrópicas. No campo, a quantificação ou isolamento de todos os fatores bióticos e abióticos que afetam o desenvolvimento de uma floresta é uma tarefa extremamente desafiadora. Nesse sentido, torna-se essencial identificar e entender os fatores que impulsionam a persistência dos ecossistemas florestais em meio às perturbações. O objetivo deste trabalho foi avaliar a técnica de análise de sobrevivência intervalar para modelar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. Foram utilizados dados de 54 parcelas permanentes de quatro fragmentos florestais em Minas Gerais, Brasil. Foram utilizados como covariáveis ambientais características químicas do solo (pH do solo em H<sub>2</sub>O, cátions trocáveis, teor de P e matéria orgânica do solo) e clima (temperatura média, precipitação < 100 mm, precipitação nos três meses mais secos, precipitação média anual total e déficit hídrico - CWD). Foram ajustados modelos de sobrevivência intervalar utilizando as distribuições exponencial, Weibull, Log-normal e Log-logístico para estudar a influência das covariáveis ambientais na mortalidade e recrutamento. A escolha do melhor modelo ajustado foi baseada no Critério de Informação de Akaike (AIC) e na análise gráfica. O modelo de Weibull foi o melhor modelo para descrever o recrutamento e o Log-normal para mortalidade. A influência das covariáveis ambientais na mortalidade e no recrutamento variaram com o grupo ecológico das árvores. As espécies pioneiras, no recrutamento, sofreram maior influência do pH, P, K<sup>+</sup>, Al<sup>3+</sup>, precipitação nos meses mais secos e CWD\_1 e do P, K<sup>+</sup>, Al<sup>3+</sup>, MO, precipitação nos meses mais secos e CWD\_3, na mortalidade. As espécies secundárias iniciais foram influenciadas pelo pH, Ca<sup>2+</sup> e precipitação nos meses mais secos, no recrutamento e na mortalidade por Ca<sup>2+</sup>, Al<sup>3+</sup>, P-Rem e CWD\_3. As secundárias tardias foram influenciadas pelo pH, P e CWD\_3 no recrutamento e na mortalidade. A metodologia de análise de sobrevivência intervalar foi eficiente em avaliar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. Esta técnica se mostrou uma ferramenta eficaz para analisar como os efeitos

ambientais influenciam os processos demográficos de diferentes grupos ecológicos.

**Palavras-chave:** Dados censurados, Modelos paramétricos, Mortalidade, Recrutamento.

## INTRODUÇÃO

As florestas tropicais fornecem múltiplas funções e serviços ecossistêmicos, como conservação da biodiversidade, fornecimento de produtos florestais madeireiros e não madeireiros, regulação do clima por meio do sequestro de carbono, ciclagem de nutrientes e decomposição de matéria orgânica e serviços hidrológicos (Borja et al., 2021; Byrnes et al., 2014; Costa et al., 2017; Kozłowski, 2002). Nos últimos anos, as florestas tropicais vêm sofrendo grandes perdas pelo desmatamento intenso (Ribeiro et al., 2019; Solórzano et al., 2021), que tem ocasionado a fragmentação da cobertura florestal (Taubert et al., 2018), e como consequência a perda de biodiversidade nessas florestas (Haddad et al., 2015; Melo et al., 2019).

Dentre as florestas tropicais brasileiras, a Mata Atlântica é considerada uma das mais importantes em termos de biodiversidade (Delgado et al., 2018; Joly et al., 2014). No entanto, sofreu uma drástica redução da sua área devido a atividades antrópicas (Dos Santos et al., 2020), e hoje é composta por pequenos fragmentos florestais com diferentes níveis de perturbação (Arroyo-Rodríguez et al., 2017; Torres et al., 2023; Tabarelli et al., 2010), ocupando apenas 12,4% da cobertura original (SOS Mata Atlântica, 2024).

Os níveis de perturbação ou distúrbios, podem influenciar os processos demográficos (crescimento, mortalidade e recrutamento) que são responsáveis pelas mudanças na composição das espécies durante a sucessão em florestas tropicais (Dalmaso et al., 2020; Souza et al., 2021). Estes processos são afetados por covariáveis climáticas (Rocha et al., 2020) e de solo (Capellesso et al., 2020; Van der Sande et al., 2017).

Embora seja relativamente simples e fácil enumerar os fatores que influenciam o crescimento ou a mortalidade da floresta, a compreensão e avaliação da soma de interação desses fatores e seus efeitos no local são complexos (Freitas et al., 2020). Ainda mais em florestas tropicais, que apresentam grande heterogeneidade, variabilidade temporal e espacial (Fien et al., 2019; Van der Sande et al., 2018).

As estratégias de manejo florestal devem, portanto, considerar os efeitos ambientais para tentar garantir um grau adequado de estabilidade das florestas. Uma compreensão detalhada dos processos que determinam a dinâmica do ecossistema florestal é essencial para o desenvolvimento de planos de manejo bem ajustados. Nesse ponto a modelagem do crescimento, recrutamento e da mortalidade das árvores é importante tanto para o manejo

florestal quanto para a inferência ecológica.

Algumas pesquisas fazem uso de modelos baseados nos métodos de mínimos quadrados para prever a mortalidade em nível de povoamento. Porém, dados de medidas repetidas de parcelas permanentes, muitas vezes exibem vários graus de dispersão e assimetria e, portanto, abordagens de modelagem de mortalidade baseadas no método dos mínimos quadrados podem não ser apropriadas para tais dados (Thapa et al., 2016). Além disso, estes modelos não levam em consideração a censura, que corresponde a observação parcial da resposta de interesse.

Na presença de censura, técnicas baseadas em métodos dos mínimos quadrados se tornam inviáveis, pois não permitem o uso de observações parciais. Colosimo e Giolo (2024) ressaltam o fato de que todas as observações provenientes de um estudo (completas e parciais) devem ser utilizadas na análise estatística. Esses autores afirmam que, mesmo sendo incompletas, as censuras fornecem informações relevantes sobre o tempo do evento e sua omissão no cálculo das estatísticas de interesse pode acarretar conclusões viesadas. Sendo assim, são necessários métodos estatísticos que possibilitem incorporar na análise a informação contida tanto nas observações completas quanto nas censuras.

A análise de sobrevivência é um conjunto de técnicas e modelos estatísticos usados na análise de experimentos, em que a variável resposta é o tempo avaliado até a ocorrência de um evento de interesse, denominado tempo de falha (Colosimo e Giolo, 2024). Essa metodologia tem como principal característica a presença de dados censurados. A censura ocorre quando o estudo é interrompido, seja porque o estudo terminou para a análise dos dados sem a ocorrência do evento de interesse, ou a unidade experimental morreu por uma causa diferente e não associada àquela estudada. Isto significa que a informação referente à resposta se resume ao conhecimento de que o tempo até o evento é superior ao observado, e este conhecimento precisa ser incorporado na análise.

A fim de quantificar e comparar a suscetibilidade de diferentes espécies ou tipos de floresta a riscos naturais, a técnica de análise de sobrevivência têm sido cada vez mais usadas nas ciências florestais (Brandl et al., 2020; Griess et al., 2012; Neuner et al., 2015; Neumann et al., 2017; Nothdurft 2013; Paul et al., 2019; Staupendahl e Möhring, 2011; Thiele et al., 2017). No entanto, todos estes estudos foram realizados em florestas de clima temperado, onde a idade das árvores na maioria das vezes é conhecida. Outro ponto é a que a modelagem utilizada nestes trabalhos foi realizada sempre considerando o tempo de falha exato, que pode nem sempre ser a solução mais apropriada, uma vez que os dados de inventário florestal, geralmente são realizados em intervalos de três, quatro ou cinco anos. Assumir que dados coletados em intervalos tenham tempo exato de falha pode conduzir a vícios, bem como resultados e

conclusões não confiáveis (Bogaerts et al., 2017; Colosimo e Giolo, 2024).

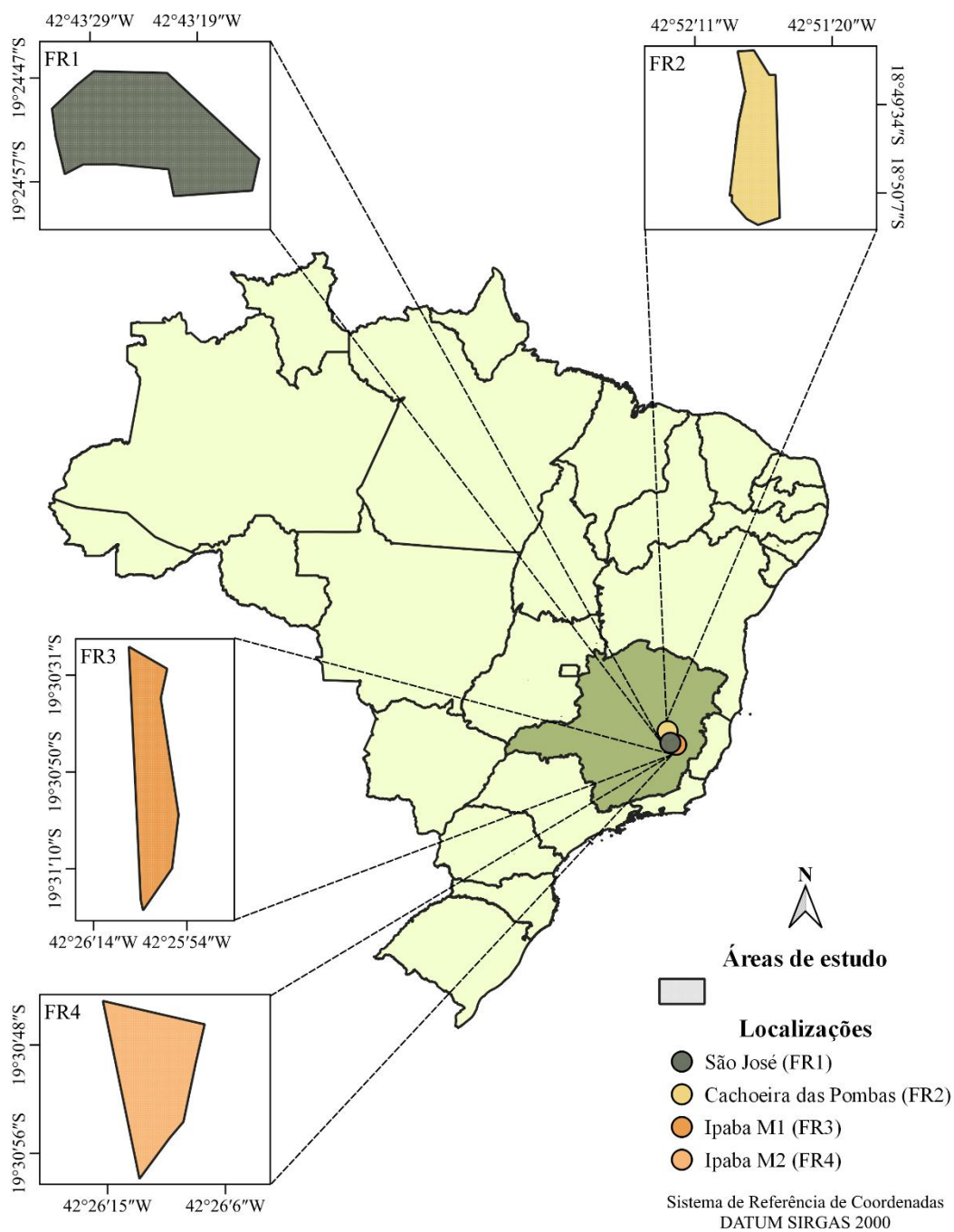
Desta forma, são necessárias pesquisas com uso de modelos de análise de sobrevivência intervalar para avaliar o efeito de covariáveis ambientais em florestas de clima tropical. O objetivo deste trabalho foi avaliar a técnica de análise de sobrevivência intervalar para estudar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. Foi abordado neste trabalho as seguintes questões: (i) Os grupos ecológicos diferem nas taxas de recrutamento e mortalidade? (ii) Quais são os efeitos das covariáveis ambientais no recrutamento e mortalidade em cada grupo ecológico?

## **MATERIAL E MÉTODOS**

### **Local de estudo e características das parcelas**

Foram utilizados dados de quatro fragmentos de Mata Atlântica, localizados em Minas Gerais, Brasil (Figura 1). A vegetação é classificada como floresta estacional semidecidual (IBGE, 2012) e os fragmentos variam em clima e solo (Tabela 1). Todos os fragmentos são florestas secundárias em regeneração após serem desmatados e/ou usados para extração de madeira, agricultura, plantações de eucalipto. Atualmente, todos os fragmentos são áreas de conservação ou proteção, garantindo a sucessão ecológica nos últimos anos.

Nos quatro fragmentos florestais foram implantadas 54 parcelas com 500 m<sup>2</sup> (10 × 50 m). Todos os fustes com diâmetro à altura do peito  $\geq 5$  cm (1,3 m de altura, DAP) foram recrutados, medidos e identificados botanicamente. As parcelas foram monitoradas ao longo dos anos de 2002, 2007, 2012 e 2017. As espécies foram classificadas em grupos ecológicos seguindo a divisão proposta por Gandolfi et al. (1995), para florestas secundárias brasileiras em: pioneiras (P); secundárias iniciais (Si) e secundárias tardias (St). A mortalidade e o recrutamento foram avaliados considerando cada fuste das árvores como um indivíduo. Na mortalidade, os fustes que permaneceram vivos até o final do monitoramento foram classificados como censurados. Para o recrutamento, foram considerados censurados os fustes que atingiram o DAP mínimo para recrutamento antes de 2002.



**Figura 1** – Localização dos quatro fragmentos de Mata Atlântica no estado de Minas Gerais, Brasil. FR1: São José; FR2: Cachoeira das Pombas; FR3: Ipaba Mata1; FR4: Ipaba Mata 2.

**Tabela 1** – Caracterização de quatro fragmentos de Mata Atlântica em Minas Gerais, Brasil, por meio de média e desvio padrão, das características de clima e solo. CWD = déficit hídrico climático; MOS – Matéria orgânica do solo; BA – Área basal.

	Cachoeira das Pombas	Ipaba Mata 1	Ipaba Mata 2	São José
<b>Clima</b>				
Precipitação (mm)	1112.74 ±214.50	1194.38 ±298.96	1194.38 ±298.96	1263.92 ±226.07
CWD (mm)	-1021.36 ±432.50	-2020.77 ±563.16	-1499.09 ±502.88	-940.19 ±395.45
Precipitação meses mais secos (mm)	30.84 ±33.57	33.99 ±30.13	33.99 ±30.13	39.84 ±30.99
Precipitação Menos 100 mm	7.82 ±1.17	7.53 ±1.28	7.53 ±1.28	7.38 ±1.18
Temperatura Média (°C)	20.85 ±1.59	24.15 ±1.73	24.15 ±1.73	19.51 ±1.97
<b>Solo</b>				
pH (H <sub>2</sub> O)	4.52 ±0.25	3.94 ±0.15	3.76 ±0.08	4.23 ±0.55
K (cmolc.dm <sup>-3</sup> )	0.16 ±0.11	0.08 ±0.01	0.07 ±0.01	0.09 ±0.03
Ca <sup>2+</sup> (cmolc.dm <sup>-3</sup> )	0.49 ±0.45	0.31 ±0.08	0.16 ±0.07	0.49 ±0.56
Mg <sup>2+</sup> (cmolc.dm <sup>-3</sup> )	0.24 ±0.27	0.17 ±0.06	0.11 ±0.01	0.16 ±0.10
Al <sup>3+</sup> (cmolc.dm <sup>-3</sup> )	1.06 ±0.43	1.40 ±0.38	1.82 ±0.17	0.91 ±0.34
P-rem (mg.L <sup>-1</sup> )	10.69 ±2.07	21.54 ±2.32	21.85 ±1.65	21.23 ±5.17
MOS (dag.kg <sup>-1</sup> )	5.64 ±1.75	3.89 ±0.57	3.00 ±0.28	5.14 ±1.56
<b>Paisagem</b>				
Nº Fustes (ha)	2008 ±664	1765 ±454	1825 ±264	1616 ±344
BA (m <sup>2</sup> .ha <sup>-1</sup> )	18.85 ±6.32	16.46 ±6.05	25.38 ±7.61	22.99 ±8.45
Parcelas	20	16	6	12

### Covariáveis climáticas

De estações meteorológicas próximas aos fragmentos, foram obtidos a precipitação anual, o número de meses com precipitação < 100 mm, a precipitação nos três meses mais secos e a temperatura média anual de cada fragmento da estação climatológica mais próxima. A partir destes dados foi estimado o déficit hídrico climático (CWD), que é a diferença entre a Evapotranspiração Potencial (ETP) e a Evapotranspiração Real (ETR), de acordo a metodologia proposta por Lutz et al. (2010), utilizando as funções CWD e AET (Redmond, 2022) do *software R* versão 4.4.2. (R Development Core Team, 2024). O CWD reflete as condições de seca com mais precisão do que a precipitação anual total (Chave et al., 2014). CWD com menores valores negativos indica condições de alto estresse hídrico, e valores

próximos a 0 (zero) indicam nenhum estresse hídrico (Poorter et al., 2017). A precipitação média anual total e CWD, foram calculados para quatro, três, dois e um ano antes do inventário de cada parcela.

### **Covariáveis de solo e topografia**

Foram coletados aleatoriamente de 20 a 30 amostras de solo para obter uma amostra composta para duas profundidades (0–20 cm e 20–40 cm) em cada parcela. A partir destas amostras foram determinados: pH do solo em H<sub>2</sub>O, cátions trocáveis (Ca<sup>2+</sup>, Mg<sup>2+</sup> e Al<sup>3+</sup>), acidez total (H<sup>+</sup> + Al<sup>3+</sup>), capacidade de troca catiônica (CTC), saturação por bases (V), fósforo disponível (P), P remanescente em solução (P-rem) e matéria orgânica do solo (MOS), determinados por métodos padrão (Teixeira et al., 2017). As análises de solo foram realizadas no laboratório de Fertilidade do Solo da Universidade Federal de Viçosa (UFV), Brasil.

Para levar em conta as diferenças na topografia, foi calculado para cada parcela a elevação, a inclinação e o aspecto da inclinação usando a ferramenta de inclinação no ArcGIS 10.3.1. Para o cálculo foi utilizando um modelo digital de elevação (DEM) da Missão Espacial da NASA -Shuttle Radar Topography Mission (SRTM) para a análise.

### **Análise dos dados**

As covariáveis quantitativas foram padronizadas para o score  $Z$ , para melhorar a convergência do modelo.

$$Z_i = \frac{x_i - \bar{x}}{\sigma} \quad (1)$$

onde,  $Z_i$  é valor padronizado da observação  $i$ ;  $x_i$  é o valor da observação  $i$ ;  $\bar{x}$  é a média dos valores observados;  $\sigma$  é o desvio padrão.

### **Análise de Sobrevivência**

A forma mais eficiente de acomodar o efeito de várias covariáveis, é utilizar modelos de regressão para dados de sobrevivência. Uma classe importante de modelos propostas na literatura são a dos modelos paramétricos. Os principais modelos paramétricos usados na análise de sobrevivência são: modelo exponencial, de Weibull, Log-normal e Log-logístico, (Tabela 1). Estes modelos foram ajustados utilizando uma abordagem intervalar, para estudar a influência das covariáveis climáticas e de solo na mortalidade e recrutamento dos fustes das árvores. Os pacotes *icenReg* versão 2.0.16 (Anderson-Bergman, 2017) e *survival* versão 3.5-8

(Therneau, 2024) do do *software R* versão 4.4.2. (R Development Core Team, 2024), foram utilizados para ajustar os modelos.

**Tabela 2** - Funções de densidade de probabilidade e de sobrevivência para os modelos paramétricos.

Distribuição	Função Densidade de Probabilidade <sup>1</sup>	Função de Sobrevivência
Exponencial $\lambda > 0, t \geq 0$	$f(t) = \lambda \exp\{-\lambda t\}$	$S(t x) = \exp\left\{-\frac{t}{\exp(\beta_0 + \beta'x)}\right\}$
Weibull $\gamma, \lambda > 0, t \geq 0$	$f(t) = \lambda \gamma t^{\gamma-1} \exp\{-\lambda t^\gamma\}$	$S(t x) = \exp\left\{-\left(\frac{t}{\exp(\beta_0 + \beta'x)}\right)^{\frac{1}{\sigma}}\right\}$
Log-normal $\sigma > 0, t \geq 0$	$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln(t) - \mu}{\sigma}\right)^2\right\}$	$S(t x) = \Phi\left(\frac{-\ln(t) + \beta_0 + \beta'x}{\sigma}\right)$
Log-logística $\gamma, \lambda > 0, t \geq 0$	$f(t) = \frac{\gamma t^{\gamma-1} \lambda}{(1 + \lambda t^\gamma)^2}$	$S(t x) = \frac{1}{1 + \left(\frac{t}{\exp(\beta_0 + \beta'x)}\right)^{\frac{1}{\sigma}}}$

<sup>1/</sup>  $\lambda = \frac{1}{[\exp(\beta_0 + \beta'x)]^\gamma}$ ;  $\gamma = \frac{1}{\sigma}$ ;  $\ln$  = Logaritmo natural;  $\Phi$  = função de distribuição acumulada da distribuição normal padrão;  $\beta'$  = vetor de parâmetros a ser estimado;  $x$  = vetor de covariáveis.

A função de verossimilhança para os modelos de sobrevivência paramétricos intervalar é dada por:

$$L(\theta) = \prod_{i=1}^n [S(l_i|x_i) - S(u_i|x_i)]^{\delta_i} [S(l_i|x_i)]^{1-\delta_i} \quad (2)$$

em que  $x_i$  é o vetor de covariáveis associado ao  $i$ -ésimo indivíduo, bem como  $\delta_i = 1$ , se o evento ocorreu no intervalo de tempo  $(l_i; u_i]$ , e  $\delta_i = 0$ , se ocorreu em um tempo superior a  $l_i$ .

### Seleção de covariáveis

Inicialmente foi calculada a correlação de Spearman para verificar quais covariáveis seriam utilizadas nos modelos. Nas covariáveis com alta correlação ( $\geq 0,80$  em magnitude), somente uma delas foi mantida, baseado em sua consistência biológica. Para as covariáveis de solo, foram utilizados somente os dados da camada 0-20 cm, pois estavam correlacionados com os dados da camada de 20-40 cm. Foram excluídos os dados de acidez total, capacidade de troca catiônica e saturação por bases, por apresentar alta correlação com os cátions trocáveis.

As covariáveis climáticas números de meses com precipitação  $< 100$  mm, precipitação média anual total para um, dois, três e quatro anos antes do inventário florestal, apresentaram

alta correlação entre si, assim como o CWD para o intervalo de anos considerados. Para evitar multicolinearidade (Dormann et al., 2013) apenas uma variável de cada grupo foi selecionada por vez.

O segundo passo foi utilizar o método derivado da estratégia de seleção de covariáveis proposta por Collett (2023) e adaptada por Colosimo e Giolo (2024), no qual as informações do pesquisador podem ser incluídas no processo de tomada de decisão. A recomendação de incluir ou não um termo no modelo deve ser seguido de um nível de significância individual próximo de 0,10.

Um nível de significância de 5% foi usado para determinar a significância estatística das covariáveis selecionadas no modelo final, o teste utilizado foi o de Wald.

### **Seleção e avaliação dos modelos**

A escolha do melhor modelo ajustado foi baseada no Critério de Informação de Akaike (AIC) (Akaike, 1974). Essa medida usa a estimativa do logaritmo da verossimilhança ( $LL$ ), adicionando um termo penalizador associado ao número de parâmetros do modelo ( $p$ ). É bem conhecido que adicionando covariáveis, pode-se melhorar o ajuste dos modelos. Assim, o AIC tenta equilibrar a qualidade do ajuste *versus* a inclusão de covariáveis no modelo. Valores menores indicam melhor ajuste do modelo. O AIC é calculado da seguinte forma:

$$AIC = -2(LL) + 2p \quad (3)$$

Avaliar a adequação de um modelo ajustado é essencial na análise de dados. A adequação do melhor modelo para descrever a mortalidade e o recrutamento dos fustes foi realizada usando verificações gráficas, comparando a curva do modelo paramétrico ajustado com o método não-paramétrico (Bogaerts et al., 2017), por meio do pacote *icenReg* versão 2.0.16 (Anderson-Bergman, 2017).

## **RESULTADOS**

### **Recrutamento**

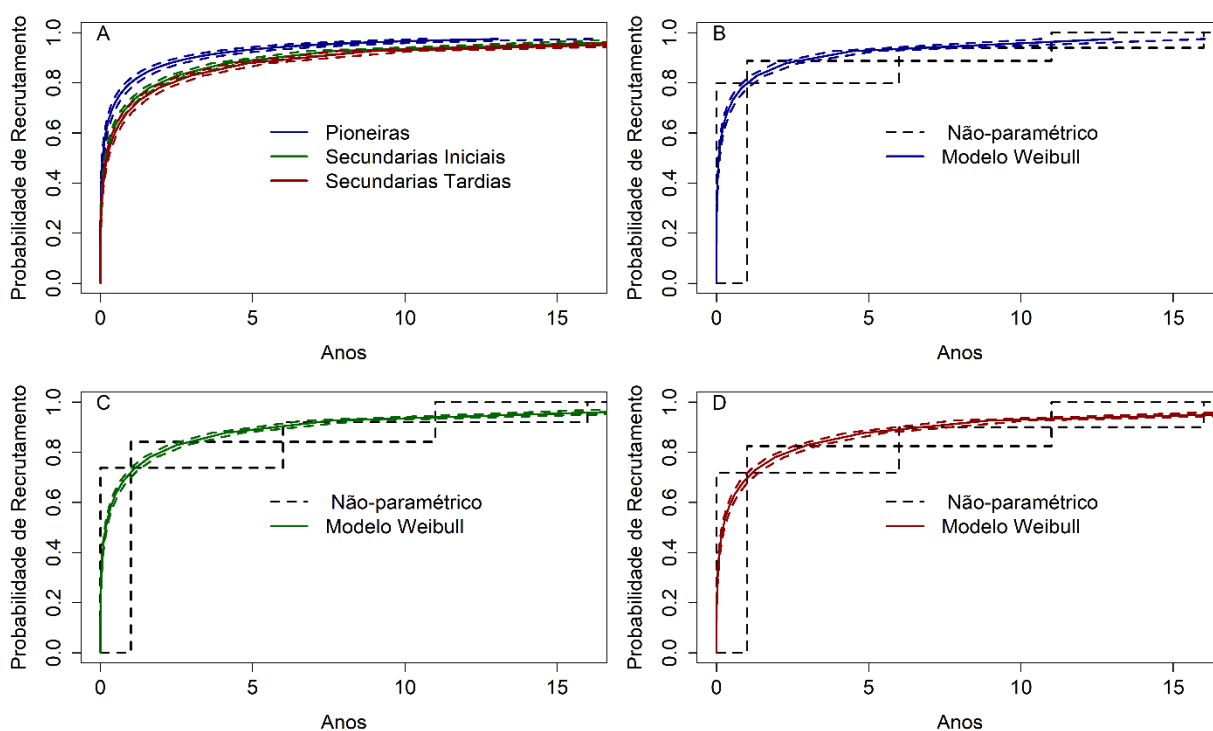
Os modelos paramétricos utilizados neste estudo foram ajustados para cada grupo ecológico em função das covariáveis ambientais e seus efeitos no recrutamento. A Tabela 3 apresenta as estimativas dos parâmetros necessários para estimar as funções de sobrevivência do modelo selecionado. O modelo escolhido para todos os grupos sucessionais foi o de Weibull, por ter apresentado menor valor de AIC.

**Tabela 3** - Parâmetros, estimativas dos parâmetros, fator de aceleração, erros-padrão (SE) e valor-*p* do teste de Wald para o modelo de Weibull no recrutamento de cada grupo ecológico.

Parâmetros	Estimativas	Fator de aceleração	SE	Valor- <i>p</i>
Pioneiras				
$\beta_0$	-1.4784		0.1402	<0.0001
$\beta_1$ (pH)	0.2649	1.3033	0.1481	0.02297
$\beta_2$ (P)	-0.6601	0.5168	0.1216	<0.0001
$\beta_3$ (K <sup>+</sup> )	-0.6887	0.5022	0.1548	<0.0001
$\beta_4$ (Al <sup>3+</sup> )	-0.6498	0.5221	0.1351	<0.0001
$\beta_5$ (Precip_Dry)	-1.5157	0.2196	0.1883	<0.0001
$\beta_6$ (CWD_1)	0.5475	1.7289	0.1331	<0.0001
$\sigma$	3.1		0.0434	<0.0001
Secundarias Iniciais				
$\beta_0$	-0.6485		0.0697	<0.0001
$\beta_1$ (pH)	0.5365	1.7104	0.0864	<0.0001
$\beta_2$ (Ca <sup>2+</sup> )	-0.1891	0.8270	0.0768	0.0140
$\beta_3$ (Precip_Dry)	-0.3643	0.6946	0.0946	<0.0001
$\sigma$	2.10		0.0284	<0.0001
Secundarias Tardias				
$\beta_0$	-0.7207		0.0744	<0.0001
$\beta_1$ (pH)	0.1559	1.1691	0.0904	<0.0001
$\beta_2$ (P)	-0.1948	0.8230	0.0804	0.0150
$\beta_3$ (CWD_3)	-0.2333	0.7923	0.0936	<0.0001
$\sigma$	2.18		0.0288	<0.0001

O fator de aceleração corresponde ao exp (estimativa). Um fator > 1 leva a uma diminuição do tempo de recrutamento, enquanto um fator < 1 leva a um aumento do tempo de recrutamento.

A qualidade do ajuste dos modelos selecionados é ilustrada na Figura 2. As curvas de recrutamento obtidas pelo método não-paramétrico e pelo modelo Weibull, apresentaram um bom ajuste para todos os grupos ecológicos.



**Figura 2** – Curvas para o recrutamento de diferentes grupos ecológicos em função de covariáveis ambientais, com intervalo de confiança de 95%. A - Estimadas pelo modelo Weibull comparando os grupos ecológicos; B, C e D - Modelo Weibull ajustado para espécies pioneiras, secundarias iniciais e secundarias tardias comparando a qualidade do ajuste com o método não-paramétrico intervalar

Para as espécies pioneiras, o pH, P,  $K^+$ ,  $Al^{3+}$ , precipitação nos três meses mais secos e o déficit hídrico de um ano foram selecionadas como covariáveis ambientais que afetaram o recrutamento das árvores. Baixos teores de P,  $K^+$ ,  $Al^{3+}$ , menores valores de precipitação nos meses mais secos e altos valores de pH e CWD de um ano anterior ao inventário florestal, diminuiram o tempo de recrutamento (Tabela 3).

Para as espécies secundarias iniciais foram selecionadas as covariáveis o pH,  $Ca^{2+}$  e precipitação nos três meses mais secos. À medida que aumenta os valores de pH e diminui os valores de  $Ca^{2+}$  e precipitação nos três meses mais secos, também aumentou a probabilidade de recrutamento (Tabela 3).

Para as espécies secundarias tardias, foram selecionados o pH, P e o déficit hídrico de três anos anteriores ao inventário. A capacidade de recrutamento foi afetada de forma positiva pelo aumento do pH e do déficit hídrico de três anos (menores valores de CWD) e diminuição do P (Tabela 3).

## Mortalidade

Os modelos paramétricos utilizados neste estudo pela análise de sobrevivência foram ajustados para cada grupo ecológico em função das covariáveis ambientais e seus efeitos na mortalidade. A Tabela 4 apresenta as estimativas dos parâmetros necessários para estimar as funções de sobrevivência do modelo selecionado. O modelo escolhido para todos os grupos sucessionais foi o Log-normal que apresentou menor AIC.

**Tabela 4** - Parâmetros, estimativas dos parâmetros, fator de aceleração, erros-padrão (SE) e valor-*p* do teste de Wald para o modelo Log-normal na mortalidade de cada grupo ecológico.

Parâmetros	Estimativas	Fator de aceleração	SE	Valor- <i>p</i>
Pioneiras				
$\beta_0$	2.0430		0.0303	<0.0001
$\beta_1$ ( <i>P</i> )	-0.2000	0.8188	0.0382	<0.0001
$\beta_2$ ( $K^+$ )	-0.1455	0.8646	0.0373	<0.0001
$\beta_3$ ( $Al^{3+}$ )	-0.2229	0.8002	0.0384	<0.0001
$\beta_4$ ( <i>MO</i> )	0.2558	1.2915	0.0427	<0.0001
$\beta_5$ ( <i>Precp_Dry</i> )	-0.4654	0.6279	0.0592	<0.0001
$\beta_6$ ( <i>CWD_3</i> )	0.2437	1.2760	0.041	<0.0001
$\sigma$	0.714		0.0303	<0.0001
Secundarias Iniciais				
$\beta_0$	2.4019		0.0281	<0.0001
$\beta_1$ ( $Ca^{2+}$ )	-0.0867	0.9169	0.0342	0.0113
$\beta_2$ ( $Al^{3+}$ )	-0.0901	0.9138	0.0340	<0.0001
$\beta_3$ ( <i>P - Rem</i> )	-0.0869	0.9167	0.0293	<0.0001
$\beta_4$ ( <i>CWD_3</i> )	0.0680	1.0703	0.0326	0.0371
$\sigma$	0.772		0.0427	<0.0001
Secundarias Tardias				
$\beta_0$	2.8327		0.0528	<0.0001
$\beta_1$ ( <i>pH</i> )	0.0888	1.0928	0.0536	0.0275
$\beta_2$ ( <i>P</i> )	-0.1523	0.8587	0.0455	<0.0001
$\beta_3$ ( <i>CWD_3</i> )	-0.1204	0.8866	0.0482	0.0125
$\sigma$	0.932		0.0528	<0.0001

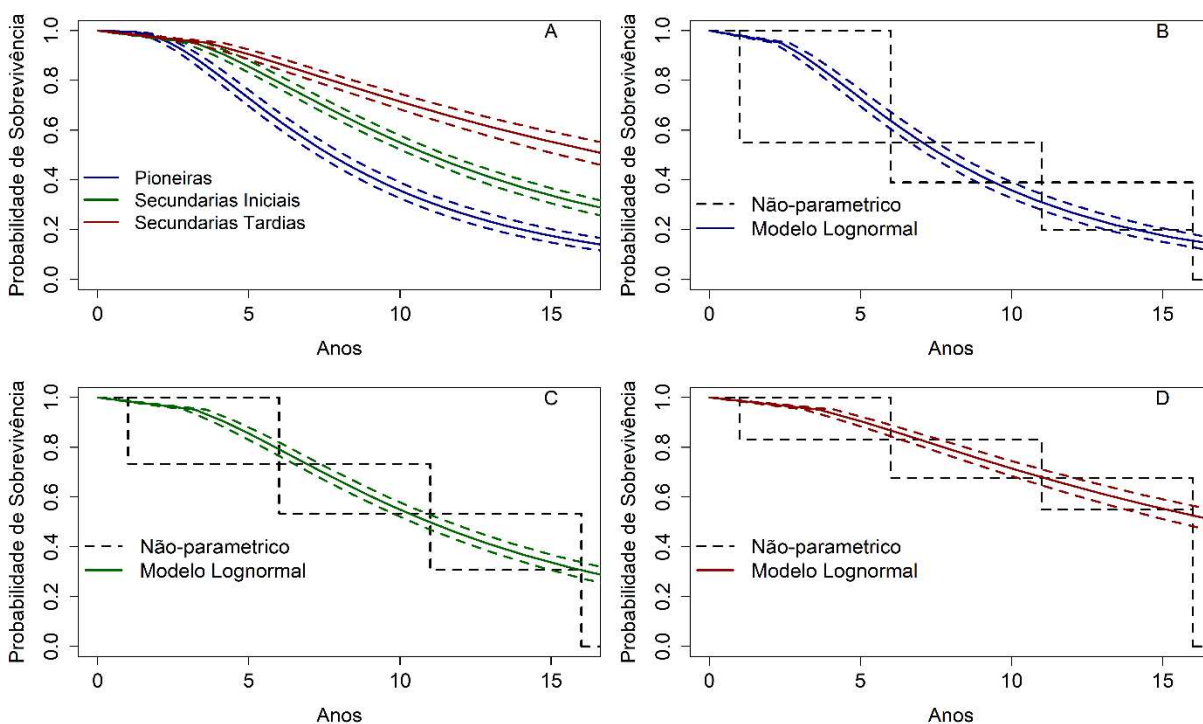
O fator de aceleração corresponde ao exp (estimativa). Um fator > 1 leva a um aumento do tempo de sobrevivência, enquanto um fator < 1 leva a uma diminuição do tempo de sobrevivência.

A qualidade do ajuste dos modelos selecionados é ilustrada na Figura 3. As curvas de sobrevivência obtidas pelo método de não-paramétrico intervalar e pelo modelo Log-normal, apresentaram um bom ajuste para todos os grupos ecológicos.

Para as espécies pioneiras, o P,  $K^+$ ,  $Al^{3+}$ , MO, precipitação nos três meses mais secos e o déficit hídrico de três anos foram selecionadas como covariáveis ambientais que afetam a mortalidade das árvores. Altos valores de P,  $K^+$ ,  $Al^{3+}$ , precipitação nos meses mais secos e eventos de seca aumentaram a probabilidade de mortalidade das arvores (Tabela 4). Enquanto que altos níveis de MO reduziram.

Para as espécies secundarias iniciais foram selecionadas as covariáveis o  $Ca^{2+}$ ,  $Al^{3+}$ , P-Rem e o déficit hídrico de três anos. As condições secas (menores valores de CWD\_3) e o aumento dos teores de  $Ca^{2+}$ ,  $Al^{3+}$ , P-Rem diminuíram a probabilidade de sobrevivência (Tabela 4).

Para as espécies secundarias tardias, foram selecionados o pH, P e o déficit hídrico de três anos. A probabilidade de mortalidade foi reduzida pelo aumento do pH e eventos de seca (Tabela 4).



**Figura 3** – Curvas para a sobrevivência de diferentes grupos ecológicos em função de covariáveis ambientais, com intervalo de confiança de 95%. A - Estimadas pelo modelo Log-normal comparando os grupos ecológicos; B, C e D - Modelo Log-normal ajustado para espécies pioneiras, secundarias iniciais e secundarias tardias mostrando a qualidade do ajuste pelo método não-paramétrico intervalar.

## DISCUSSÃO

Este estudo demonstrou o potencial da análise de sobrevivência intervalar para modelar a sobrevivência e a mortalidade de árvores em fragmentos de Mata Atlântica. É importante notar que, para todos os grupos ecológicos, os intervalos de confiança dos efeitos das covariáveis são pequenos (Figuras 2 e 3), inferindo que o ajuste do modelo foi adequado, com alta precisão das estimativas.

### Covariáveis de solo

O P, K<sup>+</sup>, Ca<sup>2+</sup> no solo tiveram influência positiva no aumento da probabilidade de recrutamento (Tabela 3) e na de mortalidade (Tabela 4). Fragmentos florestais presentes em solos de boa fertilidade apresentam altas taxas de rotatividade de povoamentos, favorecendo o recrutamento e a mortalidade de árvores (Peña e Duque, 2013, Ribeiro et al., 2019).

Neste estudo, observou-se um aumento na probabilidade de mortalidade das espécies pioneiras com a elevação dos teores de P e K<sup>+</sup> no solo (Tabela 3). Esse efeito pode estar relacionado ao fato de que concentrações mais altas desses nutrientes estimulam o rápido crescimento de espécies de baixa densidade da madeira, que possuem ciclos de vida mais curtos (Quesada et al., 2012).

A matéria orgânica do solo teve um efeito positivo no aumento do tempo de sobrevivência das árvores pioneiras (Tabela 4). Considerada um bom indicador da qualidade do solo (Teixeira et al., 2020), sua maior concentração está associada ao aumento da retenção de água, da fertilidade e do controle da erosão (Adhikari e Hartemink, 2016). Essas condições favorecem altas taxas de crescimento das árvores (Torres et al., 2023) e reduzem o risco de mortalidade.

O aumento da mortalidade e do tempo de recrutamento em espécies pioneiras e secundárias iniciais, em resposta à elevação dos teores de Al<sup>3+</sup> no solo (Tabelas 3 e 4), pode estar associado à toxicidade desse elemento para as plantas. Em altas concentrações, o Al<sup>3+</sup> pode limitar o crescimento radicular de algumas espécies, tornando-as mais vulneráveis à seca (Chandra e Keshavkant, 2021). Como consequência, essas espécies podem apresentar taxas de crescimento reduzidas e maior risco de mortalidade.

Para as espécies secundárias iniciais, o aumento do teor de Ca<sup>2+</sup> no solo resultou em uma diminuição na probabilidade de sobrevivência e recrutamento (Tabelas 3 e 4). Bizuti et al. (2018) observaram que o aumento das doses de cálcio ocasionou redução no número de raízes finas em espécies secundárias iniciais. Como essas raízes são responsáveis pela absorção de nutrientes, uma diminuição em sua quantidade pode prejudicar o desenvolvimento da planta.

O aumento da probabilidade de recrutamento nos três grupos ecológicos, em resposta à elevação do pH do solo (Tabela 3), pode estar relacionado à maior absorção e incorporação de nutrientes por essas espécies. Esse processo contribui para uma maior resistência a eventos de estresse fisiológico, como a seca (Schulze et al., 2019), além de favorecer o crescimento.

### **Covariáveis climáticas**

O CWD foi a variável climática com efeito significativo na mortalidade dos três grupos ecológicos (Tabela 4). Eventos de seca três anos antes do censo (CWD\_3) ocasionaram menor probabilidade de sobrevivência nas espécies pioneiras e secundárias iniciais. Esse efeito defasado na mortalidade, em relação ao período de seca, foi observado na floresta amazônica (Aleixo et al., 2019) e em um fragmento de Mata Atlântica em Minas Gerais (Rocha et al., 2020), em que ocorreu aumento do percentual de árvores mortas dois e quatro anos após os eventos de seca, respectivamente.

O efeito da seca foi maior nas espécies pioneiras do que nas secundárias iniciais (Tabela 4). A maior morte desse grupo ecológico, pode estar relacionado as altas taxas de condutância estomática e fotossíntese presentes nessas, isso faz com que, o risco de morte aumente em eventos de seca (Lohbeck et al., 2013, Ouédraogo et al., 2013, Uriarte et al., 2016). A elevada taxa de mortalidade das pioneiras pode afetar a composição e diversidade de espécies das florestas tropicais (Aleixo et al., 2019), alterando a dinâmica sucessional (Uriarte et al., 2016).

A seca afetou a mortalidade das espécies secundárias tardias de forma negativa, esse comportamento pode ser explicado pelo fato desse grupo ecológico ser dominado por espécies resistentes à seca, o que pode compensar os efeitos adversos da baixa disponibilidade de água (Johnson et al., 2017; Esquivel-Muelbert et al., 2017).

Devido as mudanças climáticas, eventos de seca devem se tornar cada vez mais frequentes. Estes eventos terão um efeito negativo sobre os fragmentos florestais na mata atlântica, por manter a floresta nos estágios iniciais do desenvolvimento sucessional devido a taxas de mortalidade mais altas (Torres et al., 2023).

Em relação ao recrutamento, um menor déficit hídrico aumentou a probabilidade de recrutamento de espécies pioneiras (Tabela 3). A disponibilidade de água é um importante no crescimento e no desenvolvimento de mudas dessas espécies, por estar diretamente relacionada no processo de transpiração, quanto para a absorção de CO<sub>2</sub> na fotossíntese (Schulze et al., 2019).

O aumento da precipitação nos meses mais secos ocasionou em um aumento do tempo de recrutamento das espécies pioneiras e secundárias iniciais (Tabela 3). Uma possível

explicação pode estar relacionada com aumento da lixiviação de nutrientes com o aumento da precipitação (Chadwick et al., 1999). A lixiviação de cátions básicos pode resultar em baixos níveis de pH, altas concentrações de alumínio e efeitos de toxicidade que podem afetar o crescimento das plantas (Holste et al., 2011).

A relação negativa entre o recrutamento e a disponibilidade de água para as espécies secundárias tardias (Tabela 3) também foi observada por Pineda-García et al. (2016). Esses autores constataram que espécies tolerantes à seca possuem menor capacidade de transporte de água e aquisição de carbono, em comparação com aquelas que consomem mais água.

Este é o primeiro trabalho a utilizar modelos paramétricos de sobrevivência intervalar para modelar a sobrevivência de árvores em fragmentos de Mata Atlântica. A construção do modelo com covariáveis de clima e solo representam uma alternativa promissora para prever a sobrevivência de árvores em florestas tropicais. Contudo, novas pesquisas adicionando fragilidade aos modelos podem trazer novos ganhos e ajudar a compreender melhor a dinâmica florestal neste bioma.

## CONCLUSÃO

A metodologia de análise de sobrevivência intervalar foi eficiente em avaliar o efeito de covariáveis ambientais no recrutamento e na mortalidade de árvores em fragmentos de Mata Atlântica. Esta técnica se mostrou uma ferramenta eficaz para analisar como os efeitos ambientais influenciam os processos demográficos de diferentes grupos ecológicos.

A influência das covariáveis ambientais na mortalidade e no recrutamento variaram com o grupo ecológico das árvores. As espécies pioneiras sofrem maior influência do pH, P, K<sup>+</sup>, Al<sup>3+</sup>, Precip\_Dry e CWD\_1 no recrutamento e do P, K<sup>+</sup>, Al<sup>3+</sup>, MO, Precip\_Dry e CWD\_3 na mortalidade. Enquanto as espécies secundárias iniciais são influenciadas pelo pH, Ca<sup>2+</sup> e Precip\_Dry no recrutamento e Ca<sup>2+</sup>, Al<sup>3+</sup>, P-Rem e CWD\_3 na mortalidade. As secundárias tardias são influenciadas pelo pH, P e CWD\_3 no recrutamento e mortalidade

## REFERÊNCIAS

- Adhikari, K., Hartemink, A.E., 2016. Linking soils to ecosystem services — A global review. *Geoderma* 262, 101–111. <https://doi.org/10.1016/j.geoderma.2015.08.009>
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Aleixo, I., Norris, D., Hemerik, L., Barbosa, A., Prata, E., Costa, F., Poorter, L., 2019. Amazonian rainforest tree mortality driven by climate and functional traits. *Nat. Clim. Chang.* 9, 384–388. <https://doi.org/10.1038/s41558-019-0458-0>

- Anderson-Bergman, C., 2017. *icenReg: Regression Models for Interval Censored Data in R*. Journal of Statistical Software 81, 1–23. <https://doi.org/10.18637/jss.v081.i12>
- Arroyo-Rodríguez, V., Melo, F.P.L., Martínez-Ramos, M., Bongers, F., Chazdon, R.L., Meave, J.A., Norden, N., Santos, B.A., Leal, I.R., Tabarelli, M., 2017. Multiple successional pathways in human-modified tropical landscapes: new insights from forest succession, forest fragmentation and landscape ecology research. *Biological Reviews* 92, 326–340. <https://doi.org/10.1111/brv.12231>
- Bogaerts, K., Komarek, A., Lesaffre, E., 2017. *Survival Analysis with Interval-Censored Data: A Practical Approach with Examples in R, SAS, and BUGS*. Chapman and Hall/CRC, New York. <https://doi.org/10.1201/9781315116945>
- Brandl, S., Paul, C., Knoke, T., Falk, W., 2020. The influence of climate and management on survival probability for Germany's most important tree species. *Forest Ecology and Management* 458, 117652. <https://doi.org/10.1016/j.foreco.2019.117652>
- Byrnes, J.E.K., Gamfeldt, L., Isbell, F., Lefcheck, J.S., Griffin, J.N., Hector, A., Cardinale, B.J., Hooper, D.U., Dee, L.E., Emmett Duffy, J., 2014. Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods in Ecology and Evolution* 5, 111–124. <https://doi.org/10.1111/2041-210X.12143>
- Capellesso, E.S., Cequinel, A., Marques, R., Marques, M.C.M., 2020. Temporal and environmental correlates of carbon stocks in a regenerating tropical forest. *Applied Vegetation Science* 23, 353–362. <https://doi.org/10.1111/avsc.12487>
- Chadwick, O.A., Derry, L.A., Vitousek, P.M., Huebert, B.J., Hedin, L.O., 1999. Changing sources of nutrients during four million years of ecosystem development. *Nature* 397, 491–497. <https://doi.org/10.1038/17276>
- Chandra, J., Keshavkant, S., 2021. Mechanisms underlying the phytotoxicity and genotoxicity of aluminum and their alleviation strategies: A review. *Chemosphere* 278, 130384. <https://doi.org/10.1016/j.chemosphere.2021.130384>
- 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-Yrizar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Péliissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* 20, 3177–3190. <https://doi.org/10.1111/gcb.12629>
- Collett, D., 2023. *Modelling Survival Data in Medical Research*, 4th ed. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781003282525>
- Colosimo, E.A., Giolo, S.R., 2024. *Análise de sobrevivência aplicada*, 2nd ed. Editora Edgard Blucher, São Paulo, SP.
- Costa, R.L., Prevedello, J.A., de Souza, B.G., Cabral, D.C., 2017. Forest transitions in tropical landscapes: A test in the Atlantic Forest biodiversity hotspot. *Applied Geography* 82, 93–100. <https://doi.org/10.1016/j.apgeog.2017.03.006>
- Dalmaso, C.A., Marques, M.C.M., Higuchi, P., Zwiener, V.P., Marques, R., 2020. Spatial and temporal structure of diversity and demographic dynamics along a successional gradient of tropical forests in southern Brazil. *Ecology and Evolution* 10, 3164–3177. <https://doi.org/10.1002/ece3.5816>
- Delgado, R.C., Pereira, M.G., Teodoro, P.E., dos Santos, G.L., de Carvalho, D.C., Magistrali,

- I.C., Vilanova, R.S., 2018. Seasonality of gross primary production in the Atlantic Forest of Brazil. *Global Ecology and Conservation* 14, e00392. <https://doi.org/10.1016/j.gecco.2018.e00392>
- Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36, 27–46. <https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- dos Santos, A.R., Araújo, E.F., Barros, Q.S., Fernandes, M.M., de Moura Fernandes, M.R., Moreira, T.R., de Souza, K.B., da Silva, E.F., Silva, J.P.M., Santos, J.S., Billo, D., Silva, R.F., Nascimento, G.S.P., da Silva Gandine, S.M., Pinheiro, A.A., Ribeiro, W.R., Gonçalves, M.S., da Silva, S.F., Senhorelo, A.P., Heitor, F.D., Berude, L.C., de AlmeidaTelles, L.A., 2020. Fuzzy concept applied in determining potential forest fragments for deployment of a network of ecological corridors in the Brazilian Atlantic Forest. *Ecological Indicators* 115, 106423. <https://doi.org/10.1016/j.ecolind.2020.106423>
- Esquivel-Muelbert, A., Baker, T.R., Dexter, K.G., Lewis, S.L., ter Steege, H., Lopez-Gonzalez, G., Monteagudo Mendoza, A., Brienen, R., Feldpausch, T.R., Pitman, N., Alonso, A., van der Heijden, G., Peña-Claros, M., Ahuite, M., Alexiaides, M., Álvarez Dávila, E., Murakami, A.A., Arroyo, L., Aulestia, M., Balslev, H., Barroso, J., Boot, R., Cano, A., Chama Moscoso, V., Comiskey, J.A., Cornejo, F., Dallmeier, F., Daly, D.C., Dávila, N., Duivenvoorden, J.F., Duque Montoya, A.J., Erwin, T., Di Fiore, A., Fredericksen, T., Fuentes, A., García-Villacorta, R., Gonzales, T., Guevara Andino, J.E., Honorio Coronado, E.N., Huamantupa-Chuquimaco, I., Eliana Maria Jiménez, R., Killeen, T.J., Malhi, Y., Mendoza, C., Mogollón, H., Jørgensen, P.M., Montero, J.C., Mostacedo, B., Nauray, W., Neill, D., Vargas, P.N., Palacios, S., Palacios Cuenca, W., Pallqui Camacho, N.C., Peacock, J., Phillips, J.F., Pickavance, G., Quesada, C.A., Ramírez-Angulo, H., Restrepo, Z., Reynel Rodriguez, C., Paredes, M.R., Peñuela-Mora, M.C., Sierra, R., Silveira, M., Stevenson, P., Stropp, J., Terborgh, J., Tirado, M., Toledo, M., Torres-Lezama, A., Umaña, M.N., Urrego, L.E., Vasquez Martinez, R., Gamarra, L.V., Vela, C.I.A., Vilanova Torre, E., Vos, V., von Hildebrand, P., Vriesendorp, C., Wang, O., Young, K.R., Zartman, C.E., Phillips, O.L., 2017. Seasonal drought limits tree species across the Neotropics. *Ecography* 40, 618–629. <https://doi.org/10.1111/ecog.01904>
- Fahey, T.J., Heinz, A.K., Battles, J.J., Fisk, M.C., Driscoll, C.T., Blum, J.D., Johnson, C.E., 2016. Fine root biomass declined in response to restoration of soil calcium in a northern hardwood forest. *Can. J. For. Res.* 46, 738–744. <https://doi.org/10.1139/cjfr-2015-0434>
- Fien, E.K.P., Fraver, S., Teets, A., Weiskittel, A.R., Hollinger, D.Y., 2019. Drivers of individual tree growth and mortality in an uneven-aged, mixed-species conifer forest. *Forest Ecology and Management* 449, 117446. <https://doi.org/10.1016/j.foreco.2019.06.043>
- Freitas, E.C.S. de, Paiva, H.N. de, Neves, J.C.L., Marcatti, G.E., Leite, H.G., 2020. Modeling of eucalyptus productivity with artificial neural networks. *Industrial Crops and Products* 146, 112149. <https://doi.org/10.1016/j.indcrop.2020.112149>
- Gandolfi, S., Leitão Filho, H. D. F., Bezerra, C. L. F., 1995. Levantamento florístico e caráter sucessional das espécies arbustivo-arbóreas de uma floresta mesófila semidecídua no município de Guarulhos, SP. *Revista brasileira de biologia*, 55(4), 753-767.

- Gonçalves Bizuti, D., Casagrande, J., Soares, M., Sartorio, S., Brugnaro, C., Gomes César, R., 2018. The effect of calcium on the growth of native species in a tropical forest hotspot. *iForest* 11, 221–226. <https://doi.org/10.3832/ifor2074-010>
- Griess, V.C., Acevedo, R., Härtl, F., Staupendahl, K., Knoke, T., 2012. Does mixing tree species enhance stand resistance against natural hazards? A case study for spruce. *Forest Ecology and Management* 267, 284–296. <https://doi.org/10.1016/j.foreco.2011.11.035>
- 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., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances* 1, e1500052. <https://doi.org/10.1126/sciadv.1500052>
- Holste, E.K., Kobe, R.K., Vriesendorp, C.F., 2011. Seedling growth responses to soil resources in the understory of a wet tropical forest. *Ecology* 92, 1828–1838. <https://doi.org/10.1890/10-1697.1>
- Ibge, 2012. *Manual Técnico da Vegetação Brasileira* (2a. ed.), IBGE, Rio de Janeiro.
- Johnson, D.J., Condit, R., Hubbell, S.P., Comita, L.S., 2017. Abiotic niche partitioning and negative density dependence drive tree seedling survival in a tropical forest. *Proceedings of the Royal Society B: Biological Sciences* 284, 20172210. <https://doi.org/10.1098/rspb.2017.2210>
- Joly, C.A., Metzger, J.P., Tabarelli, M., 2014. Experiences from the Brazilian Atlantic Forest: ecological findings and conservation initiatives. *New Phytologist* 204, 459–473. <https://doi.org/10.1111/nph.12989>
- Kozłowski, T.T., 2002. Physiological-ecological impacts of flooding on riparian forest ecosystems. *Wetlands* 22, 550–561. [https://doi.org/10.1672/0277-5212\(2002\)022\[0550:PEIOFO\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2002)022[0550:PEIOFO]2.0.CO;2)
- Lohbeck, M., Poorter, L., Lebrija-Trejos, E., Martínez-Ramos, M., Meave, J.A., Paz, H., Pérez-García, E.A., Romero-Pérez, I.E., Tauro, A., Bongers, F., 2013. Successional changes in functional composition contrast for dry and wet tropical forest. *Ecology* 94, 1211–1216. <https://doi.org/10.1890/12-1850.1>
- Lucas Borja, M.E., Van Stan, J.T., Plaza-Álvarez, P.A., Manso, R., 2021. Simultaneous estimation of *Pinus nigra* Arn. ssp. *salzmannii* natural regeneration emergence and survival through lifetime analysis. *Forest Ecology and Management* 499, 119613. <https://doi.org/10.1016/j.foreco.2021.119613>
- Lutz, J.A., van Wagtendonk, J.W., Franklin, J.F., 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography* 37, 936–950. <https://doi.org/10.1111/j.1365-2699.2009.02268.x>
- Melo, D.H.A., Filgueiras, B.K.C., Iserhard, C.A., Iannuzzi, L., Freitas, A.V.L., Leal, I.R., 2019. Effect of habitat loss and fragmentation on fruit-feeding butterflies in the Brazilian Atlantic Forest. *Can. J. Zool.* 97, 588–596. <https://doi.org/10.1139/cjz-2018-0202>
- Neumann, M., Mues, V., Moreno, A., Hasenauer, H., Seidl, R., 2017. Climate variability drives recent tree mortality in Europe. *Global Change Biology* 23, 4788–4797. <https://doi.org/10.1111/gcb.13724>
- Neuner, S., Albrecht, A., Cullmann, D., Engels, F., Griess, V.C., Hahn, W.A., Hanewinkel, M., Härtl, F., Kölling, C., Staupendahl, K., Knoke, T., 2015. Survival of Norway spruce

- remains higher in mixed stands under a dryer and warmer climate. *Global Change Biology* 21, 935–946. <https://doi.org/10.1111/gcb.12751>
- Nothdurft, A., 2013. Spatio-temporal prediction of tree mortality based on long-term sample plots, climate change scenarios and parametric frailty modeling. *Forest Ecology and Management* 291, 43–54. <https://doi.org/10.1016/j.foreco.2012.11.028>
- Ouédraogo, D.-Y., Mortier, F., Gourlet-Fleury, S., Freycon, V., Picard, N., 2013. Slow-growing species cope best with drought: evidence from long-term measurements in a tropical semi-deciduous moist forest of Central Africa. *Journal of Ecology* 101, 1459–1470. <https://doi.org/10.1111/1365-2745.12165>
- Paul, C., Brandl, S., Friedrich, S., Falk, W., Härtl, F., Knoke, T., 2019. Climate change and mixed forests: how do altered survival probabilities impact economically desirable species proportions of Norway spruce and European beech? *Annals of Forest Science* 76, 1–15. <https://doi.org/10.1007/s13595-018-0793-8>
- Peña, M.A., Duque, A., 2013. Patterns of stocks of aboveground tree biomass, dynamics, and their determinants in secondary Andean forests. *Forest Ecology and Management* 302, 54–61. <https://doi.org/10.1016/j.foreco.2013.03.025>
- Pineda-García, F., Paz, H., Meinzer, F.C., Angeles, G., 2016. Exploiting water versus tolerating drought: water-use strategies of trees in a secondary successional tropical dry forest. *Tree Physiology* 36, 208–217. <https://doi.org/10.1093/treephys/tpv124>
- 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., Nyctch, C.J., de Oliveira, A.A., Pérez-García, E.A., Prado-Junior, J., Rodríguez-Velázquez, J., Ruschel, A.R., Salgado-Negret, B., Schiavini, I., Swenson, N.G., Tenorio, E.A., Thompson, J., Toledo, M., Uriarte, M., Hout, P. van der, Zimmerman, J.K., Peña-Claros, M., 2017. Biodiversity and climate determine the functioning of Neotropical forests. *Global Ecology and Biogeography* 26, 1423–1434. <https://doi.org/10.1111/geb.12668>
- Quesada, C.A., Phillips, O.L., Schwarz, M., Czimczik, C.I., Baker, T.R., Patiño, S., Fyllas, N.M., Hodnett, M.G., Herrera, R., Almeida, S., Alvarez Dávila, E., Arneeth, A., Arroyo, L., Chao, K.J., Dezzeo, N., Erwin, T., di Fiore, A., Higuchi, N., Honorio Coronado, E., Jimenez, E.M., Killeen, T., Lezama, A.T., Lloyd, G., López-González, G., Luizão, F.J., Malhi, Y., Monteagudo, A., Neill, D.A., Núñez Vargas, P., Paiva, R., Peacock, J., Peñuela, M.C., Peña Cruz, A., Pitman, N., Priante Filho, N., Prieto, A., Ramírez, H., Rudas, A., Salomão, R., Santos, A.J.B., Schmerler, J., Silva, N., Silveira, M., Vásquez, R., Vieira, I., Terborgh, J., Lloyd, J., 2012. Basin-wide variations in Amazon forest structure and function are mediated by both soils and climate. *Biogeosciences* 9, 2203–2246. <https://doi.org/10.5194/bg-9-2203-2012>
- R Development Core Team, 2024. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria. <https://www.r-project.org/>
- Redmond, M.D., 2022. CWD and AET function. <https://doi.org/10.5281/zenodo.6416352>
- Ribeiro, J.C.T., Nunes-Freitas, A.F., Fidalgo, E.C.C., Uzêda, M.C., 2019. Forest fragmentation and impacts of intensive agriculture: Responses from different tree functional groups. *PLOS ONE* 14, e0212725. <https://doi.org/10.1371/journal.pone.0212725>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Schettini, B.L.S., Jacovine, L.A.G., Leite, H.G., Gelcer, E.M., Reis, L.P., Neves, K.M., Comini, I.B., Silva, L.F. da, 2020. Drought effects on carbon dynamics of trees in a secondary Atlantic Forest. *Forest*

- Ecology and Management 465, 118097. <https://doi.org/10.1016/j.foreco.2020.118097>
- Schulze, E.-D., Beck, E., Buchmann, N., Clemens, S., Müller-Hohenstein, K., Scherer-Lorenzen, M., 2019. Plant Ecology. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-662-56233-8>
- Solórzano, A., Brasil, L.S.C. de A., de Oliveira, R.R., 2021. The Atlantic Forest Ecological History: From Pre-colonial Times to the Anthropocene, in: Marques, M.C.M., Grelle, C.E.V. (Eds.), The Atlantic Forest: History, Biodiversity, Threats and Opportunities of the Mega-Diverse Forest. Springer International Publishing, Cham, pp. 25–44. [https://doi.org/10.1007/978-3-030-55322-7\\_2](https://doi.org/10.1007/978-3-030-55322-7_2)
- SOS Mata Atlântica, 2024. Atlas da Mata Atlântica. <https://www.sosma.org.br/iniciativas/atlas-da-mata-atlantica>
- Souza, C.R., Maia, V.A., Aguiar-Campos, N. de, Santos, A.B.M., Rodrigues, A.F., Farrapo, C.L., Gianasi, F.M., Paula, G.G.P. de, Fagundes, N.C.A., Silva, W.B., Santos, R.M., 2021. Long-term ecological trends of small secondary forests of the atlantic forest hotspot: A 30-year study case. Forest Ecology and Management 489, 119043. <https://doi.org/10.1016/j.foreco.2021.119043>
- Staupendahl, K., Möhring, B., 2011. Integrating natural risks into silvicultural decision models: A survival function approach. Forest Policy and Economics 13, 496–502. <https://doi.org/10.1016/j.forpol.2011.05.007>
- Tabarelli, M., Aguiar, A.V., Ribeiro, M.C., Metzger, J.P., Peres, C.A., 2010. Prospects for biodiversity conservation in the Atlantic Forest: Lessons from aging human-modified landscapes. Biological Conservation, Conserving complexity: Global change and community-scale interactions 143, 2328–2340. <https://doi.org/10.1016/j.biocon.2010.02.005>
- Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Müller, M.S., Rödig, E., Wiegand, T., Huth, A., 2018. Global patterns of tropical forest fragmentation. Nature 554, 519–522. <https://doi.org/10.1038/nature25508>
- Teixeira, H.M., Cardoso, I.M., Bianchi, F.J.J.A., da Cruz Silva, A., Jamme, D., Peña-Claros, M., 2020. Linking vegetation and soil functions during secondary forest succession in the Atlantic forest. Forest Ecology and Management 457, 117696. <https://doi.org/10.1016/j.foreco.2019.117696>
- Teixeira, P.C., Donagemma, G.K., Fontana, A., Teixeira, W.G. (Eds.), 2017. Manual de métodos de análise de solo, 3ª. ed. Embrapa, Brasília, Brazil.
- Thapa, R., Burkhart, H.E., Li, J., Hong, Y., 2016. Modeling Clustered Survival Times of Loblolly Pine with Time-dependent Covariates and Shared Frailties. JABES 21, 92–110. <https://doi.org/10.1007/s13253-015-0217-2>
- Therneau, T., 2024. A Package for Survival Analysis in R. R package version 3.8-3. <https://CRAN.R-project.org/package=survival>
- Thiele, J.C., Nuske, R.S., Ahrends, B., Panferov, O., Albert, M., Staupendahl, K., Junghans, U., Jansen, M., Saborowski, J., 2017. Climate change impact assessment—A simulation experiment with Norway spruce for a forest district in Central Europe. Ecological Modelling 346, 30–47. <https://doi.org/10.1016/j.ecolmodel.2016.11.013>
- Torres, C.M.M.E., Medina-Vega, J.A., Rocha, S.J.S.S. da, Costa, W. da S., Soares, C.P.B., Souza, A.L. de, Jacovine, L.A.G., Lana, J.M. de, Peña-Claros, M., 2023. Drivers of tree demographic processes in forest fragments of the Brazilian Atlantic forest. Forest

- Ecology and Management 534, 120893. <https://doi.org/10.1016/j.foreco.2023.120893>
- Uriarte, M., Schwartz, N., Powers, J.S., Marín-Spiotta, E., Liao, W., Werden, L.K., 2016. Impacts of climate variability on tree demography in second growth tropical forests: the importance of regional context for predicting successional trajectories. *Biotropica* 48, 780–797. <https://doi.org/10.1111/btp.12380>
- van der Sande, M.T., Arets, E.J.M.M., Peña-Claros, M., Hoosbeek, M.R., Cáceres-Siani, Y., van der Hout, P., Poorter, L., 2018. Soil fertility and species traits, but not diversity, drive productivity and biomass stocks in a Guyanese tropical rainforest. *Functional Ecology* 32, 461–474. <https://doi.org/10.1111/1365-2435.12968>

## MATERIAL SUPLEMENTAR

**Tabela 1 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para o recrutamento de espécies pioneiras.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	0.7074	0.0295	-1.4784	0.1402	-1.1699	0.0895	-1.0979	0.0855
$\beta_1$	0.122	0.0537	0.2649	0.1481	0.2803	0.1238	0.3005	0.1187
$\beta_2$	-0.3091	0.0441	-0.6601	0.1216	-0.3589	0.1021	-0.337	0.1018
$\beta_3$	-0.3275	0.0456	-0.6887	0.1548	-0.4597	0.1072	-0.4831	0.1132
$\beta_4$	-0.2628	0.0482	-0.6498	0.1351	-0.4408	0.1142	-0.4774	0.115
$\beta_5$	-0.8878	0.0673	-1.5157	0.1883	-1.3595	0.1564	-1.3843	0.1555
$\beta_6$	0.3967	0.0474	0.5475	0.1331	0.4958	0.1151	0.4693	0.1161
$\sigma$	1.000	---	3.1	0.0434	2.08	0.0333	1.16	0.0368
Log L	-1753.3	---	-1297.8	---	-1340.1	---	-1318.5	---
AIC	3522.58	---	<b>2613.54</b>	---	2698.18	---	2655.06	---

$$\text{Modelo: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 pH + \hat{\beta}_2 P + \hat{\beta}_3 K^+ + \hat{\beta}_4 Al^{3+} + \hat{\beta}_5 Precip\_Dry + \hat{\beta}_6 CWD\_1$$

**Tabela 2 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para o recrutamento de espécies secundarias iniciais.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	1.0398	0.0243	-0.6485	0.0697	-0.6485	0.0697	-0.5849	0.0647
$\beta_1$	0.3063	0.0364	0.5365	0.0864	0.5357	0.0864	0.5783	0.0856
$\beta_2$	-0.143	0.0319	-0.1891	0.0768	-0.1891	0.0768	-0.188	0.0768
$\beta_3$	0.222	0.0296	-0.3643	0.0946	0.4162	0.0682	0.4469	0.0674
$\sigma$	1.00	---	2.10	0.0284	2.46	0.0284	1.21	0.0296
Log L	-2808.1	---	-2256.5	---	-2265	---	-2300.4	---
AIC	5628.29	---	<b>4527.02</b>	---	4544.04	---	4614.81	---

$$\text{Modelo: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 pH + \hat{\beta}_2 Ca^{2+} + \hat{\beta}_3 Precip\_Dry$$

**Tabela 3 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para o recrutamento de espécies secundárias tardias.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	1.0663	0.0252	-0.7207	0.0744	-0.8207	0.0744	-0.6558	0.0697
$\beta_1$	0.0834	0.037	0.1559	0.0904	0.1659	0.0904	0.1729	0.0912
$\beta_2$	-0.1138	0.0321	-0.1948	0.0804	-0.1848	0.0804	-0.2037	0.0814
$\beta_3$	-0.0799	0.038	-0.2333	0.0936	-0.2633	0.0936	-0.2677	0.0938
$\sigma$	1.0000	---	2.18	0.0288	2.78	0.0288	1.25	0.0304
Log L	-2661.6	---	-2094.2	---	-2109.1	---	-2137.5	---
AIC	5333.12	---	<b>4200.40</b>	---	4230.11	---	4287.07	---

$$\text{Model: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 pH + \hat{\beta}_2 P + \hat{\beta}_3 CWD\_3$$

**Tabela 4 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para mortalidade de espécies pioneiras.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	2.4978	0.0416	2.3865	0.0245	2.0430	0.0303	2.0512	0.0302
$\beta_1$	-0.1887	0.0521	-0.1797	0.034	-0.2000	0.0382	-0.2074	0.0385
$\beta_2$	-0.1256	0.0495	-0.1474	0.0352	-0.1455	0.0373	-0.1519	0.038
$\beta_3$	-0.1728	0.0535	-0.1741	0.0322	-0.2229	0.0384	-0.2288	0.0381
$\beta_4$	0.1984	0.0614	0.1931	0.0373	0.2558	0.0427	0.2680	0.0425
$\beta_5$	-0.4718	0.0818	-0.3881	0.051	-0.4654	0.0592	-0.4729	0.0593
$\beta_6$	0.2806	0.0582	0.2119	0.0344	0.2437	0.041	0.2485	0.0405
$\sigma$	1.0000	---	0.5530	0.0366	0.714	0.0303	0.403	0.0412
Log L	-811.2	---	-684.6	---	-644.7	---	-656.6	---
AIC	1636.46	---	1385.23	---	<b>1305.40</b>	---	1329.10	---

$$\text{Modelo: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 P + \hat{\beta}_2 K^+ + \hat{\beta}_3 Al^{3+} + \hat{\beta}_4 MO + \hat{\beta}_5 Precp\_Dry + \hat{\beta}_6 CWD\_3$$

**Tabela 5 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para mortalidade de espécies secundarias iniciais.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	29.3080	0.0417	2.6664	0.0230	2.4019	0.0281	2.4039	0.0262
$\beta_1$	-0.1774	0.0939	-0.1241	0.0504	-0.0867	0.0342	-0.1420	0.0609
$\beta_2$	-0.1759	0.0610	-0.0858	0.0332	-0.0901	0.0340	-0.1391	0.0461
$\beta_3$	0.0610	0.0556	0.0352	0.0298	-0.0869	0.0293	0.0444	0.0356
$\beta_4$	0.1209	0.0669	0.0744	0.0359	0.0680	0.0326	0.0995	0.0418
$\sigma$	1.000	---	0.526	0.0252	0.772	0.0427	0.0413	0.0303
Log L	-1036.5	---	-896.3	---	-868.5	---	-878.2	---
AIC	2086.99	---	1808.67	---	<b>1752.97</b>	---	1772.49	---

$$\text{Modelo: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 Ca^{2+} + \hat{\beta}_2 Al^{3+} + \hat{\beta}_3 P\_Rem + \hat{\beta}_4 CWD\_3$$

**Tabela 6 - Estimativas dos parâmetros e erros padrão (SE) dos modelos, critério de informação de Akaike (AIC) para mortalidade de espécies secundarias tardias.**

Parâmetros	Exponencial		Weibull		Log-normal		Log-logístico	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
$\beta_0$	3.5196	0.0618	3.0048	0.0445	2.8327	0.0528	2.7984	0.0452
$\beta_1$	0.1116	0.075	0.0801	0.0422	0.0888	0.0536	0.0837	0.048
$\beta_2$	-0.1966	0.0586	-0.1189	0.0334	-0.1523	0.0455	-0.1473	0.042
$\beta_3$	-0.1684	0.0688	-0.0857	0.0396	-0.1204	0.0482	-0.1045	0.044
$\sigma$	1.0000	---	0.562	---	0.932	---	0.496	---
Log L	-620.6	---	-567	---	-553.3	---	-561.5	---
AIC	1249.18	---	1143.91	---	<b>1116.61</b>	---	1132.94	---

$$\text{Modelo: } \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 pH + \hat{\beta}_2 P + \hat{\beta}_3 CWD\_3$$

## **CAPITULO 2 - ANÁLISE DE SOBREVIVÊNCIA COM DADOS GRUPADOS: MORTALIDADE DE ÁRVORES EM UM FRAGMENTO DE MATA ATLÂNTICA NO BRASIL**

### **RESUMO**

Modelos para prever a mortalidade de árvores podem ajudar a entender a dinâmica da vegetação e a prever efeitos das mudanças climáticas em florestas nativas. No entanto, os padrões de mortalidade, são altamente covariáveis e difíceis de descrever. Apesar de inúmeras investigações visando desenvolver modelos de sobrevivência para arvores, ainda há lacunas importantes que precisam ser preenchidas. O objetivo deste trabalho foi avaliar a mortalidade individual de árvores de um fragmento de Mata Atlântica, por meio da técnica de análise de sobrevivência com dados discretos. Foram utilizados dados de dez parcelas permanentes de um fragmento de Mata Atlântica em Minas Gerais, Brasil. Como covariáveis explicativas utilizaram-se O DAP (diâmetro a altura do peito), área basal, distancia da borda, as características químicas do solo (pH do solo, cátions trocáveis, teor de P e matéria orgânica) e clima (temperatura média, precipitação < 100 mm, precipitação nos três meses mais secos, precipitação média anual total e déficit hídrico - CWD). Foram ajustados modelos de riscos proporcionais de Cox e o modelo de chances proporcionais para estudar a influência destas covariáveis na mortalidade das arvores. A escolha do melhor modelo foi baseada no AIC, AUC e capacidade preditiva. O modelo de taxas de falhas proporcionais de Cox e o modelo logístico de chances proporcionais apresentaram desempenho bem semelhante, na sobrevivência e mortalidade dos fustes. Os valores de AUC superiores a 0,97 e de acurácia balanceada acima de 0,81, demonstram eficiência dos modelos na classificação individual das arvores. No modelo considerando todos os grupos ecológicos, o melhor ajuste foi obtido utilizando o pH, Matéria orgânica do solo, temperatura média, precipitação média e déficit hídrico de dois anos como covariáveis preditoras. As mesmas covariáveis climáticas também foram selecionadas como preditoras para os modelos ajustados para cada grupo ecológico. Além das covariáveis de clima, a matéria orgânica do solo foi selecionada como variável preditora para as espécies pioneiras, a distância da borda,  $K^+$  e  $Ca^{2+}$  para secundarias iniciais e o  $Ca^{2+}$  para as secundarias tardias. A metodologia de análise de sobrevivência com dados discretos foi eficiente em avaliar a mortalidade de árvores na Mata Atlântica. Esta técnica pode ser uma ferramenta útil na tomada de decisões dos gestores florestais.

**Palavras-chave:** Dados censurados; Floresta tropical; Regressão discreta.

## INTRODUÇÃO

Atualmente, a Mata Atlântica é considerada um dos maiores *hotspots*, sendo um bioma em que a maior parte de suas espécies são endêmicas (Laurance, 2009). É uma floresta diversa e heterogênea composta por múltiplos fragmentos florestais (Laurance, 2009). Segundo o Atlas dos Remanescentes Florestais da Mata Atlântica (2024), restam apenas 12,4% de sua vegetação nativa, considerando apenas fragmentos acima de 3 ha. Devido ao intenso desmatamento ao longo dos anos, ela se tornou um bioma extremamente ameaçado e um dos prioritários mundialmente quanto à sua restauração, visando a conservação da biodiversidade frente às mudanças climáticas (SOS Mata Atlântica, 2024).

Estudos sobre a dinâmica das florestas tropicais são importantes para entender a evolução do ecossistema florestal após perturbações ambientais ou antrópicas. Com isso, o desenvolvimento de modelos individuais, em nível de árvore, para estimar a sobrevivência e a mortalidade são necessários para o prognóstico correto do número de árvores, área basal, distribuição de diâmetros e produção (Reis et al., 2018). Estas informações, são fundamentais para estratégias de conservação, manejo e preservação das florestas tropicais (Meir et al., 2015; Rocha et al., 2018; Vieira et al., 2018).

Uma das dificuldades em modelar a mortalidade em uma floresta é que este evento pode ser causado por vários fatores, podendo ser de forma regular causada pelo envelhecimento, supressão e competição por recursos (Reis et al., 2018), ou devido a eventos climáticos extremos, como incêndios, tempestades ou eventos de seca (Brandl et al., 2020, Bretfeld et al., 2018; Hartmann et al., 2015; Hendrik e Maxime, 2017; Manso et al., 2015). Atualmente, tais eventos extremos estão se acentuando e tornando-se cada vez mais frequentes (Meir et al., 2015). Intensificadas pela ação humana, as mudanças climáticas tem provocado o aumento das temperaturas médias globais, mudanças na quantidade, intensidade e sazonalidade das precipitações anuais nas regiões tropicais, e ocorrência de períodos de seca prolongados (Flores et al., 2024). Em consequência, estes fatores podem levar ao estresse hídrico da planta, tornando as mais susceptíveis ao ataque de pragas, doenças e a morte.

Nas florestas tropicais, a modelagem da mortalidade em nível de árvore individual, geralmente tem sido realizada por meio de regressão logística (Figueredo et al., 2020; Rocha et al., 2020) ou redes neurais artificiais (Castro et al., 2015; Rocha et al., 2018; Reis et al., 2018). No entanto, estes modelos tem apresentado baixa precisão, o que pode estar ligada ao fato de a mortalidade em florestas tropicais ser um evento raro (King et al., 2000, Reis et al., 2018). Outro ponto comum nessas abordagens é que elas não levam em consideração a censura, que corresponde a uma observação parcial da resposta de interesse.

Por meio da técnica da análise de sobrevivência, é possível estimar a probabilidade de sobrevivência e da mortalidade, assim como avaliar a influência de covariáveis no estudo (Brandl et al., 2020). Nessa análise, a variável resposta é o tempo de sobrevivência ou tempo até o evento de interesse, que geralmente é definido como sendo a morte do indivíduo em estudo (Colosimo e Giolo, 2024). Essa metodologia tem como principal característica, a inclusão dos chamados dados censurados, os quais são conhecidos como observações parciais ou censura, em que o objeto em estudo não atingiu o evento de interesse (Colosimo e Giolo, 2024). A inclusão da censura torna o modelo mais flexível em relação aos modelos de regressão logística (Brandl et al., 2020).

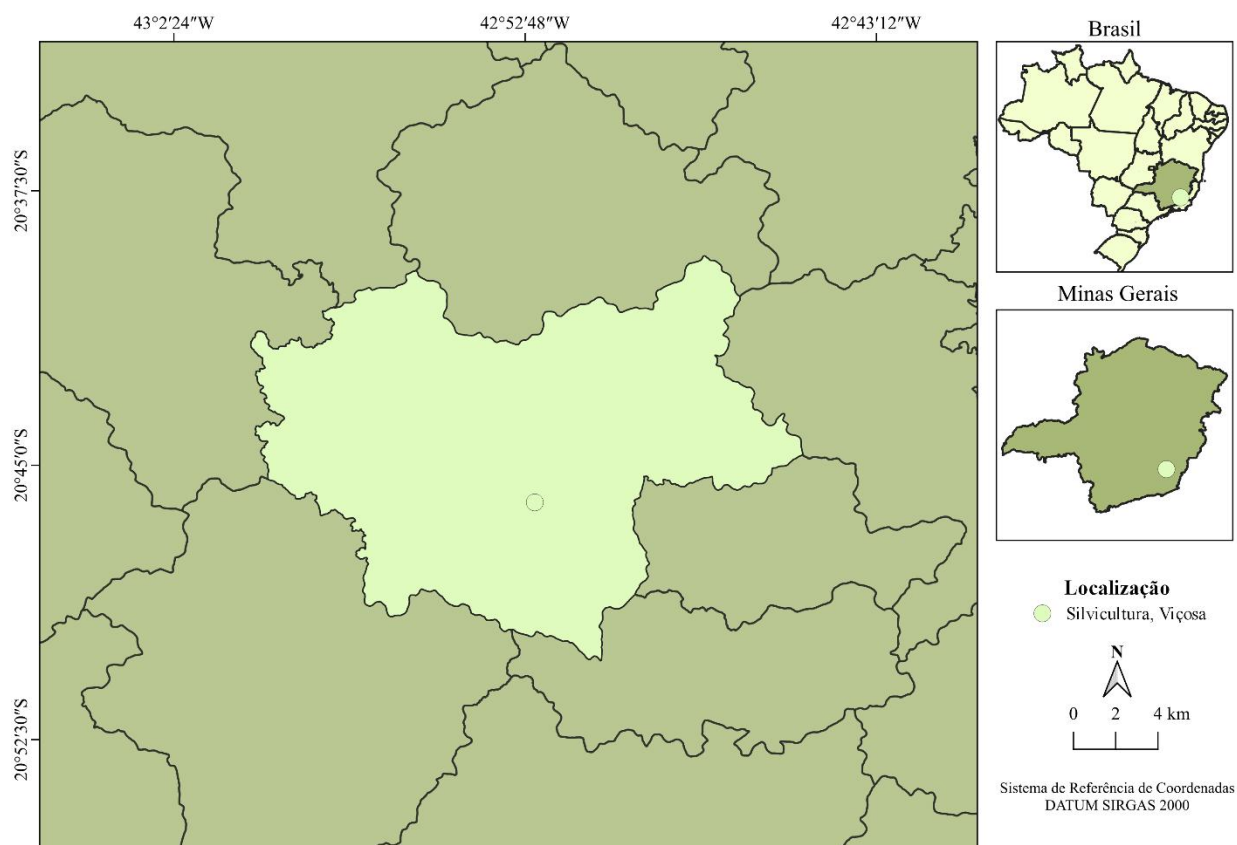
Ao utilizar observações parciais nos modelos, tem-se ganhos de informações relevantes sobre o tempo do evento, o que pode fornecer conclusões menos viesadas (Colosimo e Giolo, 2024). Sendo assim, o uso de métodos estatísticos que possibilitem incorporar na análise a informação contida tanto nas observações completas quanto nas censuras, se torna interessante.

Neste sentido, nota-se a importância do estudo, análise e compreensão dos fatores relacionados à mortalidade de árvores neste bioma, de modo a realizar o melhor manejo e adoção de práticas visando aumentar a sua conservação. Com isso, o objetivo deste trabalho foi avaliar a mortalidade individual de árvores de um fragmento de Mata Atlântica, por meio da técnica de análise de sobrevivência com dados discretos, a fim de oferecer subsídios para decisões no manejo florestal.

## **MATERIAL E MÉTODOS**

### **Local de estudo e características das parcelas**

O estudo foi conduzido na Mata da Silvicultura, uma Mata Atlântica secundária, com 17 ha, em Viçosa, Minas Gerais, Brasil (20°45'S e 42°51'W) (Figura 1). Essa área foi adquirida pela Universidade Federal de Viçosa em 1936 e tem sido protegida da exploração madeireira durante todo esse período.



**Figura 1** – Localização do fragmento de Mata Atlântica no estado de Minas Gerais, Brasil.

### Coleta dos dados

Dentro da área de estudo, foram estabelecidas de forma aleatória, dez parcelas permanentes, cada uma com 0,1 ha (20 m x 50 m). Os monitoramentos foram realizados nos anos de 1994, 1997, 2000, 2004, 2008, 2010, 2013 e 2016. Nas dez parcelas, todas as árvores com DAP (diâmetro à altura do peito; 1,3 m) maior ou igual a 5,0 cm tiveram seu CAP (circunferência na altura do peito; 1,3 m) e altura total (Ht) mensurados e foram identificados botanicamente. As espécies foram classificadas em grupos ecológicos seguindo a divisão proposta por Gandolfi et al. (1995), para florestas secundárias brasileiras em: pioneiras (P); secundárias iniciais (Si) e secundárias tardias (St) (Tabela 1). A mortalidade foi avaliada considerando cada fuste como um indivíduo. A área basal da parcela e distância até a borda foram utilizadas como covariáveis antropogênicas.

**Tabela 1** - Composição Florística em um fragmento de Mata Atlântica no estado de Minas Gerais, Brasil.

	Ano da mensuração							
	1994	1997	2000	2004	2008	2010	2013	2016
Famílias Botânicas	50	49	49	48	47	47	45	46
Gêneros Botânicos	104	103	107	107	111	111	107	113
Espécies	144	143	146	144	158	156	149	168
Pioneiras	18	16	16	14	15	15	15	17
Secundarias Iniciais	75	76	76	75	77	75	74	82
Secundarias Tardias	39	40	41	41	44	45	41	47

### Covariáveis climáticas

De uma estação meteorológica próxima ao fragmento, foram obtidos em cada monitoramento a precipitação anual, o número de meses com precipitação < 100 mm, a precipitação nos três meses mais secos e a temperatura média anual da estação climatológica mais próxima. A partir destes dados foi estimado o déficit hídrico climático (CWD), que é a diferença entre a Evapotranspiração Potencial (ETP) e a Evapotranspiração Real (ETR), de acordo a metodologia proposta por Lutz et al. (2010), utilizando as funções CWD e AET (Redmond, 2022) do software R versão 4.4.2. (R Development Core Team, 2024). O CWD reflete as condições de seca com mais precisão do que a precipitação anual total (Chave et al., 2014). O CWD quanto mais negativo, indica condições de alto estresse hídrico, e valores próximos a 0 (zero) indicam nenhum estresse (Poorter et al., 2017). Foi calculada a precipitação média anual total e CWD para quatro, três, dois e um ano antes do inventário de cada parcela. Como o CWD é por definição negativo, para facilitar sua interpretação foi multiplicado por -1. Assim, locais com valores maiores valores de CWD são locais com alto estresse hídrico.

### Covariáveis de solo

Foram coletadas aleatoriamente 20 a 30 amostras de solo para obter uma amostra composta para duas profundidades (0–20 cm e 20–40 cm) em cada parcela. pH do solo em H<sub>2</sub>O, cátions trocáveis (Ca<sup>2+</sup>, Mg<sup>2+</sup> e Al<sup>3+</sup>), acidez total (H<sup>+</sup> + Al<sup>3+</sup>), capacidade de troca catiônica (CTC), saturação por bases (V), fósforo disponível (P), P remanescente em solução (P-res) e matéria orgânica do solo (MOS). As análises de solo foram realizadas no laboratório de Fertilidade do Solo da Universidade Federal de Viçosa (UFV), Brasil.

### Análise Estatística

As covariáveis quantitativas foram padronizadas para o *score*  $Z$ , para melhorar a convergência do modelo.

$$Z_i = \frac{x_i - \bar{x}}{\sigma}$$

onde,  $Z_i$  é valor padronizado da observação  $i$ ;  $x_i$  é o valor da observação  $i$ ;  $\bar{x}$  é a média dos valores observados;  $\sigma$  é o desvio padrão.

Após a padronização das covariáveis o banco de dados foi separado em dois grupos: um para ajuste do modelo (70%) e outro para validação (30%).

### Modelo de sobrevivência discreto

A abordagem em nível de árvore individual foi usada nesse estudo, adotando o método de análise de sobrevivência para dados grupados ou discreto. Os dados foram grupados em 7 intervalos de tempo  $I_j = [t_{j-1}, t_j)$ ,  $j = 1, \dots, 7$  onde  $0 = t_0 < \dots < t_k$  e  $t_{k+1} = \infty$ . Além disso, foi considerado  $R_j$  o risco definido no tempo  $t_{j-1}$ , e  $\delta_{ij} = 1$  se o evento (mortalidade) do  $i$ -ésimo fuste ocorreu no intervalo  $I_j$  e  $\delta_{ij} = 0$ , caso contrário. A probabilidade de mortalidade do  $i$ -ésimo fuste no intervalo  $I_j$ , considerando as covariáveis analisadas, condicionado ao fato de que esse fuste não morreu em  $t_{j-1}$  é representada por  $p_j(\mathbf{x}_i)$ . A função de verossimilhança para tal modelo é:

$$L = \prod_{j=1}^7 \prod_{i \in R_i} \{p_i(\mathbf{x}_i)\}^{\delta_{ij}} \{1 - p_i(\mathbf{x}_i)\}^{(1-\delta_{ij})}$$

Essa probabilidade foi estimada por meio do modelo de taxas de falhas proporcionais de Cox e pelo modelo logístico de chances proporcionais.

No modelo de taxas de falhas proporcionais de Cox, proposto por Prentice e Gloeckler (1978), a probabilidade  $p_j(\mathbf{x}_i)$  é dada por:

$$p_j(\mathbf{x}_i) = 1 - \gamma_j^{\exp(\boldsymbol{\beta}'\mathbf{x}_i)}$$

em que  $\gamma_j = \frac{S_0(t_j)}{S_0(t_{j-1})}$ ,  $j = 1, \dots, 7$ ;  $S_0(t)$  é a função de sobrevivência básica.

Já no modelo logístico de chances proporcionais, proposto por Hosmer e Lemeshow (2000), a  $p_j(\mathbf{x}_i)$  assume a forma:

$$p_j(\mathbf{x}_i) = 1 - \frac{1}{\{1 + \gamma_j \exp(\boldsymbol{\beta}'\mathbf{x}_i)\}}$$

em que  $\gamma_j = \frac{p_j(0)}{1 - p_j(0)}$ ,  $j = 1, \dots, 7$ .

Esses modelos foram linearizados, utilizando transformações que permitiram utilizar métodos usuais para a modelagem de resposta binária.

No modelo de riscos proporcionais de Cox, utilizou-se a função de ligação complemento log-log:

$$\ln [-\ln (1 - p_j(\mathbf{x}_i))] = \gamma_j^* + \boldsymbol{\beta}' \mathbf{x}_i = \eta_{ij}$$

em que,  $\gamma_j^* = \ln(\ln(\gamma_j))$  é o efeito do  $j$ -ésimo intervalo de tempo,  $\eta_{ij}$  é o preditor linear.

No modelo logístico de chances proporcionais, esse resultado foi alcançado por meio da função de ligação logito:

$$\ln \left( \frac{p_j(\mathbf{x}_i)}{1 - p_j(\mathbf{x}_i)} \right) = \gamma_j^* + \boldsymbol{\beta}' \mathbf{x}_i = \eta_{ij}$$

em que,  $\gamma_j^* = \ln(\gamma_j)$  é o efeito do  $j$ -ésimo intervalo de tempo,  $\eta_{ij}$  é o preditor linear.

### Seleção de covariáveis

O primeiro passo para a seleção de covariáveis foi realizar uma análise de correlação de Spearman para verificar quais covariáveis seriam utilizadas nos modelos. Nas covariáveis com alta correlação ( $\pm 0.8$ ), somente uma delas foi mantida, baseado em sua consistência biológica. Para as covariáveis de solo, foram utilizados somente os dados da camada 0-20 cm, pois estavam correlacionados com os dados da camada de 20-40 cm. Foram excluídos os dados de acidez total, capacidade de troca catiônica e saturação por bases, por apresentar alta correlação com os cátions trocáveis.

As covariáveis climáticas números de meses com precipitação  $< 100$  mm, precipitação média anual total para um, dois, três e quatro anos apresentaram alta correlação entre si, assim como o CWD para o intervalo de anos considerados. Para evitar multicolinearidade (Dormann et al., 2013) apenas uma variável de cada grupo foi testada por vez.

O segundo passo foi utilizar o método derivado da estratégia de seleção de covariáveis proposta por Collett (2023) e adaptada por Colosimo e Giolo (2024), no qual as informações do pesquisador podem ser incluídas no processo de tomada de decisão, mesmo quando não significativa. A recomendação de incluir ou não um termo no modelo deve ser seguido de um nível de significância individual próximo de 0,10.

Um nível de significância de 5% foi usado para determinar a significância estatística das covariáveis selecionadas no modelo final de regressão discreta, o teste utilizado foi o de Wald.

### Seleção e avaliação dos modelos

A escolha do melhor modelo ajustado foi baseada no Critério de Informação de Akaike (AIC) (Akaike, 1974). Essa medida usa a estimativa do logaritmo da verossimilhança ( $LL$ ), adicionando um termo penalizador associado ao número de parâmetros do modelo ( $p$ ). É bem conhecido que adicionando covariáveis, pode-se melhorar o ajuste dos modelos. Assim, o AIC tenta equilibrar a qualidade do ajuste *versus* a inclusão de covariáveis no modelo. Valores menores indicam melhor ajuste do modelo. O AIC foi calculado da seguinte forma:

$$AIC = -2(LL) + 2p$$

### Avaliação de desempenho do modelo

A qualidade de ajuste dos modelos foi avaliada por meio do teste de Hosmer e Lemeshow (Hosmer Junior et al., 2013), utilizando o pacote *generalhoslem* (Jay, 2019). Este teste corresponde a um teste do  $\chi^2$  com  $g - 2$  graus de liberdade e consiste em dividir o número de observações em dez grupos ( $g = 10$ ) e depois comparar as frequências preditas com as observadas. O objetivo deste teste é verificar se existem diferenças significativas entre as classificações realizadas pelo modelo e a realidade observada. Para este teste o interessante é não rejeitar a hipótese  $H_0$ , ou seja, de que não existe diferenças significativas entre a classificação realizada pelo modelo e a observada.

Para avaliar a capacidade preditiva dos modelos selecionados, foram utilizadas as seguintes métricas de desempenho: Sensibilidade ( $S_n$ ), Especificidade ( $E_p$ ), acurácia balanceada ( $AB$ ) e Area abaixo da curva ( $AUC$ ). As métricas foram obtidas utilizando o pacote *ROCR* (Sing et al., 2020).

Um classificador eficiente é aquele capaz realizar previsões corretas a partir de dados conhecidos da realidade do objeto de estudo, podendo identificar casos como verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN) (Vujović, 2021).

Sensibilidade ( $S_n$ ): é a porcentagem de todos os VP que foram preditos corretamente, dada por:

$$S_n = \frac{VP}{(VP + FN)}$$

Especificidade ( $E_p$ ): é a porcentagem de todos os VN que foram preditos corretamente, dada por:

$$Ep = \frac{VN}{(VN + FP)}$$

Acurácia balanceada (AB): avalia a acurácia média em ambas as classes positivas e negativas, considerando o desequilíbrio das classes. Essa é frequentemente utilizada para conjuntos de dados desbalanceados, pois combina sensibilidade e especificidade para fornecer uma compreensão mais detalhada do desempenho dos modelos, dada por

$$AB = \frac{1}{2} \left( \frac{VP}{VP + FN} + \frac{VN}{VN + FP} \right).$$

A curva ROC, baseada nas medidas de sensibilidade e especificidade, possibilita a análise empírica da precisão e qualidade dos modelos, ao avaliar sua capacidade de discriminar corretamente entre árvores sobreviventes e mortas. A área sob a curva (AUC) foi calculada para representar a eficácia geral dos modelos na classificação dos resultados de mortalidade.

De acordo com Hosmer Junior et al. (2013) a regra geral para avaliação do resultado da área sob a Curva ROC é:

- Se  $AUC = 0,5$ : não há discriminação;
- Se  $0,5 \leq AUC \leq 0,7$ : baixa discriminação;
- Se  $0,7 \leq AUC \leq 0,8$ : discriminação aceitável;
- Se  $AUC \geq 0,8$ : excelente discriminação;

## RESULTADOS

Um resumo dos percentuais de falha e censura pode ser visualizado na Tabela 2, podendo notar que os grupos ecológicos secundários iniciais e tardios, apresentaram altos percentuais de censura.

**Tabela 2** – Porcentagem de falhas (mortalidade) e de censuras (sobrevivência) em um fragmento de Mata Atlântica no estado de Minas Gerais, Brasil.

	Pioneiras	Secundarias Iniciais	Secundarias Tardias	Total
Censura	39,30%	63,40%	63,50%	60,30%
Falha	60,70%	36,60%	36,50%	39,70%

De acordo com os testes de Hosmer Lemeshow, houve boa adequação dos modelos ajustados (Tabela 3). O teste apresentou  $p > 0.099$ , o que indicou que não houve grande desvio entre as predições dos modelos e as taxas de mortalidade e sobrevivência observadas.

**Tabela 3** – Resumo do teste de Hosmer e Lemeshow e critério de informação de Akaike para os modelos ajustados para avaliar a sobrevivência e mortalidade individual de arvores em um fragmento de Mata Atlântica.

Modelo	$\chi^2$	GL	Valor- <i>p</i>	AIC
Geral - Cox	9.747	8	0.2832	654.14
Geral - Logístico	12.274	8	0.1394	662.04
Pi - Cox	4.068	8	0.8509	95.70
Pi - Logístico	2.544	8	0.8635	98.12
Si – Cox	7.792	8	0.0995	292.23
Si - Logístico	7.487	8	0.1123	299.38
St – Cox	0.902	8	0.9242	136.99
St - Logístico	0.757	8	0.944	138.59

O modelo de taxas de falhas proporcionais de Cox e o modelo logístico de chances proporcionais apresentaram desempenho bem semelhante, na sobrevivência e mortalidade dos fustes (Tabela 4). Os valores de AUC superiores a 0,97 e de acurácia balanceada acima de 0,81, demonstraram a eficiência dos modelos na classificação individual das arvores.

**Tabela 4** – Medidas de precisão de treinamento e validação dos modelos de sobrevivência discretos na classificação de sobrevivência e mortalidade de árvores individuais no fragmento de Mata Atlântica. Sensibilidade (Sn), Especificidade (Ep), Area abaixo da curva (AUC), acurácia balanceada dos modelos ajustados.

Modelos	Sn	Ajuste			Validação			
		Ep	AUC	Acurácia	Sn	Ep	AUC	Acurácia
Geral -Cox	0.670	0.996	0.980	0.833	0.730	0.996	0.976	0.863
Geral - Logístico	0.670	0.995	0.981	0.832	0.730	0.994	0.977	0.862
P - Cox	0.833	0.995	0.992	0.914	0.869	0.990	0.987	0.929
P - Logístico	0.833	0.995	0.992	0.914	0.869	0.990	0.993	0.929
Si - Cox	0.754	1.000	0.992	0.877	0.764	1.000	0.992	0.877
Si - Logístico	0.754	1.000	0.992	0.877	0.764	1.000	0.992	0.877
St - Cox	0.623	1.000	0.986	0.811	0.757	1.000	0.982	0.878
St - Logístico	0.623	0.997	0.986	0.810	0.757	1.000	0.982	0.878

A precisão da classificação do número de árvores mortas (Sensibilidade), quando considerado todos os grupos ecológicos num único modelo, foi superior a 67% no ajuste e 73% na validação (Tabela 4). Quando os grupos ecológicos são analisados de forma separada, o acerto na taxa de mortalidade foi maior para as espécies pioneiras e secundárias tardias (Tabela 4). O maior percentual de acerto na validação em relação ao ajuste, caracteriza uma generalização eficiente dos modelos para dados independentes.

No modelo considerando todos os grupos ecológicos, o melhor ajuste foi obtido utilizando o pH, Matéria orgânica do solo (MOS), temperatura média, precipitação média e déficit hídrico de dois anos como covariáveis preditoras (Tabela 5). O aumento de um desvio-padrão nos valores de pH, MOS, precipitação e temperatura média diminuem o risco de mortalidade, ou seja, de o fuste morrer. Já o aumento de um desvio padrão do CWD\_2, ou seja, maior déficit hídrico, aumenta o risco de morte.

As mesmas covariáveis climáticas também foram selecionadas como preditoras para os modelos ajustados para cada grupo ecológico e apresentaram o mesmo comportamento em relação ao risco de morte das árvores (Tabela 6 e 7). Além das covariáveis de clima, a matéria orgânica do solo foi selecionada como variável preditora para as espécies pioneiras, a distância da borda,  $K^+$  e  $Ca^{2+}$  para secundárias iniciais e o  $Ca^{2+}$  para as secundárias tardias.

O aumento de um desvio padrão na MOS diminuiu o risco de mortalidade dos fustes para as espécies pioneiras e o  $K^+$  e distância da borda para espécies secundárias iniciais (Tabela 6 e 7). Quanto ao  $Ca^{2+}$ , o aumento de um desvio padrão aumentou o risco de morte das espécies secundárias iniciais e tardias.

**Tabela 5** – Valores estimados pelo modelo de Cox e Logístico para avaliar a sobrevivência e mortalidade individual de árvores em um fragmento de Mata Atlântica.

Cox				Logístico			
Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )	Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )
$\gamma_1^*$	<b>-30.8666</b>	3.4318	-	$\gamma_1^*$	<b>-34.7088</b>	3.9286	-
$\gamma_2^*$	<b>-20.7002</b>	1.5956	-	$\gamma_2^*$	<b>-23.9431</b>	1.9364	-
$\gamma_3^*$	<b>-13.9429</b>	1.3262	-	$\gamma_3^*$	<b>-15.6012</b>	1.5573	-
$\gamma_4^*$	<b>-6.6485</b>	0.5017	-	$\gamma_4^*$	<b>-7.5948</b>	0.6445	-
$\gamma_5^*$	<b>-3.5805</b>	0.3254	-	$\gamma_5^*$	<b>-3.3609</b>	0.4054	-
$\gamma_6^*$	<b>-2.1079</b>	0.2654	-	$\gamma_6^*$	<b>-2.0840</b>	0.3358	-
$\gamma_7^*$	<b>-0.07643</b>	0.1510	-	$\gamma_7^*$	<b>0.5051</b>	0.1858	-
$\beta_1$	<b>-0.36368</b>	0.0951	0.70	$\beta_1$	<b>-0.4573</b>	0.1174	0.63
$\beta_2$	<b>-1.77996</b>	0.1789	0.17	$\beta_2$	<b>-2.0855</b>	0.2168	0.12
$\beta_3$	<b>-3.88349</b>	0.3412	0.02	$\beta_3$	<b>-4.6269</b>	0.4202	0.01
$\beta_4$	<b>-1.65778</b>	0.1640	0.19	$\beta_4$	<b>-2.2803</b>	0.2334	0.10
$\beta_5$	<b>3.61898</b>	0.3108	37.30	$\beta_5$	<b>4.0185</b>	0.3737	55.62

$$\eta_{li} = \gamma_1^* * int1 + \gamma_2^* * int2 + \gamma_3^* * int3 + \gamma_4^* * int4 + \gamma_5^* * int5 + \gamma_6^* * int6 + \gamma_7^* * int7 + \beta_1 * pH + \beta_2 * \text{Materia orgânica} + \beta_3 * \text{Precipitação média} + \beta_4 * \text{Temperatura média} + \beta_5 * CWD\_2 ; \gamma_i^* = \ln(\gamma_i) \text{ é o efeito do } i\text{-ésimo intervalo de tempo.}$$

Valores em negrito:  $p < 0.001$

**Tabela 6** - Valores estimados pelo modelo de Cox para avaliar a sobrevivência e mortalidade individual de árvores em um fragmento de Mata Atlântica.

Pioneiras				Secundarias Iniciais				Secundarias Tardias			
Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )	Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )	Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )
$\gamma_1^*$	<b>-40.7102</b>	13.6903	-	$\gamma_1^*$	<b>-40.5875</b>	10.4696	-	$\gamma_1^*$	<b>-34.6220</b>	14.8564	-
$\gamma_2^*$	<b>-27.4979</b>	6.1305	-	$\gamma_2^*$	<b>-27.2352</b>	2.9826	-	$\gamma_2^*$	<b>-23.4458</b>	4.8636	-
$\gamma_3^*$	<b>-18.455</b>	4.6782	-	$\gamma_3^*$	<b>-18.2284</b>	3.3103	-	$\gamma_3^*$	<b>-16.4418</b>	5.0876	-
$\gamma_4^*$	<b>-9.1662</b>	1.9436	-	$\gamma_4^*$	<b>-8.5541</b>	0.9142	-	$\gamma_4^*$	<b>-7.3748</b>	1.3157	-
$\gamma_5^*$	<b>-3.8748</b>	1.0474	-	$\gamma_5^*$	<b>-4.6515</b>	0.5678	-	$\gamma_5^*$	<b>-3.6153</b>	0.8064	-
$\gamma_6^*$	<b>-3.0151</b>	0.932	-	$\gamma_6^*$	<b>-1.8078</b>	0.4154	-	$\gamma_6^*$	<b>-1.9412</b>	0.6920	-
$\gamma_7^*$	<b>-0.2685</b>	0.4156	-	$\gamma_7^*$	<b>0.4549</b>	0.2547	-	$\gamma_7^*$	<b>0.7757</b>	0.4071	-
$\beta_1$	<b>-2.2203</b>	0.5989	0.11	$\beta_1$	<b>-0.7916</b>	0.3635		$\beta_1$	<b>2.3409</b>	0.5013	10.39
$\beta_2$	<b>-5.2272</b>	1.262	0.01	$\beta_2$	<b>3.8927</b>	0.5464		$\beta_2$	<b>-4.3200</b>	0.9725	0.01
$\beta_3$	<b>-1.9287</b>	0.5448	0.15	$\beta_3$	<b>-5.4301</b>	0.6331		$\beta_3$	<b>-2.0540</b>	0.4116	0.13
$\beta_4$	<b>5.4006</b>	1.1682	221.53	$\beta_4$	<b>-1.3465</b>	0.2455		$\beta_4$	<b>3.8228</b>	0.8308	45.73
				$\beta_5$	<b>5.7276</b>	0.6056					
				$\beta_6$	<b>-0.6325</b>	0.1555					

$P = \eta_{li} = \gamma_1^* * int1 + \gamma_2^* * int2 + \gamma_3^* * int3 + \gamma_4^* * int4 + \gamma_5^* * int5 + \gamma_6^* * int6 + \gamma_7^* * int7 + \beta_1 * Matéri\ orgânica + \beta_2 * Precipitação\ média + \beta_3 * Temperatura\ média + \beta_4 * CWD\_2$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

$Si = \eta_{li} = \gamma_1^* * int1 + \gamma_2^* * int2 + \gamma_3^* * int3 + \gamma_4^* * int4 + \gamma_5^* * int5 + \gamma_6^* * int6 + \gamma_7^* * int7 + \beta_1 * K^+ + \beta_2 * Ca^{2+} + \beta_3 * Precipitação\ média + \beta_4 * Temperatura\ média + \beta_5 * CWD_2 + \beta_6 * Distância\ da\ borda$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

$St = \eta_{li} = \gamma_1^* * int1 + \gamma_2^* * int2 + \gamma_3^* * int3 + \gamma_4^* * int4 + \gamma_5^* * int5 + \gamma_6^* * int6 + \gamma_7^* * int7 + \beta_1 * Ca^{2+} + \beta_2 * Precipitação\ média + \beta_3 * Temperatura\ média + \beta_4 * CWD\_2$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

Valores em negrito:  $p < 0.001$

**Tabela 7** - Valores estimados pelo Logístico para avaliar a sobrevivência e mortalidade individual de arvores em um fragmento de Mata Atlântica.

Pioneiras				Secundarias Iniciais				Secundarias Tardias			
Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )	Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )	Parâmetros	Estimativa	Erro Padrão	Exp ( $\beta$ )
$\gamma_1^*$	<b>-46.0966</b>	14.5425	-	$\gamma_1^*$	<b>-44.3558</b>	10.3403	-	$\gamma_1^*$	<b>-35.6674</b>	12.1583	-
$\gamma_2^*$	<b>-31.6625</b>	6.8078	-	$\gamma_2^*$	<b>-30.6059</b>	3.5315	-	$\gamma_2^*$	<b>-24.9754</b>	5.4420	-
$\gamma_3^*$	<b>-21.07</b>	5.1597	-	$\gamma_3^*$	<b>-19.4089</b>	3.7756	-	$\gamma_3^*$	<b>-16.6672</b>	4.9405	-
$\gamma_4^*$	<b>-10.4258</b>	2.3255	-	$\gamma_4^*$	<b>-9.3300</b>	1.0966	-	$\gamma_4^*$	<b>-8.0082</b>	1.5761	-
$\gamma_5^*$	<b>-3.8858</b>	1.3057	-	$\gamma_5^*$	<b>-4.6825</b>	0.7126	-	$\gamma_5^*$	<b>-3.3727</b>	1.0038	-
$\gamma_6^*$	<b>-3.1556</b>	1.1454	-	$\gamma_6^*$	<b>-1.6126</b>	0.4950	-	$\gamma_6^*$	<b>-1.9181</b>	0.8527	-
$\gamma_7^*$	<b>0.1859</b>	0.5531	-	$\gamma_7^*$	<b>1.0244</b>	0.3123	-	$\gamma_7^*$	<b>1.2722</b>	0.4802	-
$\beta_1$	<b>-2.3272</b>	0.6903	0.10	$\beta_1$	<b>-0.9649</b>	0.4355	0.38	$\beta_1$	<b>2.4997</b>	0.6053	12.18
$\beta_2$	<b>-6.3814</b>	1.4873	0.01	$\beta_2$	<b>4.4782</b>	0.6332	88.07	$\beta_2$	<b>-4.7845</b>	1.1438	0.01
$\beta_3$	<b>-2.5097</b>	0.7311	0.08	$\beta_3$	<b>-6.3316</b>	0.7733	0.01	$\beta_3$	<b>-2.6098</b>	0.5676	0.07
$\beta_4$	<b>6.2633</b>	1.372	524.96	$\beta_4$	<b>-1.8486</b>	0.3314	0.16	$\beta_4$	<b>3.9834</b>	0.9873	53.70
				$\beta_5$	<b>6.3505</b>	0.7389	572.76				
				$\beta_6$	<b>-0.7388</b>	0.1866	0.48				

$P = \eta_{li} = \gamma_1^* \cdot \text{int1} + \gamma_2^* \cdot \text{int2} + \gamma_3^* \cdot \text{int3} + \gamma_4^* \cdot \text{int4} + \gamma_5^* \cdot \text{int5} + \gamma_6^* \cdot \text{int6} + \gamma_7^* \cdot \text{int7} + \beta_1 \cdot \text{Matéri orgânica} + \beta_2 \cdot \text{Precipitação média} + \beta_3 \cdot \text{Temperatura média} + \beta_4 \cdot \text{CWD}_2$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

$Si = \eta_{li} = \gamma_1^* \cdot \text{int1} + \gamma_2^* \cdot \text{int2} + \gamma_3^* \cdot \text{int3} + \gamma_4^* \cdot \text{int4} + \gamma_5^* \cdot \text{int5} + \gamma_6^* \cdot \text{int6} + \gamma_7^* \cdot \text{int7} + \beta_1 \cdot K^+ + \beta_2 \cdot Ca^{2+} + \beta_3 \cdot \text{Precipitação média} + \beta_4 \cdot \text{Temperatura média} + \beta_5 \cdot \text{CWD}_2 + \beta_6 \cdot \text{Distância da borda}$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

$St = \eta_{li} = \gamma_1^* \cdot \text{int1} + \gamma_2^* \cdot \text{int2} + \gamma_3^* \cdot \text{int3} + \gamma_4^* \cdot \text{int4} + \gamma_5^* \cdot \text{int5} + \gamma_6^* \cdot \text{int6} + \gamma_7^* \cdot \text{int7} + \beta_1 \cdot Ca^{2+} + \beta_2 \cdot \text{Precipitação média} + \beta_3 \cdot \text{Temperatura média} + \beta_4 \cdot \text{CWD}_2$ ;  $\gamma_i^* = \ln(\gamma_i)$  é o efeito do i-ésimo intervalo de tempo.

Valores em negrito:  $p < 0.001$

## DISCUSSÃO

O desempenho semelhante do modelo de taxas de falhas proporcionais de Cox e o modelo logístico de chances proporcionais (Tabela 4) também foram verificados por Chalita et al. (2006) e hashimoto et al. (2011) em estudos de simulação.

Os resultados deste estudo indicam que a probabilidade de mortalidade individual pode ser predita prevista com uma boa precisão em fragmentos de Mata Atlântica, usando um modelo de sobrevivência discreto (Tabela 4). O percentual de acertos da mortalidade foi superior ao encontrado por Castro et al., 2015; Figueredo et al., 2020 e Rocha et al., 2018 que também usaram modelos de árvore individual para prever a mortalidade em fragmentos de mata atlântica.

A maior precisão encontrada pelos modelos de sobrevivência discretos neste trabalho em relação ao modelo de regressão logística (Figueredo et al., 2020) e de redes neurais (Castro et al., 2015; Rocha et al., 2018) pode estar relacionada a inclusão da censura no modelo. Mesmo sendo incompletas, as censuras fornecem informações relevantes sobre o tempo do evento e sua utilização no cálculo das estatísticas de interesse pode acarretar em conclusões menos viesadas (Colosimo e Giolo, 2024).

Outro fato que pode estar relacionado a melhor capacidade de predição dos modelos estudados é utilização dos intervalos de tempo ( $\gamma_j$ ) como parâmetros no modelo. Os intervalos podem abranger influências climáticas agregadas como condições de seca, chuvas excessivas, tempestades de vento, surtos de pragas e doenças e declínio por um ou vários anos. Estas influências podem aumentar as taxas de mortalidade para um intervalo específico, independentemente das condições do povoamento e do local (Fortin et al., 2008). Dessa forma a adição do tempo na função de verossimilhança traz melhorias na estimação dos parâmetros.

A melhor capacidade preditiva (Tabela 4) dos modelos ajustados (Tabela 6 e 7) para o grupo ecológico das espécies pioneiras pode estar relacionado ao menor percentual de censuras (Tabela 2) e menor número de espécies (Tabela 1). Quanto maior o percentual de censura nos dados, maior o erro de predição do modelo (Chalita et al., 2006). A presença de um grande número de espécies torna a modelagem da mortalidade individual das arvores mais complexa e desafiadora, pois geralmente, resulta de interações entre diversos fatores, muitas vezes em um processo gradual (Kim et al., 2017). Mesmo estando dentro do mesmo grupo ecológico, as espécies podem ter respostas diferentes em relação aos os efeitos ambientais, entomológicos, fisiológicos e patológicos (Hallinger et al., 2016; Hülsmann et al., 2016) que dificultam a predição, modelagem e prognóstico da mortalidade de árvores em florestas tropicais (Manso et al., 2015; Reis et al., 2018, Rocha et al., 2018).

Nos oito modelos selecionados, as covariáveis climáticas, déficit hídrico de dois anos, precipitação média e temperatura média anual foram selecionadas como covariáveis preditoras (Tabela 5, 6 e 7). Esses resultados corroboram com os encontrados por Brandt et al. (2020); Rocha et al. (2018) e Maringer et al. (2021), que ao incluir covariáveis climáticas nos modelos obtiveram maior capacidade preditiva, uma vez que a mortalidade de árvores está associada ao efeito climático (Cortini et al., 2017). Isso evidencia a importância da precipitação para manutenção das dinâmicas florestais, visto que períodos de seca podem contribuir para a redução na vitalidade de árvores, aumentando a sua susceptibilidade (Brandl et al., 2020). Vatandaşlar et al. (2023) também concluíram que o déficit hídrico do solo levou à maior mortalidade de árvores, ao correlacionar mudanças edáficas com mudanças climáticas.

Estes resultados demonstram o potencial dos modelos de sobrevivência discretos para o prognóstico da mortalidade e sobrevivência de árvores em florestas do bioma Mata Atlântica. Esses modelos podem contribuir para o manejo florestal, por meio da criação de cenários e simulações de eventos climáticos extremos (Anderegg et al., 2012) e para prever a longevidade individual (Holzwarth et al., 2013).

Em relação as covariáveis de solo selecionadas como preditoras, o aumento do pH e a redução do risco da mortalidade das árvores, pode estar ligado ao fato de que algumas espécies pioneiras e secundárias iniciais apresentam maior desenvolvimento em solos menos ácidos (Bizuti et al., 2018; Morais Junior et al., 2020). O aumento do pH pode favorecer a absorção e incorporação de nutrientes por estas espécies, tornando-os um pouco mais resistentes a eventos que possam causar algum tipo de estresse fisiológico (Schulze et al., 2019), como a seca por exemplo. Solos com pH mais ácidos podem conter maior disponibilidade de  $Al^{3+}$ , que em concentrações mais elevadas pode limitar o crescimento das raízes de algumas espécies e como consequência serem mais sensíveis a eventos de seca (Chandra e Keshavkant, 2021).

A matéria orgânica é um bom indicador da qualidade do solo (Teixeira et al., 2020), quanto maior o seu teor, maior a retenção de água, fertilidade e controle da erosão do solo (Adhikari e Hartemink, 2016). Essas condições favorecem altas taxas de crescimento das árvores (Torres et al., 2023) e redução do risco de mortalidade.

Ao estudar como que fatores antropogênicos e ambientais afetam os processos demográficos de árvores em fragmentos florestais da Mata Atlântica, Torres et al. (2023) verificaram que o aumento dos teores de  $Ca^{2+}$  e  $K^+$  no solo foi positivamente relacionado às taxas de mortalidade. Os resultados para o  $Ca^{2+}$  corroboram com os encontrados neste trabalho, em que o aumento dos teores desse nutriente no solo aumentou o risco de mortalidade das árvores para as secundárias iniciais e tardias (Tabelas 6 e 7). No entanto, para o  $K^+$  encontramos

resultados contrários, com o aumento do teor desse nutriente, o risco de morte das espécies secundárias iniciais diminuiu. A redução da mortalidade pode estar associada ao fato deste nutriente ter um papel importante no estado energético, na translocação, armazenamento de assimilados e na manutenção da água nos tecidos vegetais (Zahoor et al. 2017). Desta forma, com o aumento dos teores de potássio no solo, os efeitos fisiológicos causados pelo déficit hídrico seriam amenizados para algumas espécies secundárias iniciais.

A diminuição no risco de mortalidade para as espécies secundárias iniciais com o aumento da distância da parcela até a borda da floresta era esperado. Árvores mais próximas às bordas da floresta geralmente são mais vulneráveis a danos causado pelo vento (Magnago et al., 2015, Silvério et al., 2019), infestação de lianas (Campbell et al., 2018) e incêndios (Guedes et al., 2020), que podem reduzir o seu vigor e levar a morte.

Este é o primeiro trabalho a utilizar modelos de sobrevivência discretos para modelar a sobrevivência de árvores em um fragmento de Mata Atlântica. A construção do modelo com covariáveis ambientais, representa uma alternativa promissora para prever a sobrevivência e mortalidade de árvores em florestas tropicais. Contudo, novas pesquisas adicionando fragilidade aos modelos podem trazer novos ganhos e ajudar a compreender melhor a dinâmica florestal neste bioma.

## CONCLUSÃO

A metodologia de análise de sobrevivência de dados discretos foi eficiente avaliar a sobrevivência e mortalidade individual de árvores em um fragmento de Mata Atlântica. Esta técnica pode ser uma ferramenta útil na tomada de decisões dos gestores florestais.

## REFERENCIAS

- Adhikari, K., Hartemink, A.E., 2016. Linking soils to ecosystem services — A global review. *Geoderma* 262, 101–111. <https://doi.org/10.1016/j.geoderma.2015.08.009>
- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Anderegg, W.R.L., Berry, J.A., Smith, D.D., Sperry, J.S., Anderegg, L.D.L., Field, C.B., 2012. The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off. *Proceedings of the National Academy of Sciences* 109, 233–237. <https://doi.org/10.1073/pnas.1107891109>
- Brandl, S., Paul, C., Knoke, T., Falk, W., 2020. The influence of climate and management on survival probability for Germany’s most important tree species. *Forest Ecology and Management* 458, 117652. <https://doi.org/10.1016/j.foreco.2019.117652>
- Bretfeld, M., Ewers, B.E., Hall, J.S., 2018. Plant water use responses along secondary forest

- succession during the 2015–2016 El Niño drought in Panama. *New Phytologist* 219, 885–899. <https://doi.org/10.1111/nph.15071>
- Campbell, M.J., Edwards, W., Magrath, A., Alamgir, M., Porolak, G., Mohandass, D., Laurance, W.F., 2018. Edge disturbance drives liana abundance increase and alteration of liana–host tree interactions in tropical forest fragments. *Ecology and Evolution* 8, 4237–4251. <https://doi.org/10.1002/ece3.3959>
- Castro, R.V.O., Soares, C.P.B., Leite, H.G., de Souza, A.L., Martins, F.B., Nogueira, G.S., 2015. Artificial Neural Networks Effectiveness to Estimate Mortality in a Semi-Deciduous Seasonal Forest. *Australian Journal of Basic and Applied Sciences*.
- Chalita, L.V.A.S., Colosimo, E.A., Demétrio, C.G.B., 2002. Likelihood Approximations and Discrete Models for Tied Survival Data. *Communications in Statistics - Theory and Methods* 31, 1215–1229. <https://doi.org/10.1081/STA-120004920>
- Chandra, J., Keshavkant, S., 2021. Mechanisms underlying the phytotoxicity and genotoxicity of aluminum and their alleviation strategies: A review. *Chemosphere* 278, 130384. <https://doi.org/10.1016/j.chemosphere.2021.130384>
- 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-Yrizar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Péliissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* 20, 3177–3190. <https://doi.org/10.1111/gcb.12629>
- Collett, D., 2023. *Modelling Survival Data in Medical Research*, 4th ed. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781003282525>
- Colosimo, E.A., Giolo, S.R., 2024. *Análise de sobrevivência aplicada*, 2nd ed. Editora Edgard Blucher, São Paulo, SP.
- Cortini, F., Comeau, P.G., Strimbu, V.C., Hogg, E.H. (Ted), Bokalo, M., Huang, S., 2017. Survival functions for boreal tree species in northwestern North America. *Forest Ecology and Management* 402, 177–185. <https://doi.org/10.1016/j.foreco.2017.06.036>
- de Moraes Junior, V.T.M., Jacovine, L.A.G., Alves, E.B.B.M., Torres, C.M.M.E., Faustino, I.S., França, L.C. de J., Rocha, S.J.S.S. da, Simiqueli, G.F., Silva, L.B., Cruz, R.A. la, 2020. Growth and survival of potential tree species for carbon-offset in degraded areas from southeast Brazil. *Ecological Indicators* 117, 106514. <https://doi.org/10.1016/j.ecolind.2020.106514>
- Figueiredo, L.T.M.D., Soares, C.P.B., Castro, R.V.O., Leite, H.G., Souza, A.L.D., Silva, G.F.D., Torres, C.M.M.E., 2020. Modeling Mortality of Individual Trees in a Seasonal Semi-Deciduous Forest. *Floresta Ambient.* 27, e20170980. <https://doi.org/10.1590/2179-8087.098017>
- Flores, B.M., Montoya, E., Sakschewski, B., Nascimento, N., Staal, A., Betts, R.A., Levis, C., Lapola, D.M., Esquivel-Muelbert, A., Jakovac, C., Nobre, C.A., Oliveira, R.S., Borma, L.S., Nian, D., Boers, N., Hecht, S.B., ter Steege, H., Arriera, J., Lucas, I.L., Berenguer, E., Marengo, J.A., Gatti, L.V., Mattos, C.R.C., Hirota, M., 2024. Critical transitions in the Amazon forest system. *Nature* 626, 555–564. <https://doi.org/10.1038/s41586-023-06970-0>
- Fortin, M., Bédard, S., DeBlois, J., Meunier, S., 2008. Predicting individual tree mortality in

- northern hardwood stands under uneven-aged management in southern Québec, Canada. *Ann. For. Sci.* 65, 205–205. <https://doi.org/10.1051/forest:2007088>
- Gandolfi, S., Leitão Filho, H. D. F., Bezerra, C. L. F., 1995. Levantamento florístico e caráter sucessional das espécies arbustivo-arbóreas de uma floresta mesófila semidecídua no município de Guarulhos, SP. *Revista brasileira de biologia*, 55(4), 753-767.
- Gonçalves Bizuti, D.T., Casagrande, J.C., Soares, M.R., Sartorio, S.D., Brugnaro, C., Gomes César, R., 2018. The effect of calcium on the growth of native species in a tropical forest hotspot. *iForest - Biogeosciences and Forestry* 11, 221. <https://doi.org/10.3832/ifor2074-010>
- Guedes, B.J., Massi, K.G., Evers, C., Nielsen-Pincus, M., 2020. Vulnerability of small forest patches to fire in the Paraíba do Sul River Valley, southeast Brazil: Implications for restoration of the Atlantic Forest biome. *Forest Ecology and Management* 465, 118095. <https://doi.org/10.1016/j.foreco.2020.118095>
- Hallinger, M., Johansson, V., Schmalholz, M., Sjöberg, S., Ranius, T., 2016a. Factors driving tree mortality in retained forest fragments. *Forest Ecology and Management* 368, 163–172. <https://doi.org/10.1016/j.foreco.2016.03.023>
- Hallinger, M., Johansson, V., Schmalholz, M., Sjöberg, S., Ranius, T., 2016b. Factors driving tree mortality in retained forest fragments. *Forest Ecology and Management* 368, 163–172. <https://doi.org/10.1016/j.foreco.2016.03.023>
- Hartmann, H., Adams, H.D., Anderegg, W.R.L., Jansen, S., Zeppel, M.J.B., 2015. Research frontiers in drought-induced tree mortality: crossing scales and disciplines. *New Phytologist* 205, 965–969. <https://doi.org/10.1111/nph.13246>
- Hashimoto, E.M., Ortega, E.M.M., Paula, G.A., Barreto, M.L., 2011. Regression models for grouped survival data: Estimation and sensitivity analysis. *Computational Statistics & Data Analysis* 55, 993–1007. <https://doi.org/10.1016/j.csda.2010.08.004>
- Hendrik, D., Maxime, C., 2017. Assessing drought-driven mortality trees with physiological process-based models. *Agricultural and Forest Meteorology* 232, 279–290. <https://doi.org/10.1016/j.agrformet.2016.08.019>
- Holzwarth, F., Kahl, A., Bauhus, J., Wirth, C., 2013. Many ways to die – partitioning tree mortality dynamics in a near-natural mixed deciduous forest. *Journal of Ecology* 101, 220–230. <https://doi.org/10.1111/1365-2745.12015>
- Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. *Applied logistic regression*, Third edition. ed, Wiley series in probability and statistics. Wiley, Hoboken, New Jersey.
- Hülsmann, L., Bugmann, H.K.M., Commarmot, B., Meyer, P., Zimmermann, S., Brang, P., 2016a. Does one model fit all? Patterns of beech mortality in natural forests of three European regions. *Ecological Applications* 26, 2465–2479. <https://doi.org/10.1002/eap.1388>
- Hülsmann, L., Bugmann, H.K.M., Commarmot, B., Meyer, P., Zimmermann, S., Brang, P., 2016b. Does one model fit all? Patterns of beech mortality in natural forests of three European regions. *Ecological Applications* 26, 2465–2479. <https://doi.org/10.1002/eap.1388>
- Jay, M., 2019. *generalhoslem: Goodness of Fit Tests for Logistic Regression Models*.
- Kim, M., Lee, W.-K., Choi, G.-M., Song, C., Lim, C.-H., Moon, J., Piao, D., Kraxner, F., Shividenko, A., Forsell, N., 2017. Modeling stand-level mortality based on maximum

- stem number and seasonal temperature. *Forest Ecology and Management* 386, 37–50. <https://doi.org/10.1016/j.foreco.2016.12.001>
- King, S.L., Bennett, K.P., List, S., 2000. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. *Computers and Electronics in Agriculture* 27, 401–406. [https://doi.org/10.1016/S0168-1699\(00\)00106-X](https://doi.org/10.1016/S0168-1699(00)00106-X)
- Laurance, W.F., 2009. Conserving the hottest of the hotspots. *Biological Conservation, Conservation Issues in the Brazilian Atlantic Forest* 142, 1137. <https://doi.org/10.1016/j.biocon.2008.10.011>
- Lutz, J.A., van Wagtenonk, J.W., Franklin, J.F., 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography* 37, 936–950. <https://doi.org/10.1111/j.1365-2699.2009.02268.x>
- Magnago, L.F.S., Rocha, M.F., Meyer, L., Martins, S.V., Meira-Neto, J.A.A., 2015. Microclimatic conditions at forest edges have significant impacts on vegetation structure in large Atlantic forest fragments. *Biodivers Conserv* 24, 2305–2318. <https://doi.org/10.1007/s10531-015-0961-1>
- Manso, R., Morneau, F., Ningre, F., Fortin, M., 2015. Incorporating stochasticity from extreme climatic events and multi-species competition relationships into single-tree mortality models. *Forest Ecology and Management* 354, 243–253. <https://doi.org/10.1016/j.foreco.2015.06.008>
- Maringer, J., Stelzer, A.-S., Paul, C., Albrecht, A.T., 2021. Ninety-five years of observed disturbance-based tree mortality modeled with climate-sensitive accelerated failure time models. *Eur J Forest Res* 140, 255–272. <https://doi.org/10.1007/s10342-020-01328-x>
- Meir, P., Mencuccini, M., Dewar, R.C., 2015. Drought-related tree mortality: addressing the gaps in understanding and prediction. *New Phytologist* 207, 28–33. <https://doi.org/10.1111/nph.13382>
- Paulo, S., n.d. Atlas dos Remanescentes Florestais da Mata Atlântica.
- 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-Velázquez, J., Ruschel, A.R., Salgado-Negret, B., Schiavini, I., Swenson, N.G., Tenorio, E.A., Thompson, J., Toledo, M., Uriarte, M., Hout, P. van der, Zimmerman, J.K., Peña-Claros, M., 2017. Biodiversity and climate determine the functioning of Neotropical forests. *Global Ecology and Biogeography* 26, 1423–1434. <https://doi.org/10.1111/geb.12668>
- Prentice, R.L., Gloeckler, L.A., 1978. Regression analysis of grouped survival data with application to breast cancer data. *Biometrics* 34, 57–67.
- R Development Core Team, 2024. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria. <https://www.r-project.org/>
- Redmond, M.D., 2022. CWD and AET function. <https://doi.org/10.5281/zenodo.6416352>
- Reis, L.P., de Souza, A.L., dos Reis, P.C.M., Mazzei, L., Soares, C.P.B., Miquelino Eleto Torres, C.M., da Silva, L.F., Ruschel, A.R., Rêgo, L.J.S., Leite, H.G., 2018. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. *Ecological Engineering* 112, 140–147. <https://doi.org/10.1016/j.ecoleng.2017.12.014>

- Rocha, S.J.S.S. da, 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. da, Reis, L.P., Zanuncio, J.C., 2018. Artificial neural networks: Modeling tree survival and mortality in the Atlantic Forest biome in Brazil. *Science of The Total Environment* 645, 655–661. <https://doi.org/10.1016/j.scitotenv.2018.07.123>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Schettini, B.L.S., Jacovine, L.A.G., Leite, H.G., Gelcer, E.M., Reis, L.P., Neves, K.M., Comini, I.B., Silva, L.F. da, 2020. Drought effects on carbon dynamics of trees in a secondary Atlantic Forest. *Forest Ecology and Management* 465, 118097. <https://doi.org/10.1016/j.foreco.2020.118097>
- Schulze, E.-D., Beck, E., Buchmann, N., Clemens, S., Müller-Hohenstein, K., Scherer-Lorenzen, M., 2019. *Plant Ecology*. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-662-56233-8>
- Silvério, D.V., Brando, P.M., Bustamante, M.M.C., Putz, F.E., Marra, D.M., Levick, S.R., Trumbore, S.E., 2019. Fire, fragmentation, and windstorms: A recipe for tropical forest degradation. *Journal of Ecology* 107, 656–667. <https://doi.org/10.1111/1365-2745.13076>
- Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T., Unterthiner, T., Ernst, F.G.M., 2020. ROCR: Visualizing the Performance of Scoring Classifiers.
- SOS Mata Atlântica, 2024. Atlas da Mata Atlântica. <https://www.sosma.org.br/iniciativas/atlas-da-mata-atlantica>.
- Teixeira, H.M., Cardoso, I.M., Bianchi, F.J.J.A., da Cruz Silva, A., Jamme, D., Peña-Claros, M., 2020. Linking vegetation and soil functions during secondary forest succession in the Atlantic forest. *Forest Ecology and Management* 457, 117696. <https://doi.org/10.1016/j.foreco.2019.117696>
- Torres, C.M.M.E., Medina-Vega, J.A., Rocha, S.J.S.S. da, Costa, W. da S., Soares, C.P.B., Souza, A.L. de, Jacovine, L.A.G., Lana, J.M. de, Peña-Claros, M., 2023. Drivers of tree demographic processes in forest fragments of the Brazilian Atlantic forest. *Forest Ecology and Management* 534, 120893. <https://doi.org/10.1016/j.foreco.2023.120893>
- Vatandaşlar, C., Türkeş, M., Semerci, A., Karahan, A., 2023. Analyzing climate-induced mortality of Taurus fir based on temporal forest management plans and climatic variations and droughts in the Central Mediterranean sub-region of Turkey. *Eur J Forest Res* 142, 61–89. <https://doi.org/10.1007/s10342-022-01506-z>
- Vieira, G.C., de Mendonça, A.R., da Silva, G.F., Zanetti, S.S., da Silva, M.M., dos Santos, A.R., 2018. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of The Total Environment* 619–620, 1473–1481. <https://doi.org/10.1016/j.scitotenv.2017.11.138>
- Vujovic, Ž.Đ., 2021. Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications (IJACSA)* 12. <https://doi.org/10.14569/IJACSA.2021.0120670>
- Zahoor, R., Zhao, W., Abid, M., Dong, H., Zhou, Z., 2017. Title: Potassium application regulates nitrogen metabolism and osmotic adjustment in cotton (*Gossypium hirsutum* L.) functional leaf under drought stress. *Journal of Plant Physiology* 215, 30–38. <https://doi.org/10.1016/j.jplph.2017.05.001>

## CAPITULO 3 – MÉTODOS DE ANÁLISE DE SOBREVIVÊNCIA BASEADOS EM APRENDIZADO DE MÁQUINA: MODELAGEM DA MORTALIDADE DE ARVORES NA MATA ATLÂNTICA

### RESUMO

A compreensão das respostas de florestas altamente fragmentadas, como a Mata Atlântica, às perturbações ambientais e antrópicas é essencial para o desenvolvimento de estratégias eficazes de conservação, manejo e restauração. Entender como esses fatores influenciam os processos demográficos, como crescimento, mortalidade e recrutamento, é fundamental para subsidiar práticas de manejo mais eficientes e sustentáveis. Neste sentido, o objetivo deste trabalho foi avaliar o uso de métodos de análise de sobrevivência baseados em aprendizado de máquina na mortalidade de árvores da Mata Atlântica, considerando covariáveis ambientais e antrópicas. Foram utilizados dados de 54 parcelas permanentes de quatro fragmentos florestais em Minas Gerais, Brasil. Foram utilizados como covariáveis antrópicas a área basal, histórico de uso da terra, idade de abandono, tamanho da floresta e distância da borda; covariáveis ambientais características químicas do solo (pH do solo em H<sub>2</sub>O, cátions trocáveis, teor de P e matéria orgânica do solo) e clima (temperatura média, precipitação < 100 mm, precipitação nos três meses mais secos, precipitação média anual total e déficit hídrico - CWD). Foram ajustados modelos de floresta de sobrevivência aleatória e regressão censurada de vetores de suporte para estudar a influência das covariáveis ambientais e antrópicas na mortalidade das árvores. A escolha do melhor modelo foi baseada no *C-index* e no *Brier-Score*. A *random survival forests* foi o algoritmo que apresentou melhor desempenho. A área basal e o grupo ecológico foram as covariáveis mais importantes. Foi possível prever a sobrevivência e mortalidade das árvores com boa precisão utilizando apenas três covariáveis: área basal, grupo ecológico e déficit hídrico de três anos. A metodologia *random survival forests* foi eficiente avaliar a sobrevivência e mortalidade individual de árvores nos fragmentos de Mata Atlântica. Esta técnica pode ser uma ferramenta útil na tomada de decisões dos gestores florestais.

**Palavras-chave:** Dados censurados; Floresta tropical; *random survival forests*.

### INTRODUÇÃO

As florestas tropicais desempenham um papel fundamental no ciclo global do carbono e abrigam mais da metade das espécies conhecidas no planeta (Pan et al., 2011). No entanto, a

intensificação das atividades antrópicas nas últimas décadas tem provocado uma crescente fragmentação desses ecossistemas (Taubert et al., 2018), especialmente na Mata Atlântica (Souza et al., 2021). Esse bioma, reconhecido como uma das 25 áreas globais mais importantes para a biodiversidade, abriga mais de 20.000 espécies de plantas (Campoe et al., 2014), embora permaneça apenas com 12,4% da cobertura florestal original (SOS Mata Atlântica, 2024).

A compreensão das respostas de florestas altamente fragmentadas, como a Mata Atlântica, às perturbações ambientais e antrópicas é essencial para o desenvolvimento de estratégias eficazes de conservação, manejo (Rocha et al., 2018; 2020; Vieira et al., 2018) e restauração (Torres et al., 2023). Devido à heterogeneidade de seus fragmentos remanescentes, a identificação dos fatores ambientais e antropogênicos que influenciam os processos demográficos, como crescimento, mortalidade e recrutamento, é fundamental para subsidiar práticas de manejo mais eficientes e sustentáveis (Torres et al., 2023).

Entre os processos demográficos, a mortalidade representa um dos fenômenos mais complexos de modelagem, por ser um evento relativamente raro (King et al., 2000, Reis et al., 2018) e ser causado por vários fatores aleatórios e complexos (Fan et al., 2006). Nesse contexto, a construção de modelos individuais em nível de árvore para estimar a sobrevivência e a mortalidade torna-se essencial para a previsão acurada do número de árvores, da área basal, da distribuição diamétrica e da produção florestal (Reis et al., 2018). Estas informações são fundamentais para entender a dinâmica nessas florestas.

A técnica de análise de sobrevivência tem sido amplamente utilizada para avaliar o efeito de covariáveis climáticas (Brandl et al., 2020; Manso et al., 2015; Neumann et al., 2017; Paul et al., 2019; Thiele et al., 2017), edáficas (Longpre e Morris, 2012; Maringer et al., 2021; Thapa et al., 2015) e de paisagem (Uzoh e Mori, 2012) sobre a mortalidade individual de árvores em florestas temperadas. Esta metodologia é baseada na análise do tempo até a ocorrência do evento de interesse e tem como principal característica a presença de dados censurados (Collet, 2023). A censura, corresponde a observação parcial da resposta de interesse e com a sua inclusão nos modelos, tem-se ganhos de informações relevantes sobre o tempo do evento, o que pode fornecer conclusões menos viesadas (Colosimo e Giolo, 2024).

No entanto, os modelos clássicos de sobrevivência como os de tempo de vida acelerado e o modelo de Cox, podem produzir resultados inconsistentes e desempenho inferior ao lidar com inúmeras covariáveis e efeitos não lineares (Huang et al., 2023; Senevirathne et al., 2024). Nesse contexto, para capturar as relações complexas entre covariáveis, métodos de análise de sobrevivência baseados em aprendizado de máquina vêm sendo utilizados, principalmente das ciências médicas (Huang et al., 2023; Pickett et al., 2021; Qiu et al., 2020). Dentre esses

métodos, podemos citar a floresta de sobrevivência aleatória (RSF) desenvolvida por Ishwaran et al. (2008) e a regressão censurada de vetores de suporte (SVCR) criada por Shivaswamy et al. (2007).

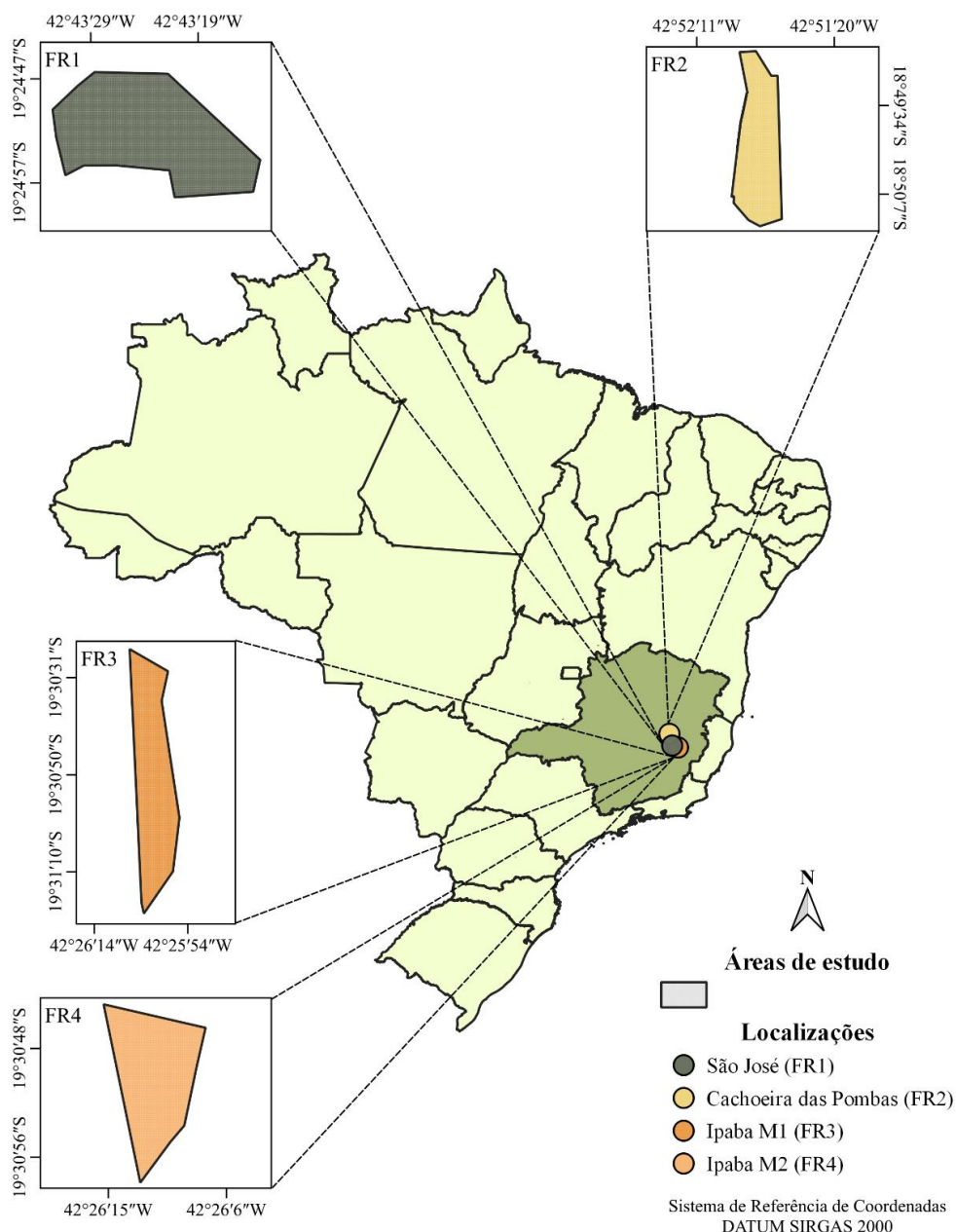
A RSF é uma extensão do algoritmo de *random forest* voltados para dados censurados. Este método de aprendizado de máquina tem facilidade em lidar com dados alta dimensão e a capacidade de incorporar relações não lineares entre as covariáveis resposta e preditoras e suas interações (Ishwaran et al. (2008). O SVCR é um método preditivo e supervisionado, capaz de lidar com dados de alta dimensionalidade e são capazes de incorporar relações não lineares e interações de forma automática nas covariáveis (Van Belle et al., 2011).

Neste sentido, o objetivo deste trabalho foi avaliar o uso de métodos de análise de sobrevivência baseados em aprendizado de máquina na mortalidade de árvores da Mata Atlântica, considerando covariáveis ambientais e antrópicas.

## **MATERIAL E METODOS**

### **Local de estudo e características das parcelas**

Foram utilizados dados de quatro fragmentos de Mata Atlântica, localizados em Minas Gerais, Brasil (Figura 1). A vegetação é classificada como floresta estacional semidecidual (IBGE, 2012) e os fragmentos variam em clima e solo (Tabela 1). Todos os fragmentos são florestas secundárias em regeneração após serem desmatados e/ou usados para extração de madeira, agricultura, plantações de eucalipto. Atualmente, todos os fragmentos são áreas de conservação ou proteção, garantindo a sucessão ecológica nos últimos anos (Torres et al., 2023).



**Figura 1** – Localização dos quatro fragmentos de Mata Atlântica no estado de Minas Gerais, Brasil. FR1: São José; FR2: Cachoeira das Pombas; FR3: Ipaba Mata1; FR4: Ipaba Mata 2.

Nos quatro fragmentos florestais foram estabelecidas de forma aleatória, 54 parcelas com 500 m<sup>2</sup> (10 × 50 m). Todos os fustes com diâmetro à altura do peito  $\geq$  5 cm (1,3 m de altura, DAP) foram recrutados, medidos e identificados botanicamente. As parcelas foram monitoradas ao longo dos anos de 2002, 2007, 2012, 2017 e 2023. As espécies foram classificadas em grupos ecológicos seguindo a divisão proposta por Gandolfi et al. (1995), para florestas secundárias brasileiras em: pioneiras (P); secundárias iniciais (Si) e secundárias tardias (St). A mortalidade foi avaliada considerando cada fuste como um indivíduo.

**Tabela 1** – Caracterização de quatro fragmentos de Mata Atlântica em Minas Gerais, Brasil, por meio de média e desvio padrão, das características de clima, solo, paisagem e atributos da floresta. CWD = déficit hídrico climático; MOS – Matéria orgânica do solo; BA – Área basal.

	Cachoeira das Pombas	Ipaba Mata 1	Ipaba Mata 2	São José
<b>Clima</b>				
Precipitação (mm)	1112.74 (±214.50)	1194.38 (±298.96)	1194.38 (±298.96)	1263.92 (±226.07)
CWD (mm)	-1021.36 (±432.50)	-2020.77 (±563.16)	-1499.09 (±502.88)	-940.19 (±395.45)
Precipitação meses mais secos (mm)	30.84 (±33.57)	33.99 (±30.13)	33.99 (±30.13)	39.84 (±30.99)
Precipitação Menos 100 mm	7.82 (±1.17)	7.53 (±1.28)	7.53 (±1.28)	7.38 (±1.18)
Temperatura Média (°C)	20.85 (±1.59)	24.15 (±1.73)	24.15 (±1.73)	19.51 (±1.97)
<b>Solo</b>				
pH (H <sub>2</sub> O)	4.52 (±0.25)	3.94 (±0.15)	3.76 (±0.08)	4.23 (±0.55)
K (cmolc.dm <sup>-3</sup> )	0.16 (±0.11)	0.08 (±0.01)	0.07 (±0.01)	0.09 (±0.03)
Ca <sup>2+</sup> (cmolc.dm <sup>-3</sup> )	0.49 (±0.45)	0.31 (±0.08)	0.16 (±0.07)	0.49 (±0.56)
Mg <sup>2+</sup> (cmolc.dm <sup>-3</sup> )	0.24 (±0.27)	0.17 (±0.06)	0.11 (±0.01)	0.16 (±0.10)
Al <sup>3+</sup> (cmolc.dm <sup>-3</sup> )	1.06 (±0.43)	1.40 (±0.38)	1.82 (±0.17)	0.91 (±0.34)
P-rem (mg.L <sup>-1</sup> )	10.69 (±2.07)	21.54 (±2.32)	21.85 (±1.65)	21.23 (±5.17)
MOS (dag.kg <sup>-1</sup> )	5.64 (±1.75)	3.89 (±0.57)	3.00 (±0.28)	5.14 (±1.56)
<b>Paisagem</b>				
Tamanho da Floresta (ha)	106.00	264.00	37.30	120.71
Elevação (m)	959.80 (±88.09)	267.75 (±19.29)	267.83 (±25.02)	846.67 (±56.93)
Declive (%)	30.52 (±9.22)	26.25 (±13.27)	29.07 (±3.94)	36.35 (±15.01)
Ângulo de Inclinação(°)	161.27 (±88.05)	265.53 (±36.29)	109.46 (±28.02)	179.92 (±26.83)
Distância da borda (m)	101.90	88.58	74.16	38.4
<b>Atributos da floresta</b>				
Espécies Arboreas	31 (±7)	25 (±8)	37 (±9)	40 (±14)
Fustes (ha)	2008 (±664)	1765 (±454)	1825 (±264)	1616 (±344)
BA (m <sup>2</sup> .ha <sup>-1</sup> )	18.85 (±6.32)	16.46 (±6.05)	25.38 (±7.61)	22.99 (±8.45)
Parcelas	20	16	6	12

### **Atributos da floresta e Covariáveis antrópicas**

Os atributos da floresta e covariáveis antrópicas foram definidos a nível de parcela, as seguintes covariáveis foram utilizadas: A área basal, histórico de uso da terra, idade de abandono, tamanho da floresta e distância da borda. A área basal total por parcela foi usada como um proxy para a idade da floresta porque a área basal está fortemente associada à idade da floresta (Menezes e Melo, 2019, Van Breugel et al., 2019).

O histórico de uso da terra foi categorizado em produção agrícola, desmatamento e corte seletivo. A categoria desmatamento se refere à remoção de floresta para uso de madeira em serrarias ou para produção de carvão. Essas categorias foram baseadas em entrevistas com proprietários, títulos de propriedades, fotografias aéreas (de 1960, 1980 e 1990) e imagens de satélite Landsat de 1985.

### **Covariáveis climáticas**

Foram obtidos a precipitação anual, o número de meses com precipitação < 100 mm, a precipitação nos três meses mais secos e a temperatura média anual de cada fragmento da estação climatológica mais próxima. A partir destes dados foi estimado o déficit hídrico climático (CWD), que é a diferença entre a Evapotranspiração Potencial (ETP) e a Evapotranspiração Real (ETR), de acordo a metodologia proposta por Lutz et al. (2010), utilizando as funções CWD e AET (Redmond, 2022) do software R versão 4.4.2. (R Development Core Team, 2024). O CWD reflete as condições de seca com mais precisão do que a precipitação anual total (Chave et al., 2014). CWD com menores valores negativos indica condições de alto estresse hídrico, e valores próximos a 0 (zero) indicam nenhum estresse hídrico (Poorter et al., 2017). A precipitação média anual total e CWD, foram calculados para quatro, três, dois e um ano antes do inventário de cada parcela.

### **Covariáveis de solo e topografia**

Foram coletados aleatoriamente de 20 a 30 amostras de solo para obter uma amostra composta para duas profundidades (0–20 cm e 20–40 cm) em cada parcela. A partir destas amostras foram determinados: pH do solo em H<sub>2</sub>O, cátions trocáveis (Ca<sup>2+</sup>, Mg<sup>2+</sup> e Al<sup>3+</sup>), acidez total (H<sup>+</sup> + Al<sup>3+</sup>), capacidade de troca catiônica (CTC), saturação por bases (V), fósforo disponível (P), P remanescente em solução (P-rem) e matéria orgânica do solo (MOS), determinados por métodos padrão (Teixeira et al., 2017). As análises de solo foram realizadas no laboratório de Fertilidade do Solo da Universidade Federal de Viçosa (UFV), Brasil.

Para levar em conta as diferenças na topografia, foi calculado para cada parcela a

elevação, a inclinação e o aspecto da inclinação usando a ferramenta de inclinação no ArcGIS 10.3.1. Para o cálculo foi utilizado um modelo digital de elevação (DEM) da Missão Espacial da NASA - Shuttle Radar Topography Mission (SRTM) para a análise.

### **Analise dos dados**

As covariáveis quantitativas foram padronizadas para o score  $Z$ , para melhorar a convergência do modelo, uma vez que muitas covariáveis tem unidades de medidas diferentes.

$$Z_i = \frac{x_i - \bar{x}}{\sigma}$$

onde,  $Z_i$  é valor padronizado da observação  $i$ ;  $x_i$  é o valor da observação  $i$ ;  $\bar{x}$  é a média dos valores observados;  $\sigma$  é o desvio padrão.

### **Random survival forests (RSF)**

A *random survival forest* é um algoritmo baseado em um conjunto de árvores que pode ser aplicado para dados de sobrevivência, sendo uma extensão do método random forest (Ishwaran et al., 2014). No RSF um conjunto de árvores de sobrevivência são construídas a partir de diferentes subconjuntos de dados e covariáveis. As previsões são agregadas para melhorar a precisão e reduzir a variância do modelo (Ishwaran et al. 2008). Inicialmente, o conjunto total de covariáveis são subdivididos de forma a se obter a melhor divisão do conjunto, com base em algum critério de divisão (log-rank, log-rank score, entre outros). Em seguida, cada árvore é construída usando uma amostra bootstrap dos indivíduos em estudo.

A implementação do algoritmo da RSF pode ser resumida nos seguintes passos:

1° – No conjunto de dados original são retiradas aleatoriamente  $B$  (número de árvores - ntree) amostras bootstrap. Em média, cada amostra bootstrap irá conter aproximadamente 2/3 dos dados. O 1/3 restante dos dados é denominado de dados *out-of-bag* (dados OOB), e é utilizada para se estimar o erro de predição do modelo;

2° - Para cada amostra bootstrap, uma árvore de decisão binária é construída, pela seleção aleatória de um subconjunto de covariáveis (*mtry*), em cada um dos nós da árvore. A divisão do conjunto de covariáveis é realizada por meio de um critério de divisão para dados censurados, de modo a maximizar as diferenças de sobrevivência entre dois nós vizinhos (filhos);

Como critério para a divisão do conjunto de observações em cada nó da árvore de sobrevivência binária foi utilizado a estatística de log-rank. Seja  $i$  denotando um indivíduo, com  $i = \{1, \dots, n\}$ , em que  $n$  denota o número total de observações em um dado nó  $h$ .

Considerando a covariável  $x_1$ , a regra de divisão se baseia em encontrar o ponto de corte  $c$ , tal que:  $x \leq c$  (indivíduo é direcionado ao nó 1) ou  $x > c$  (indivíduo é direcionado ao nó 2), que maximiza a diferença de sobrevivência entre os indivíduos dos dois nós. Este critério se fundamenta na estatística log-rank, dada por:

$$L(x, c) = \frac{\sum_{j=1}^N \left( d_{i,j} - Y_{i,j} \frac{d_i}{Y_i} \right)}{\sqrt{\sum_{i=1}^N \frac{Y_{i,j}}{Y_i} \left( 1 - \frac{Y_{i,j}}{Y_i} \right) \left( \frac{Y_i - d_i}{Y_i - 1} \right) d_i}}$$

em que:  $d_{i,j}$  e  $Y_{i,j}$  são respectivamente, o número de eventos e o número de indivíduos sob risco no tempo  $t_i$  nos nós filhos  $j = 1, 2$  e  $N$  o número de tempos de falhas distintos  $t_1 < t_2 < \dots < t_N$ . Neste critério, o valor  $|L(x, c)|$  é utilizado como medida para a divisão em cada um dos nós, quanto maior o valor, maior será a diferença entre os grupos (Ishwaran et al., 2008). O processo foi repetido para todas as covariáveis, até que fosse encontrada uma covariável  $x^*$  e o respectivo ponto de corte  $c^*$ , tais que:  $|L(x^*, c^*)| \geq |L(x, c)|$ , para qualquer covariável  $x$  e ponto de corte  $c$ . Este procedimento foi repetido em cada nó, até que os nós terminais fossem alcançados.

3° - Cada árvore cresce até o tamanho máximo, de modo que o nó terminal atinja o número mínimo de eventos com tempos de sobrevivência exclusivos;

4° passo - Para uma árvore com crescimento completo, foi calculada a função de risco acumulado ( $\hat{\Lambda}_h$ ), para um dado tempo  $t$ , utilizando o estimador de Nelson-Aalen:

$$\hat{\Lambda}_h(t) = \sum_{t_{i,h} \leq t} \frac{d_{i,h}}{Y_{i,h}}$$

onde  $t_{1,h} < t_{2,h} < \dots < t_{N(h),h}$  são os  $N(h)$  tempos de evento distintos do nó terminal  $h$ ;  $d_{i,h}$  são os números de eventos e  $Y_{i,h}$  são o número de indivíduos sob risco no tempo  $t_{i,h}$ . Deste modo, em cada nó terminal é associada uma estimativa da taxa de falha acumulada. Assim, é estimada uma sequência de taxas de falhas para cada árvore binária e do agrupamento dessas sequências, obtém-se a taxa de falha acumulada de cada árvore.

A estimativa da taxa de falha de um indivíduo específico  $i$  (na amostra  $b$ ), considerando o conjunto de covariáveis explicativas ( $x_i$ ), é denotada por  $\hat{\Lambda}_b(t|x_i)$  e pode ser obtida percorrendo a árvore, na tentativa de identificar em qual dos nós terminais o indivíduo está contido. Caso o indivíduo seja encontrado no nó  $h$ , a estimativa da taxa de falha acumulada será igual a  $\hat{\Lambda}_b(t|x_i) = \hat{\Lambda}_h(t)$ . A taxa de falha acumulada conjunta é obtida pela soma das taxas

de falha acumuladas estimadas para cada amostra bootstrap, dividida pelo número de árvores, e pode ser obtida por:

$$\hat{\Lambda}(t|x_i) = \frac{1}{B} \sum_{b=1}^B \hat{\Lambda}_b(t|x_i).$$

5º passo - Estima-se a taxa de erro, computada a partir dos dados OOB (ISHWARAN et al., 2008).

Para melhorar os resultados preditivos do modelo, uma seleção adequada de hiperparâmetros é importante na construção de modelos RSF. Com base em vários testes preliminares e na literatura (Dauda, 2022; Hao et al., 2021, Senevirathne et al., 2024), os hiperparâmetros do modelo foram determinados como  $n_{tree} = 800$ ,  $m_{try} = 6$  (padrão  $\approx \sqrt{\text{número de covariáveis}}$ ),  $n_{split} = 5$ ,  $sample_{type} = \text{"swr"}$  (amostra sem substituição),  $min_{split} = 1\%$  dos dados e  $n_{odesize} = 1$ . Os valores padrão foram usados para os parâmetros restantes.

Por meio dos dados OOB foi calculado a medida de importância da variável (VIMP), associado a cada covariável. O VIMP é definido como a mudança no erro de predição em um novo conjunto de teste, se a covariável  $x$  não estiver disponível, dado que  $x$  foi usado para aumentar a floresta original (Ishwaran et al., 2008). Este cálculo é realizado por meio da permutação de observações fora da amostra (OOB) que não estavam envolvidas na construção das árvores. Valores altos e positivos de VIMP estão relacionados à covariáveis com maior poder preditivo, já valores negativos ou nulos estão associados à covariáveis com baixo ou nenhum poder preditivo.

O pacote *randomForestSRC* (Ishwaran e Kogalur, 2024) foi utilizado para o ajuste dos modelos de RSF.

### **Regressão censurada de vetores de suporte (SVCR)**

Um método baseado em máquinas de vetores de suporte para lidar com dados censurados (SVCR - Support Vector Censure Regression) foi proposto por Shivaswamy et al. (2007). Neste método, o objetivo é encontrar uma função de predição que minimize um risco empírico modificado adequado para dados de sobrevivência.

O modelo de regressão de vetor de suporte para dados censurados é expresso por:

$$\min_{w,b,\epsilon,\epsilon^*} \frac{1}{2} w^T w + \gamma \sum_{i=1}^n (\epsilon_i + \epsilon_i^*),$$

Sujeito as restrições,

$$\begin{cases} w^T \phi(x_i) + b \geq y_i - \epsilon_i, & \forall i = 1, \dots, n \\ -\delta_i (w^T \phi(x_i) + b) \geq -\delta_i y_i - \epsilon_i^*, & \forall i = 1, \dots, n \\ \epsilon_i \geq 0, & \forall i = 1, \dots, n \\ \epsilon_i^* \geq 0, & \forall i = 1, \dots, n. \end{cases}$$

Onde  $\mathbf{w} \in \mathbb{R}^p$  e  $b \in \mathbb{R}$  são os parâmetros a serem estimados,  $\mathbf{x}_i \in \mathbb{R}^p$  o vetor composto por  $p$  covariáveis explicativas,  $y_i \in \{-1, +1\}$  o rótulo de classe, para  $i = 1, \dots, n$  observações da amostra de treinamento.

O índice prognóstico para um novo ponto  $x^*$  é encontrado como

$$\hat{u}(x^*) = \sum_i (\alpha_i - \delta_i \alpha_i^*) \phi(x_i)^T \phi(x^*) + b$$

com  $\alpha_i$  e  $\alpha_i^*$  os multiplicadores de Lagrange. Uma vantagem dos modelos baseados em máquina de vetor de suporte é que o mapa de características  $\phi(\cdot)$  não precisa ser definido explicitamente. Pode ser escrito como:

$$k(x_i, x_j) = \phi(x_i)^T \phi(x_j)$$

desde que  $k(\cdot, \cdot)$  seja um kernel positivo definido.

Neste trabalho duas funções de kernel foram consideradas no ajuste dos modelos: a de base radial e o kernel aditivo.

O pacote *survivalsvm* (Fouodo, 2018) foi utilizado para o ajuste dos modelos de SVCR.

### Critérios para a comparação dos modelos

Para comparação dos diferentes modelos, duas medidas de desempenho foram utilizadas:

Índice de concordância (*C - index*): O índice de concordância é a razão entre o número de pares concordantes e o número de pares comparáveis:

$$C - \text{index}(u) = \frac{\sum_{i=1}^n \sum_{j \neq i} \mathcal{J}[(u(x_i) - u(x_j))(y_i - y_j) \geq 0]}{\sum_{i=1}^n \sum_{j \neq i} \text{comp}(i, j)}$$

com  $0 \leq C - \text{index}(u) \leq 1$ , sendo  $u(x_i)$  é uma função de predição para  $x_i$ .

O *Brier-Score* (BS) pode ser estimado pela expressão:

$$BS(t) = \frac{1}{N} \sum_{i=1}^N w_i(t) [\hat{S}(t|x_i) - y_i(t)]^2$$

em que:  $y_i(t) = I(t_i \geq t)$  é o verdadeiro status do indivíduo  $i$  no tempo  $t$ ,  $S(t|x_i)$  representa a probabilidade de sobrevivência do indivíduo  $i$  no tempo  $t$ , dado o conjunto de covariáveis preditoras,  $N$  o número de indivíduos e  $w_i(t)$  denota o peso associado ao  $i$ -ésimo indivíduo. Os pesos são estimados com base no estimador Kaplan-Meier para a função de sobrevivência ( $S$ ), de modo que:

$$w_i(t) = \begin{cases} \delta_i / S(y_i) & \text{se } y_i \leq t \\ 1 / S(y_i) & \text{se } y_i > t \end{cases}$$

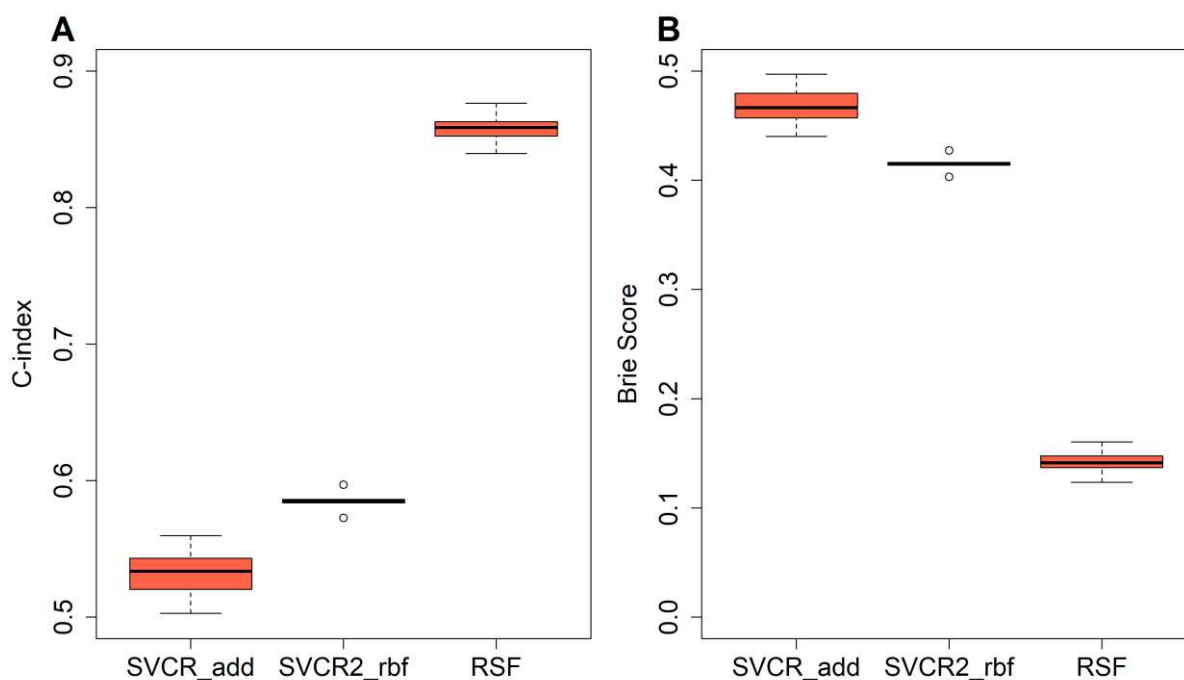
O *C-Index* é uma medida que apresenta valores no intervalo  $[0,1]$ , sendo o valor um indicando um modelo com predições perfeitas e 0,5 um modelo com predições aleatórias.

Já o *Brier-Score* apresenta valor 0 para predições perfeitas, 0,25 para predições razoáveis, e 1 para um baixo poder de predição.

O banco de dados foi separado em dois grupos, de modo que as proporções de observações censuradas nos dois conjuntos fossem aproximadamente iguais: um para treinamento dos modelos (75%) e outro para validação (25%). Esse processo foi repetido 100 vezes, obtendo-se a média das métricas para comparação de todos os modelos.

## RESULTADOS

O modelo baseado em *random survival forests* apresentou os maiores valores do índice de concordância (Figura 2A) e menores valores do *Brier-Score* (Figura 3B) para todas as repetições. Isso sugere que o modelo RSF é eficaz na predição da mortalidade de árvores na Mata Atlântica. Entre os modelos de regressão censurada de vetores de suporte, o modelo com função de kernel aditivo foi o que apresentou pior desempenho.



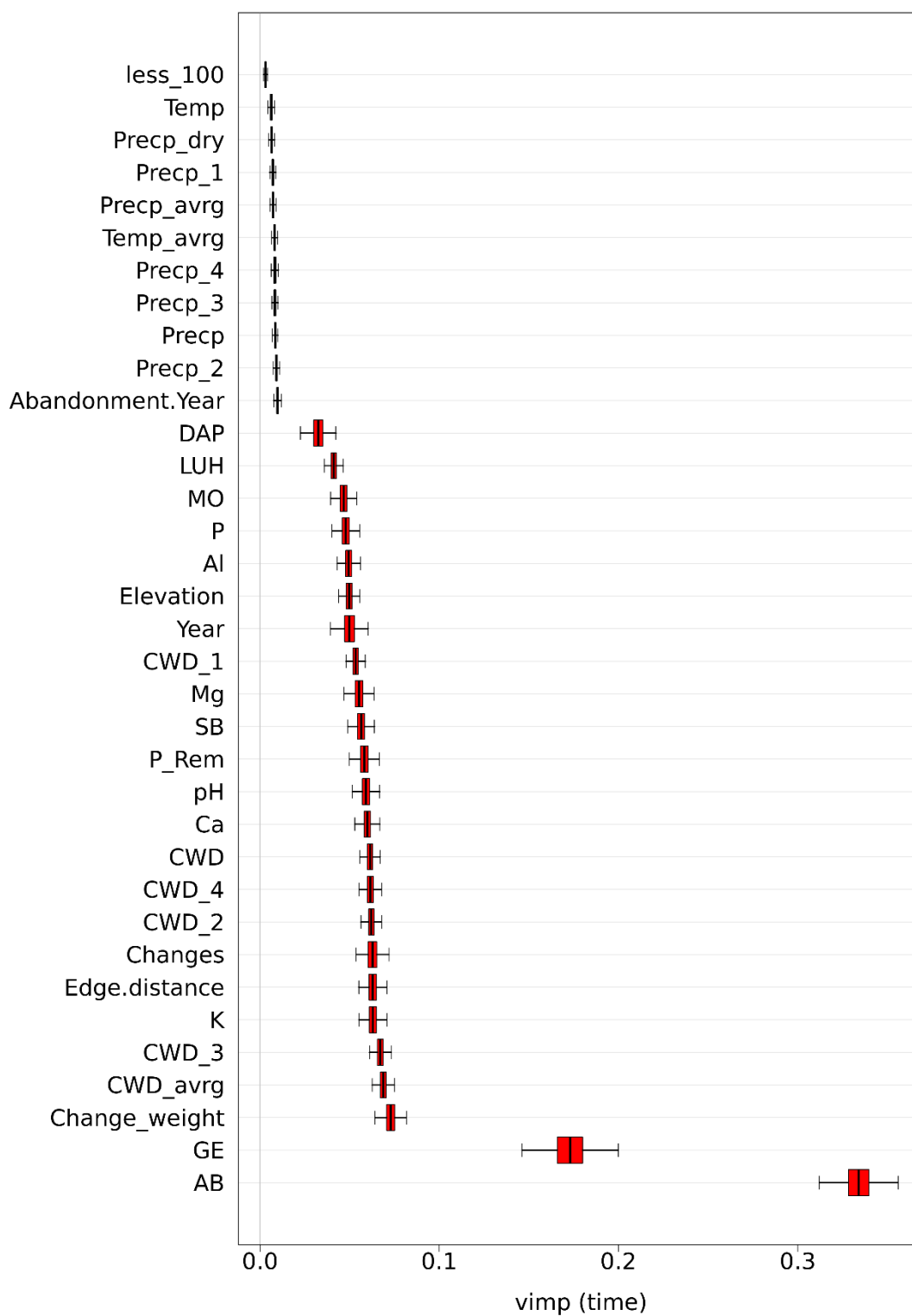
**Figura 2** - Boxplots do *C-Index* (A) e *Brier-Score* (B) para os métodos de análise de sobrevivência baseados em aprendizado de máquina na mortalidade de árvores da Mata Atlântica. SVCR\_add - Regressão censurada de vetores de suporte com função kernel aditivo, SVCR\_rbf – Regressão censurada de vetores de suporte com função de base radial, RSF – *random survival forests*.

A área basal e o grupo ecológico foram as covariáveis mais relevantes na predição da mortalidade (Figura 3). Por outro lado, a precipitação, a temperatura e a idade do abandono não demonstraram influência significativa no modelo. O déficit hídrico, a mudança no uso do solo, a distância da parcela até a borda da floresta e as covariáveis de solo apresentaram importância consideravelmente menor, sem diferenças estatisticamente significativas entre si.

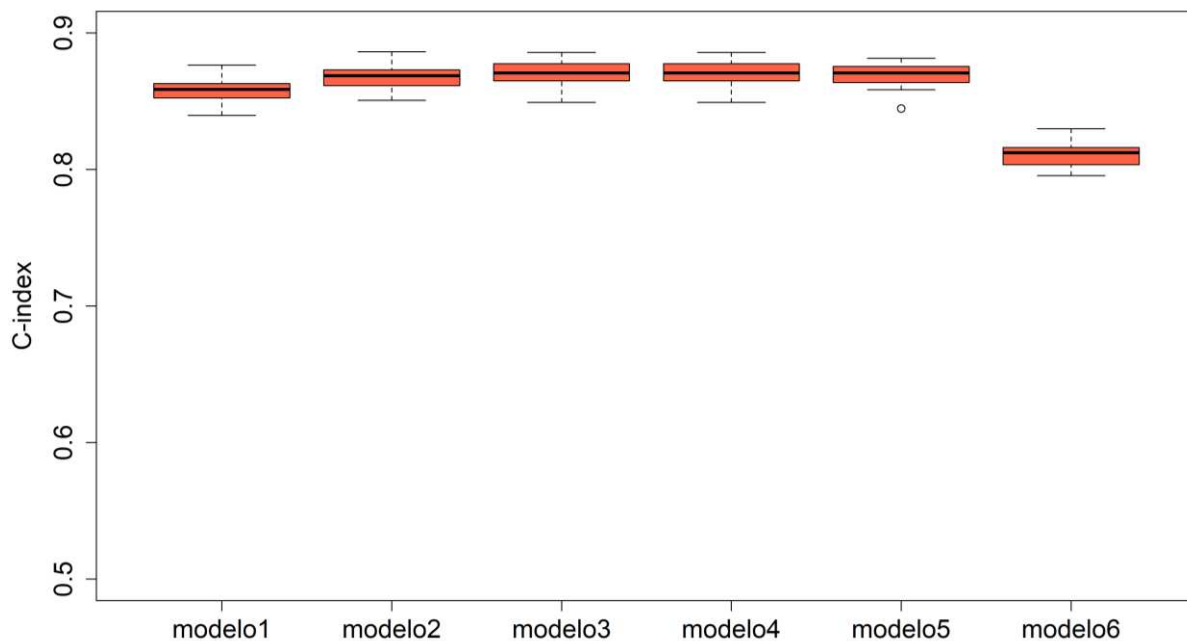
A precisão da predição da mortalidade permanece em um nível relativamente similar ao modelo completo, mesmo quando somente as três covariáveis mais importantes (Área basal, Grupo ecológico e CWD\_3) são mantidas no modelo (Tabela 2, Figura 4 e 5). No entanto, uma simplificação do modelo, para apenas as duas principais covariáveis (Área basal e Grupo ecológico) resulta em valores mais baixos de *C-index* e um *Brier-Score* mais alto, indicando menor precisão de predição em comparação aos outros modelos estudados.

**Tabela 2** – Lista de covariáveis preditoras incluídas nos modelos completo e simplificado.

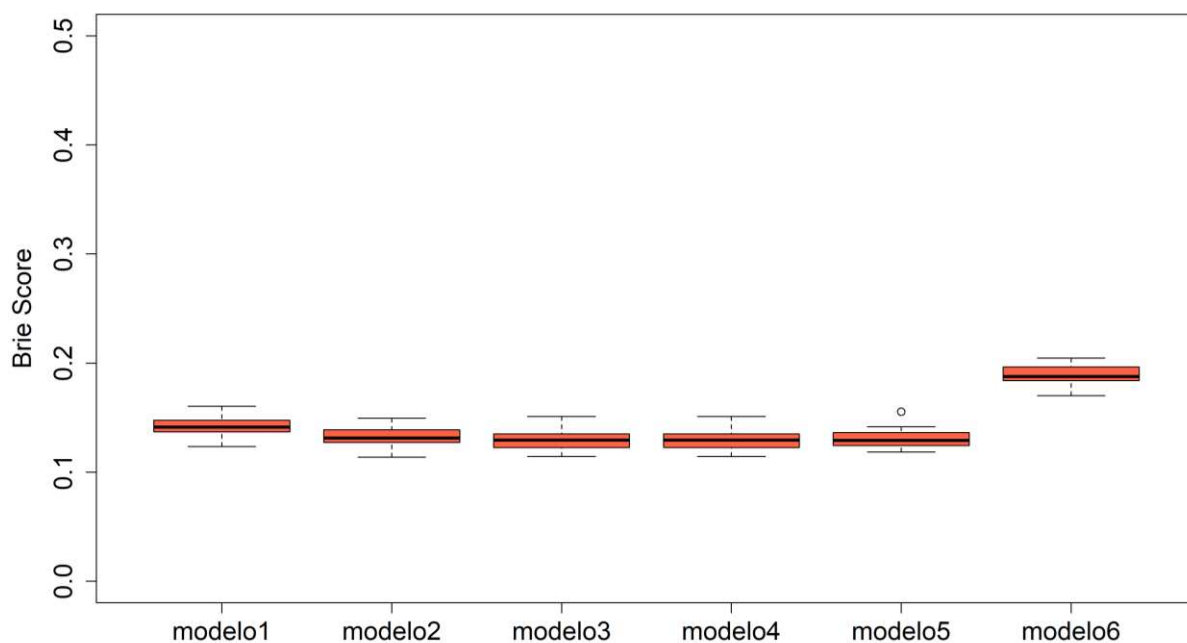
Modelos	Covariáveis
Modelo 1	Todas as covariáveis
Modelo 2	área basal, grupo ecológico, déficit hídrico de 3 anos, $K^+$ , distância até a borda, mudanças de uso do solo.
Modelo 3	área basal, grupo ecológico, déficit hídrico de 3 anos, $K^+$ , distância até a borda
Modelo 4	área basal, grupo ecológico, déficit hídrico de 3 anos, $K^+$
Modelo 5	área basal, grupo ecológico, déficit hídrico de 3 anos
Modelo 6	área basal, grupo ecológico



**Figura 3** – Valores médios de importância das variáveis (VIMP) estimados por meio do modelo de *random survival forests* (RSF) para a predição da mortalidade de árvores em fragmentos de Mata Atlântica no estado de Minas Gerais, Brasil.



**Figura 4** - Boxplots do desempenho preditivo de modelos simplificados (com base na importância da variável) versus modelos de floresta de sobrevivência aleatória completa (RSF) usando *C-Index*.



**Figura 5** - Boxplots do desempenho preditivo de modelos simplificados (com base na importância da variável) versus modelos de floresta de sobrevivência aleatória completa (RSF) usando *Brier-Score*.

## DISCUSSÃO

Este estudo evidenciou que as *random survival forests* apresentaram desempenho superior e maior capacidade de generalização em comparação à regressão censurada de vetores de suporte na predição da mortalidade de árvores em fragmentos de Mata Atlântica (Figura 2). O desempenho inferior da SVRC em relação a RSF também foi verificado em alguns estudos nas ciências médicas (Suresh et al., 2022; Li et al., 2024).

A predição do modelo RSF foi melhor do que o SVCR de acordo com o *C-Index* e o *Brier-Score* (Figura 2). Estes valores sugerem que o modelo RSF pode classificar com precisão eventos de sobrevivência e mortalidade de árvores em fragmentos de Mata Atlântica. O *C-Index* mede a precisão de predição do modelo com base na taxa de verdadeiro positivo e na taxa de falso positivo. Isso indica a precisão do modelo na classificação de sobrevivência versus mortalidade. O *Brier-Score* mede a precisão de predição do modelo com base em valores absolutos que indicam a diferença entre as probabilidades de sobrevivência/mortalidade previstas e reais. Neste trabalho, o modelo RSF demonstrou maior precisão de predição em comparação com o SVCR tanto na *C-index* quanto no *Brier-Score*.

Alguns estudos com aplicações ecológicas têm demonstrado o desempenho superior do método de *random forest* em comparação a outras técnicas de aprendizado de máquina (Aviña-Hernández et al., 2023; Chiaverini et al., 2023; da Silva et al., 2023; Dang et al., 2019; Rocha et al., 2024; Torres e Qiu, 2014). A capacidade deste algoritmo em lidar com relações complexas, robustez em relação a outliers, dados ausentes, avaliação de importância de variáveis e prevenção de overfitting contribuem para sua eficácia (Reise et al., 2019). Como um método não paramétrico, ele se beneficia de sua capacidade de levar em consideração a variabilidade dos dados e relacionamentos não lineares (Freeman et al., 2015; Mascaro et al., 2014).

Ao reduzir o número de covariáveis preditoras no modelo RSF não reduziu consideravelmente sua precisão de predição (Figura 4 e 5). Ao utilizar apenas três covariáveis: Área basal, Grupo ecológico e déficit hídrico de 3 anos, o modelo 5 atingiu um nível semelhante de predição em comparação ao modelo completo e os outros modelos reduzidos com um maior número de covariáveis. O resultado implica que a aplicação do RSF pode possivelmente reduzir o custo de aquisição e processamento de dados na prática, mantendo um alto nível de precisão na predição. Notavelmente, reduzir ainda mais o número de preditoras no modelo levou a um declínio na precisão de predição. Modelos que dependem de menos de três covariáveis (modelo 6) exibiram capacidades preditivas obviamente mais baixas, o que implica que as três

covariáveis mais importantes contêm informações críticas para prever com precisão a sobrevivência e mortalidade nesses fragmentos de Mata Atlântica.

A covariável área basal, representa a área transversal do tronco de uma árvore na altura do peito. Covariáveis de tamanho de árvore (área basal) geralmente são componentes essenciais da modelagem de mortalidade porque expressam a competição entre indivíduos (Rocha et al., 2024). Em fragmentos com alta área basal, a competição por recursos, como água, luz e nutrientes aumenta (Ruiz-Benito et al., 2013, Young et al., 2017), resultando em auto desbaste (Pilet et al., 2018), aumentando a taxa de mortalidade. Na presença de eventos de seca, a competição entre os indivíduos, pode ser intensificar devido à escassez de recursos (McDowell, 2018) e árvores com áreas basais maiores podem apresentar taxas de mortalidade menores, por possuir maior acesso aos recursos (Rocha et al., 2020).

Este estudo reforça a influência significativa do déficit hídrico na mortalidade das árvores na Mata Atlântica. Nossos achados estão alinhados com pesquisas anteriores conduzidas nesse ecossistema (Rocha et al., 2018; 2020; Torres et al., 2023), que indicam que eventos de seca têm impacto direto na dinâmica florestal.

Muitas espécies arbóreas da Mata Atlântica apresentam altas taxas de condutância hidráulica e fotossíntese, tornando-se mais vulneráveis ao estresse fisiológico induzido pela seca, o que pode resultar em um aumento na taxa de mortalidade (Lohbeck et al., 2013; Ouédraogo et al., 2013; Uriarte et al., 2016). Principalmente para espécies pioneiras, que apresentam rápido crescimento e durante a sucessão são substituídas por espécies mais adaptadas ao ambiente sombreado e úmido (Ruschel et al., 2009). A alta mortalidade de pioneiras pode afetar a composição e diversidade de espécies nas florestas tropicais (Aleixo et al., 2019, Silk, 2004), alterando a dinâmica sucessional (Uriarte et al., 2016).

Em resumo, este estudo demonstra que, o método de análise de sobrevivência baseado em *random survival forests* (RSF) foi uma ferramenta eficiente e importante para modelar a sobrevivência e mortalidade de árvores em fragmentos florestais na Mata Atlântica. Esta técnica pode ajudar no desenvolvimento de estratégias de manejo mais eficientes voltadas para a conservação e redução dos efeitos ambientais nesse importante bioma. Sua capacidade de analisar dados complexos, capturar relações não lineares e fornecer previsões precisas, pode dar suporte aos processos de tomada de decisão para a compreensão e conservação deste ecossistema único e ameaçado.

## CONCLUSÃO

A metodologia *random survival forests* foi eficiente avaliar a sobrevivência e mortalidade individual de árvores nos fragmentos de Mata Atlântica. Esta técnica pode ser uma ferramenta útil na tomada de decisões dos gestores florestais.

## REFERENCIAS

- Aleixo, I., Norris, D., Hemerik, L., Barbosa, A., Prata, E., Costa, F., Poorter, L., 2019. Amazonian rainforest tree mortality driven by climate and functional traits. *Nat. Clim. Chang.* 9, 384–388. <https://doi.org/10.1038/s41558-019-0458-0>
- Aviña-Hernández, J., Ramírez-Vargas, M., Roque-Sosa, F., Martínez-Rincón, R.O., 2023. Predictive performance of random forest on the identification of mangrove species in arid environments. *Ecological Informatics* 75, 102040. <https://doi.org/10.1016/j.ecoinf.2023.102040>
- Brandl, S., Paul, C., Knoke, T., Falk, W., 2020. The influence of climate and management on survival probability for Germany's most important tree species. *Forest Ecology and Management* 458, 117652. <https://doi.org/10.1016/j.foreco.2019.117652>
- Campoe, O.C., Iannelli, C., Stape, J.L., Cook, R.L., Mendes, J.C.T., Vivian, R., 2014. Atlantic forest tree species responses to silvicultural practices in a degraded pasture restoration plantation: From leaf physiology to survival and initial growth. *Forest Ecology and Management* 313, 233–242. <https://doi.org/10.1016/j.foreco.2013.11.016>
- 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-Yrizar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Péliissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology* 20, 3177–3190. <https://doi.org/10.1111/gcb.12629>
- Chen, Y., Jia, Z., Mercola, D., Xie, X., 2013. A Gradient Boosting Algorithm for Survival Analysis via Direct Optimization of Concordance Index. *Computational and Mathematical Methods in Medicine* 2013, 873595. <https://doi.org/10.1155/2013/873595>
- Chiaverini, L., Macdonald, D.W., Hearn, A.J., Kaszta, Ž., Ash, E., Bothwell, H.M., Can, Ö.E., Channa, P., Clements, G.R., Haidir, I.A., Kyaw, P.P., Moore, J.H., Rasphone, A., Tan, C.K.W., Cushman, S.A., 2023. Not seeing the forest for the trees: Generalised linear model out-performs random forest in species distribution modelling for Southeast Asian felids. *Ecological Informatics* 75, 102026. <https://doi.org/10.1016/j.ecoinf.2023.102026>
- Collett, D., 2023. *Modelling Survival Data in Medical Research*, 4th ed. Chapman and Hall/CRC, Boca Raton. <https://doi.org/10.1201/9781003282525>
- Colosimo, E.A., Giolo, S.R., 2024. *Análise de sobrevivência aplicada*, 2nd ed. Editora Edgard Blucher, São Paulo, SP.
- da Silva, S.D.P., Eugenio, F.C., Fantinel, R.A., Amaral, L. de P., dos Santos, A.R., Mallmann, C.L., dos Santos, F.D., Pereira, R.S., Ruoso, R., 2023. Modeling and detection of invasive trees using UAV image and machine learning in a subtropical forest in Brazil.

- Ecological Informatics 74, 101989. <https://doi.org/10.1016/j.ecoinf.2023.101989>
- Dang, A.T.N., Nandy, S., Srinet, R., Luong, N.V., Ghosh, S., Senthil Kumar, A., 2019. Forest aboveground biomass estimation using machine learning regression algorithm in Yok Don National Park, Vietnam. *Ecological Informatics* 50, 24–32. <https://doi.org/10.1016/j.ecoinf.2018.12.010>
- Department of Statistics and Mathematical Sciences, Kwara State University, Malete, Nigeria, Dauda, K.A., 2022. Optimal Tuning of Random Survival Forest Hyperparameter with an Application to Liver Disease. *MJMS* 29, 67–76. <https://doi.org/10.21315/mjms2022.29.6.7>
- Fan, Z., Kabrick, J.M., Shifley, S.R., 2006. Classification and regression tree based survival analysis in oak-dominated forests of Missouri's Ozark highlands. *Can. J. For. Res.* 36, 1740–1748. <https://doi.org/10.1139/x06-068>
- Fouodo, C.J.K., 2018. *survivalsvm: Survival Support Vector Analysis*.
- Freeman, E.A., Moisen, G.G., Coulston, J.W., Wilson, B.T., 2016. Random forests and stochastic gradient boosting for predicting tree canopy cover: comparing tuning processes and model performance. *Can. J. For. Res.* 46, 323–339. <https://doi.org/10.1139/cjfr-2014-0562>
- Gandolfi, S., Leitão Filho, H. D. F., Bezerra, C. L. F., 1995. Levantamento florístico e caráter sucessional das espécies arbustivo-arbóreas de uma floresta mesófila semidecídua no município de Guarulhos, SP. *Revista brasileira de biologia*, 55(4), 753-767.
- Hao, L., Kim, J., Kwon, S., Ha, I.D., 2021. Deep Learning-Based Survival Analysis for High-Dimensional Survival Data. *Mathematics* 9, 1244. <https://doi.org/10.3390/math9111244>
- Huang, Y., Li, J., Li, M., Aparasu, R.R., 2023. Application of machine learning in predicting survival outcomes involving real-world data: a scoping review. *BMC Medical Research Methodology* 23, 268. <https://doi.org/10.1186/s12874-023-02078-1>
- Ibge, 2012. *Manual Técnico da Vegetação Brasileira (2a. ed.)*, IBGE, Rio de Janeiro.
- Ishwaran, H., Gerds, T.A., Kogalur, U.B., Moore, R.D., Gange, S.J., Lau, B.M., 2014. Random survival forests for competing risks. *Biostatistics* 15, 757–773. <https://doi.org/10.1093/biostatistics/kxu010>
- Ishwaran, H., Kogalur, U.B., 2025. *randomForestSRC: Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*.
- Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. *The Annals of Applied Statistics* 2, 841–860. <https://doi.org/10.1214/08-AOAS169>
- King, S.L., Bennett, K.P., List, S., 2000. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. *Computers and Electronics in Agriculture* 27, 401–406. [https://doi.org/10.1016/S0168-1699\(00\)00106-X](https://doi.org/10.1016/S0168-1699(00)00106-X)
- Li, Y., Yang, X., Ma, L., 2024. Prognostic prediction for HER2-low breast cancer patients using a novel machine learning model. *BMC Cancer* 24, 1474. <https://doi.org/10.1186/s12885-024-13241-2>
- Lohbeck, M., Poorter, L., Lebrija-Trejos, E., Martínez-Ramos, M., Meave, J.A., Paz, H., Pérez-García, E.A., Romero-Pérez, I.E., Tauro, A., Bongers, F., 2013. Successional changes in functional composition contrast for dry and wet tropical forest. *Ecology* 94, 1211–

1216. <https://doi.org/10.1890/12-1850.1>
- Longpre, T.W., Morris, D.M., 2012. Environmental Drivers of Succession in Jack Pine Stands of Boreal Ontario: An Application of Survival Analysis. *Northern Journal of Applied Forestry* 29, 81–92. <https://doi.org/10.5849/njaf.11-016>
- Lutz, J.A., van Wagtenonk, J.W., Franklin, J.F., 2010. Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography* 37, 936–950. <https://doi.org/10.1111/j.1365-2699.2009.02268.x>
- Manso, R., Morneau, F., Ningre, F., Fortin, M., 2015. Incorporating stochasticity from extreme climatic events and multi-species competition relationships into single-tree mortality models. *Forest Ecology and Management* 354, 243–253. <https://doi.org/10.1016/j.foreco.2015.06.008>
- Maringer, J., Stelzer, A.-S., Paul, C., Albrecht, A.T., 2021. Ninety-five years of observed disturbance-based tree mortality modeled with climate-sensitive accelerated failure time models. *Eur J Forest Res* 140, 255–272. <https://doi.org/10.1007/s10342-020-01328-x>
- Mascaro, J., Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Martin, R.E., Anderson, C., Higgins, M., Chadwick, K.D., 2014. A Tale of Two “Forests”: Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. *PLOS ONE* 9, e85993. <https://doi.org/10.1371/journal.pone.0085993>
- McDowell, N.G., 2018. Deriving pattern from complexity in the processes underlying tropical forest drought impacts. *New Phytologist* 219, 841–844. <https://doi.org/10.1111/nph.15341>
- Neumann, M., Mues, V., Moreno, A., Hasenauer, H., Seidl, R., 2017. Climate variability drives recent tree mortality in Europe. *Global Change Biology* 23, 4788–4797. <https://doi.org/10.1111/gcb.13724>
- Ouédraogo, D.-Y., Mortier, F., Gourlet-Fleury, S., Freycon, V., Picard, N., 2013. Slow-growing species cope best with drought: evidence from long-term measurements in a tropical semi-deciduous moist forest of Central Africa. *Journal of Ecology* 101, 1459–1470. <https://doi.org/10.1111/1365-2745.12165>
- Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A Large and Persistent Carbon Sink in the World’s Forests. *Science* 333, 988–993. <https://doi.org/10.1126/science.1201609>
- Paul, C., Brandl, S., Friedrich, S., Falk, W., Härtl, F., Knoke, T., 2019. Climate change and mixed forests: how do altered survival probabilities impact economically desirable species proportions of Norway spruce and European beech? *Annals of Forest Science* 76, 1–15. <https://doi.org/10.1007/s13595-018-0793-8>
- Pickett, K.L., Suresh, K., Campbell, K.R., Davis, S., Juarez-Colunga, E., 2021. Random survival forests for dynamic predictions of a time-to-event outcome using a longitudinal biomarker. *BMC Medical Research Methodology* 21, 216. <https://doi.org/10.1186/s12874-021-01375-x>
- Pillet, M., Joetzer, E., Belmin, C., Chave, J., Ciais, P., Dourdain, A., Evans, M., Hérault, B., Luysaert, S., Poulter, B., 2018. Disentangling competitive vs. climatic drivers of tropical forest mortality. *Journal of Ecology* 106, 1165–1179. <https://doi.org/10.1111/1365-2745.12876>

- 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-Velázquez, J., Ruschel, A.R., Salgado-Negret, B., Schiavini, I., Swenson, N.G., Tenorio, E.A., Thompson, J., Toledo, M., Uriarte, M., Hout, P. van der, Zimmerman, J.K., Peña-Claros, M., 2017. Biodiversity and climate determine the functioning of Neotropical forests. *Global Ecology and Biogeography* 26, 1423–1434. <https://doi.org/10.1111/geb.12668>
- Qiu, X., Gao, J., Yang, J., Hu, J., Hu, W., Kong, L., Lu, J.J., 2020. A Comparison Study of Machine Learning (Random Survival Forest) and Classic Statistic (Cox Proportional Hazards) for Predicting Progression in High-Grade Glioma after Proton and Carbon Ion Radiotherapy. *Front. Oncol.* 10. <https://doi.org/10.3389/fonc.2020.551420>
- R: The R Project for Statistical Computing [WWW Document], n.d. URL <https://www.r-project.org/> (accessed 2.25.25).
- Reis, L.P., de Souza, A.L., dos Reis, P.C.M., Mazzei, L., Soares, C.P.B., Miquelino Eleto Torres, C.M., da Silva, L.F., Ruschel, A.R., Rêgo, L.J.S., Leite, H.G., 2018. Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. *Ecological Engineering* 112, 140–147. <https://doi.org/10.1016/j.ecoleng.2017.12.014>
- Reise, J., Kukulka, F., Flade, M., Winter, S., 2019. Characterising the richness and diversity of forest bird species using National Forest Inventory data in Germany. *Forest Ecology and Management* 432, 799–811. <https://doi.org/10.1016/j.foreco.2018.10.012>
- Rocha, S.J.S.S. da, 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. da, Reis, L.P., Zanoncio, J.C., 2018. Artificial neural networks: Modeling tree survival and mortality in the Atlantic Forest biome in Brazil. *Science of The Total Environment* 645, 655–661. <https://doi.org/10.1016/j.scitotenv.2018.07.123>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Schettini, B.L.S., Jacovine, L.A.G., Leite, H.G., Gelcer, E.M., Reis, L.P., Neves, K.M., Comini, I.B., Silva, L.F. da, 2020. Drought effects on carbon dynamics of trees in a secondary Atlantic Forest. *Forest Ecology and Management* 465, 118097. <https://doi.org/10.1016/j.foreco.2020.118097>
- Rocha, S.J.S.S. da, Torres, C.M.M.E., Villanova, P.H., Tavares Júnior, I. da S., Rufino, M.P.M.X., Romero, F.M.B., Jacovine, L.A.G., de Moraes Junior, V.T.M., França, L.C. de 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., 2024. Machine learning methods: Modeling net growth in the Atlantic Forest of Brazil. *Ecological Informatics* 81, 102564. <https://doi.org/10.1016/j.ecoinf.2024.102564>
- Ruiz-Benito, P., Lines, E.R., Gómez-Aparicio, L., Zavala, M.A., Coomes, D.A., 2013. Patterns and Drivers of Tree Mortality in Iberian Forests: Climatic Effects Are Modified by Competition. *PLOS ONE* 8, e56843. <https://doi.org/10.1371/journal.pone.0056843>
- Ruschel, A.R., Mantovani, M., Reis, M.S. dos, Nodari, R.O., 2009. Caracterização e dinâmica de duas fases sucessionais em floresta secundária da mata atlântica. *Rev. Árvore* 33, 101–115. <https://doi.org/10.1590/S0100-67622009000100011>
- Senevirathne, D.M., Yang, S.-I., Brandeis, C., Hodges, D.G., 2024. Predicting time-to-harvest in mixed-species forests using a random survival forest algorithm. *Forest Ecosystems* 11, 100236. <https://doi.org/10.1016/j.fecs.2024.100236>

- Shivaswamy, P.K., Chu, W., Jansche, M., 2007. A Support Vector Approach to Censored Targets, in: Seventh IEEE International Conference on Data Mining (ICDM 2007). Presented at the Seventh IEEE International Conference on Data Mining (ICDM 2007), pp. 655–660. <https://doi.org/10.1109/ICDM.2007.93>
- Slik, J.W.F., 2004. El Niño droughts and their effects on tree species composition and diversity in tropical rain forests. *Oecologia* 141, 114–120. <https://doi.org/10.1007/s00442-004-1635-y>
- SOS Mata Atlântica, 2024. Atlas da Mata Atlântica. <https://www.sosma.org.br/iniciativas/atlas-da-mata-atlantica>.
- Souza, C.R., Maia, V.A., Aguiar-Campos, N. de, Santos, A.B.M., Rodrigues, A.F., Farrapo, C.L., Gianasi, F.M., Paula, G.G.P. de, Fagundes, N.C.A., Silva, W.B., Santos, R.M., 2021. Long-term ecological trends of small secondary forests of the atlantic forest hotspot: A 30-year study case. *Forest Ecology and Management* 489, 119043. <https://doi.org/10.1016/j.foreco.2021.119043>
- Suresh, K., Severn, C., Ghosh, D., 2022. Survival prediction models: an introduction to discrete-time modeling. *BMC Medical Research Methodology* 22, 207. <https://doi.org/10.1186/s12874-022-01679-6>
- Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Müller, M.S., Rödig, E., Wiegand, T., Huth, A., 2018. Global patterns of tropical forest fragmentation. *Nature* 554, 519–522. <https://doi.org/10.1038/nature25508>
- Teixeira, P.C., Donagemma, G.K., Fontana, A., Teixeira, W.G. (Eds.), 2017. Manual de métodos de análise de solo, 3ª. ed. Embrapa, Brasília, Brazil.
- Thapa, R., Burkhart, H.E., Li, J., Hong, Y., 2016. Modeling Clustered Survival Times of Loblolly Pine with Time-dependent Covariates and Shared Frailties. *JABES* 21, 92–110. <https://doi.org/10.1007/s13253-015-0217-2>
- Thiele, J.C., Nuske, R.S., Ahrends, B., Panferov, O., Albert, M., Staupendahl, K., Junghans, U., Jansen, M., Saborowski, J., 2017. Climate change impact assessment—A simulation experiment with Norway spruce for a forest district in Central Europe. *Ecological Modelling* 346, 30–47. <https://doi.org/10.1016/j.ecolmodel.2016.11.013>
- Torres, C.M.M.E., Medina-Vega, J.A., Rocha, S.J.S.S. da, Costa, W. da S., Soares, C.P.B., Souza, A.L. de, Jacovine, L.A.G., Lana, J.M. de, Peña-Claros, M., 2023a. Drivers of tree demographic processes in forest fragments of the Brazilian Atlantic forest. *Forest Ecology and Management* 534, 120893. <https://doi.org/10.1016/j.foreco.2023.120893>
- Torres, C.M.M.E., Medina-Vega, J.A., Rocha, S.J.S.S. da, Costa, W. da S., Soares, C.P.B., Souza, A.L. de, Jacovine, L.A.G., Lana, J.M. de, Peña-Claros, M., 2023b. Drivers of tree demographic processes in forest fragments of the Brazilian Atlantic forest. *Forest Ecology and Management* 534, 120893. <https://doi.org/10.1016/j.foreco.2023.120893>
- Torres, M., Qiu, G., 2014. Automatic habitat classification using image analysis and random forest. *Ecological Informatics, Special Issue on Multimedia in Ecology and Environment* 23, 126–136. <https://doi.org/10.1016/j.ecoinf.2013.08.002>
- Uriarte, M., Schwartz, N., Powers, J.S., Marín-Spiotta, E., Liao, W., Werden, L.K., 2016a. Impacts of climate variability on tree demography in second growth tropical forests: the importance of regional context for predicting successional trajectories. *Biotropica* 48, 780–797. <https://doi.org/10.1111/btp.12380>
- Uriarte, M., Schwartz, N., Powers, J.S., Marín-Spiotta, E., Liao, W., Werden, L.K., 2016b.

- Impacts of climate variability on tree demography in second growth tropical forests: the importance of regional context for predicting successional trajectories. *Biotropica* 48, 780–797. <https://doi.org/10.1111/btp.12380>
- Uzoh, F.C.C., Mori, S.R., 2012. Applying survival analysis to managed even-aged stands of ponderosa pine for assessment of tree mortality in the western United States. *Forest Ecology and Management* 285, 101–122. <https://doi.org/10.1016/j.foreco.2012.08.006>
- Van Belle, V., Pelckmans, K., Van Huffel, S., Suykens, J.A.K., 2011. Improved performance on high-dimensional survival data by application of Survival-SVM. *Bioinformatics* 27, 87–94. <https://doi.org/10.1093/bioinformatics/btq617>
- van Breugel, M., Craven, D., Lai, H.R., Baillon, M., Turner, B.L., Hall, J.S., 2019. Soil nutrients and dispersal limitation shape compositional variation in secondary tropical forests across multiple scales. *Journal of Ecology* 107, 566–581. <https://doi.org/10.1111/1365-2745.13126>
- Vieira, G.C., de Mendonça, A.R., da Silva, G.F., Zanetti, S.S., da Silva, M.M., dos Santos, A.R., 2018. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of The Total Environment* 619–620, 1473–1481. <https://doi.org/10.1016/j.scitotenv.2017.11.138>
- Young, D.J.N., Stevens, J.T., Earles, J.M., Moore, J., Ellis, A., Jirka, A.L., Latimer, A.M., 2017. Long-term climate and competition explain forest mortality patterns under extreme drought. *Ecology Letters* 20, 78–86. <https://doi.org/10.1111/ele.12711>

## **CONCLUSÕES GERAIS**

A análise de sobrevivência foi eficiente para a modelagem da sobrevivência e mortalidade e recrutamento em fragmentos da Mata Atlântica. As covariáveis que influenciam esses processos demográficos variaram de acordo com o grupo ecológico das árvores.

As covariáveis ambientais (clima e características do solo), distância até a borda e área basal, apresentaram diferentes graus de influência, positiva ou negativa, entre espécies pioneiras, secundárias iniciais e secundárias tardias. Essa variação permitiu a identificação dos grupos de espécies mais suscetíveis aos impactos de cada uma dessas covariáveis.

Os métodos aplicados neste estudo podem servir como ferramentas estratégicas para a tomada de decisões por gestores florestais, contribuindo para a conservação e manejo sustentável da Mata Atlântica.