

PREDIÇÃO DE RISCO AMBIENTAL EM MUNICÍPIOS DA ZONA DE CONVERGÊNCIA SOCIOECOLÓGICA CERRADO-AMAZÔNIA UTILIZANDO MODELOS SUPERVISIONADOS DE MACHINE LEARNING

Jeferson Silva Araújo¹
Giovanni de Araújo Boggione²
Universidade Evangélica de Goiás – UniEvangélica¹²

RESUMO

A região de transição entre os biomas Cerrado e Amazônia enfrenta pressões ambientais crescentes, como desmatamento acelerado, queimadas recorrentes e perda de biodiversidade, intensificados pela expansão agropecuária e pela conversão de terras naturais em pastagens e lavouras. A identificação de áreas sob maior risco ambiental é essencial para subsidiar estratégias de conservação e formulação de políticas públicas. Este estudo tem como objetivo desenvolver e avaliar modelos supervisionados de *machine learning* capazes de prever, em escala municipal, o risco de desmatamento, queimadas e perda de cobertura florestal. A metodologia proposta integra bases de dados ambientais e socioeconômicas em um sistema de informação geográfica, seguido pelo treinamento de algoritmos Random Forest, XGBoost e Support Vector Machine, avaliados por métricas como AUC-ROC, *precision-recall*, *f1-score* e matriz de confusão. Espera-se que os resultados, espacializados em mapas temáticos de risco, auxiliem na identificação de áreas prioritárias para conservação e mitigação de impactos ambientais.

Palavras-chave: Sensoriamento remoto; variáveis socioeconômicas; geoprocessamento; modelagem preditiva.

INTRODUÇÃO

A interface entre os biomas Cerrado e Amazônia é marcada por intensa dinâmica de uso e cobertura da terra, onde o desmatamento, as queimadas e a expansão agropecuária configuram-se como os principais vetores de pressão ambiental. A transformação da paisagem compromete a biodiversidade, fragmenta habitats e reduz a provisão de serviços ecossistêmicos essenciais.

A integração de dados de sensoriamento remoto e variáveis socioeconômicas em um ambiente de banco de dados geoespacial, visando a construção de modelos supervisionados de *machine learning* para predição do risco ambiental nos 255 municípios da zona de convergência socioecológica Cerrado-Amazônia, conforme Figura 1, distribuídos entre 5 estados Maranhão, Mato Grosso, Pará, Rondônia e Tocantins, onde o período estudado abrange os anos de 1985 até 2023, apresentando uma série histórica sobre as alterações nos dois biomas envolvidos em seus municípios, alimentando os modelos de predição de uma forma eficiente.



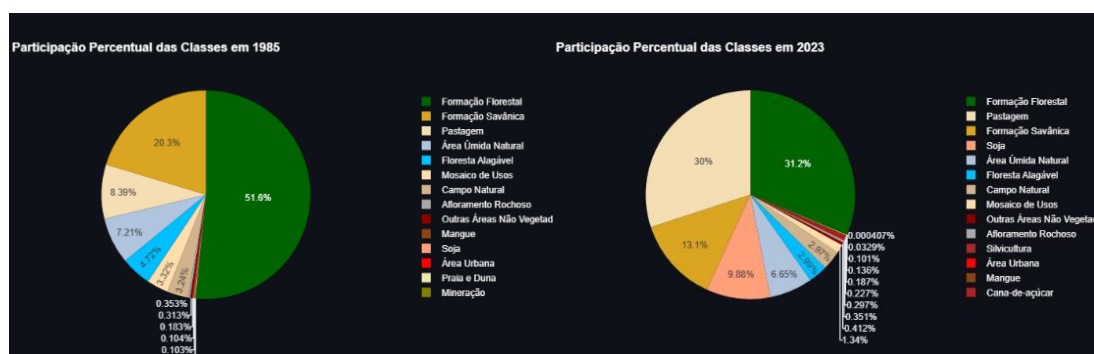
Fonte - Autor

Apesar da disponibilidade de dados ambientais em larga escala, ainda há lacunas quanto à integração de variáveis ambientais e socioeconômicas para predição espacializada de risco ambiental em escala municipal na zona de convergência socioecológica Amazônia-Cerrado. Assim, este estudo tem como objetivo desenvolver e avaliar modelos supervisionados de machine learning capazes de prever o risco de desmatamento, queimadas e perda de cobertura vegetal nos municípios da região entre os anos de 1985 e 2023, gerando mapas temáticos de risco que possam subsidiar estratégias de conservação e políticas baseadas em evidências.

MATERIAIS E MÉTODOS

No contexto das alterações de uso do solo na região da ZCS Cerrado-Amazônia, as ferramentas computacionais baseadas em *machine learning* serão amplamente utilizadas para integrar diferentes fontes de dados e gerar modelos preditivos robustos que permitem compreender a dinâmica socioambiental e prever cenários futuros de risco ambiental (BREIMAN, 2001; CHEN; GUESTIN, 2016; CARRASCO et al., 2019; GRECCI et al., 2015).

Figura 2 – Participação das Classes de Cobertura do Solo anos 1985 e 2023

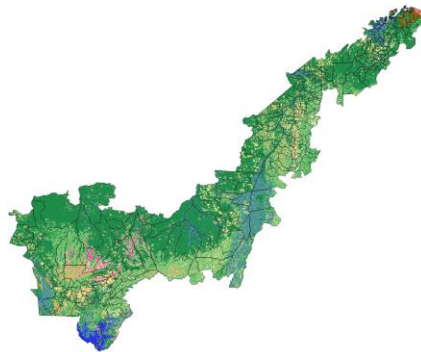


Fonte - Autor

A escolha dessa abordagem decorre do fato de que o Cerrado e a Amazônia apresentam intensa dinâmica de uso e cobertura da terra, sendo o desmatamento e a conversão de áreas naturais para a agricultura fatores centrais de transformação ambiental nas últimas décadas (GRECCI et al., 2015), conforme mostrado na Figura 2.

Para capturar tais dinâmicas, serão utilizadas séries históricas de dados provenientes de satélites Landsat e bases como MapBiomas e PRODES. A adoção de técnicas de detecção de mudanças já se mostrou eficaz para avaliar a perda de vegetação nativa em larga escala, ao evidenciar que mais de 12 milhões de hectares de vegetação natural do Cerrado foram convertidos entre 1990 e 2010 (GRECCI et al., 2015). A Figura 3 a seguir apresenta o mapa da zona de convergência socioecológica Cerrado-Amazônia no ano de 1985, mostrando a situação com grande participação ainda de formações florestais e coberturas campestres originais.

Figura 3 - Cobertura do Solo dos municípios da ZCS Amazônia – Cerrado no ano de 1985



Fonte – Autor

Já a figura 4 a seguir apresenta a situação de cobertura e uso do solo da região ZCS Cerrado-Amazônia no ano de 2023, demonstrando o grande avanço de pastagens e culturas como soja principalmente.

Figura 4 – Cobertura do Solo dos municípios da ZCS Amazônia – Cerrado no ano de 2023



Fonte - Autor

Conforme as Figuras 3 e Figura 4, pode-se ver que a ZCS sofreu grandes alterações em sua distribuição de uso e cobertura do solo. Os algoritmos de *machine learning* selecionados — Random Forest, XGBoost e Support Vector Machine — foram escolhidos por sua robustez em lidar com dados heterogêneos e de alta dimensionalidade. O Random Forest, segundo Breiman (2001), é eficiente na construção de múltiplas árvores de decisão, reduzindo a variância por meio de combinações de classificadores e apresentando bom desempenho em dados ruidosos. De forma complementar, o XGBoost utiliza uma estratégia de otimização baseada em gradiente, oferecendo maior precisão e velocidade em relação a outros métodos de *boosting* (CHEN; GUESTIN, 2016).

O processo metodológico passará por revisões bibliográficas, coletas de dados de sistemas abertos de dados georreferenciados e incluirá etapas de pré-processamento de variáveis, treinamento dos modelos supervisionados e avaliação do desempenho por meio de métricas consagradas, como AUC-ROC, *precision-recall*, *f1-score* e matriz de confusão. Essa estrutura se alinha a estudos recentes que utilizam algoritmos de aprendizado de máquina para reconstruir séries históricas de uso da terra e identificar trajetórias de transformação ambiental em regiões tropicais (CARRASCO et al., 2019). Os dados para alimentar os modelos foram coletados por sistemas abertos como MapBiomas para análise de cobertura do solo, dados do IBGE para verificação de alterações socioeconômicas

A validação dos modelos será conduzida com partição dos dados em conjuntos de treino e teste, além de validação cruzada *k-fold* para assegurar robustez estatística.

Resultados

O estudo sobre a zona de convergência socioecológica Cerrado–Amazônia entre 1985 e 2023 revelou uma perda significativa de 35% da vegetação nativa, totalizando mais de 25 milhões de hectares. Essa perda foi impulsionada, principalmente, pela expansão da agricultura e da urbanização. As atividades agrícolas, em particular, apresentaram um crescimento expressivo, com a soja e "outras lavouras" somando um acréscimo de quase 28 milhões de hectares. A expansão urbana também se destacou, com um aumento de mais de 121%. No âmbito municipal, cidades como Juara e Vila Bela da Santíssima Trindade registraram as maiores perdas de vegetação nativa, enquanto o aumento de

pastagens foi mais notável em Poconé e Cáceres. Já o crescimento urbano se concentrou em São Luís e Sinop.

Conclusão

Os resultados para a zona de convergência socioecológica Cerrado–Amazônia revelam uma dinâmica persistente de conversão do uso do solo, marcada pela redução de remanescentes de vegetação nativa, pela expansão urbana e pelo avanço da agricultura, com destaque para cultivos como soja e outras lavouras, em paralelo a uma substituição gradual de áreas de pastagem por sistemas agrícolas mais intensivos. Essa reorganização do território evidencia padrões espaciais heterogêneos, com municípios que se configuram como hotspots de pressão, e sustenta a utilização de modelos preditivos para antecipar cenários de perda adicional de cobertura natural, priorizar áreas de risco e orientar estratégias de conservação e produção sustentável. Apesar das incertezas inerentes a dados temáticos e variações locais, a consistência dos padrões observados oferece base robusta para decisões de ordenamento territorial e para políticas públicas focalizadas na convergência entre os biomas.

Referências Bibliográficas

BREIMAN, L. Random Forests. *Machine Learning*, v. 45, p. 5-32, 2001.

CARRASCO, L. et al. Reconstructing three decades of land use and land cover changes in tropical regions. *Remote Sensing of Environment*, v. 232, p. 111-120, 2019.

CHEN, T.; GUESTRIN, C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785-794, 2016.

GRECCI, R. C. et al. Assessing land cover changes in the Brazilian Cerrado between 1990 and 2010 using a remote sensing sampling approach. *Anais XVII Simpósio Brasileiro de Sensoriamento Remoto*, João Pessoa-PB, 2015.