XXIV SIMPÓSIO BRASILEIRO DE RECURSOS HIDRÍCOS

ACURÁCIA DOS DADOS DE COBERTURA DO SOLO DO PROJETO MAPBIOMAS: ESTUDO DE CASO DE UMA PEQUENA BACIA HIDROGRÁFICA DE CABECEIRA

Phelipe da Silva Anjinho[1]; Talyson de Melo Bolleli[2]; Lucas Olegário Bueno[3]; Denise Parizotto[4]; Allita Rezende dos Santos[5]; Camila Bermond Ruezzene[6]; Cássia Cortes Valadão[7]; Mariana Abibi Guimarães Araujo Barbosa[8]; Rhayane Carvalho Roque[9] & Frederico Fábio Mauad[10]

Palavras-Chave – Índice Kappa; MapBiomas; Uso e Ocupação do solo.

1. INTRODUÇÃO

O padrão espacial de uso e ocupação do solo de um território influencia os principais ciclos naturais do planeta, provocando mudanças nos fluxos de água, energia, carbono e nutrientes, tanto em escala local como global (Sun *et al*., 2017). Essas mudanças desregulam o funcionamento dos ecossistemas e afetam suas capacidades de fornecer serviços ecossistêmicos. Por modificar os ciclos de água, energia e matéria, dados de uso e ocupação do solo são importantes para o desenvolvimento de estudos ambientais. Pesquisas sobre clima, modelagem hidrossedimentológica, ecossistêmica, ecologia, biodiversidade, poluição hídrica, planejamento e gestão ambiental são exemplos de estudos que demandam dados de cobertura do solo (Ellison *et al*., 2017; Bucak *et al*., 2018; Ferreira *et al*., 2019; Fendrich *et al*., 2020; Anjinho *et al*., 2021). A qualidade dos resultados dessas pesquisas é fortemente dependente da acurácia dos dados de uso e ocupação do solo.

No Brasil, dados de cobertura do solo são fornecidos pelo Projeto de Mapeamento Anual da Cobertura e Uso do Solo do Brasil (MapBiomas). O MapBiomas é fruto de uma rede colaborativa que conta com a participação de múltiplas instituições e especialistas de diversas áreas do conhecimento. Foi criado em 2015 com o propósito de monitorar e entender a dinâmica de uso do solo no Brasil. O projeto está na sua 5ª coleção e disponibiliza dados anuais de cobertura do solo, de 1985 a 2019. Todo o processo de mapeamento é realizado em nuvem. Resumidamente, o método utilizado pelo MapBiomas inclui: i) classificação supervisionada de pixels de imagens de satélite Landsat (30 metros de resolução espacial), realizada por meio de algoritmos de aprendizagem de máquina (*machine learning*) na plataforma Google Earth Engine; ii) aplicação de filtros temporais e espaciais nos dados gerados pela classificação, iii) integração dos dados produzidos; e iv) validação dos dados por meio de análise comparativa com ~75.000 amostras distribuídas por todo território brasileiro (Souza; Azevedo, 2017).

Embora procedimentos de validação sejam realizados pelo próprio MapBiomas, pesquisas ainda são necessárias para avaliar a acurácia desses dados, tendo em vista a grande extensão do território brasileiro e suas peculiaridades locais e regionais, o que dificulta a validação precisa dos dados. Muitos estudos têm utilizado dados de uso e ocupação do solo do MapBiomas (Fendrich *et al*., 2020; Arruda *et al*., 2021; Rosa *et al*., 2021), porém poucos têm investigado a acurácia desses dados em escala local. Nesse contexto, esse trabalho teve como objetivo avaliar a acurácia dos dados de uso e ocupação do solo gerados pelo MapBiomas, por meio da comparação com os dados de uso e ocupação do solo obtidos pela digitalização manual de imagens geradas pelo satélite Sentinel-2.

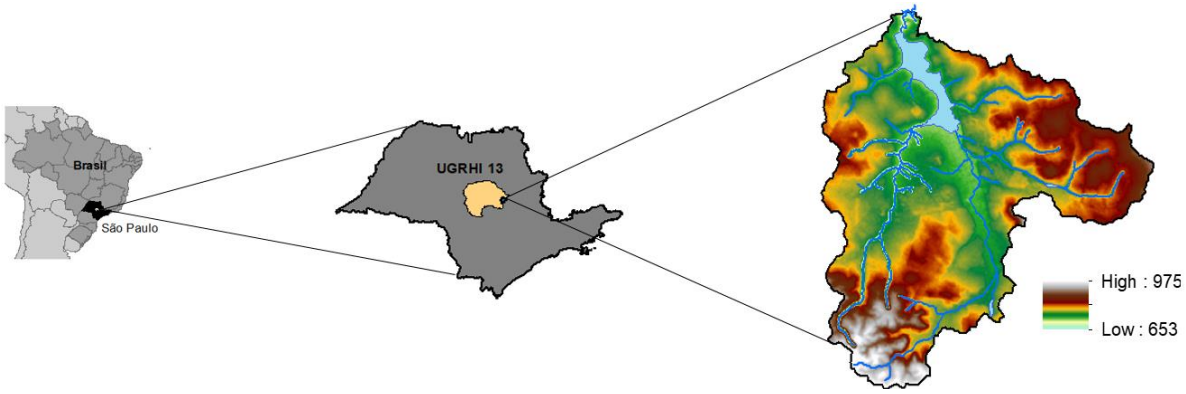
[1]) [1 a 10]) Centro de Recursos Hídricos e Estudos Ambientais, Universidade de São Paulo – EESC/USP, Rodovia Domingos Innocentini, km 13, Itirapina (SP), e-mail: phelipe.anjinho@usp.br, (16) 3371 – 8255

2. MATERIAL E MÉTODOS

2.1. Área de Estudo

O estudo foi desenvolvido na bacia hidrográfica do ribeirão do Lobo (BHRL), localizada na região central do Estado de São Paulo (Figura 1). Possui área de 221,20 km² e seus limites se situam no interior da Unidade de Gerenciamento de Recursos Hídricos do Estado de São Paulo (UGRHI 13) dos rios Tietê e Jacaré-Guaçu.

Figura 1: Localização da BHRL.



2.2. Base de Dados

Os dados utilizados nesta pesquisa foram adquiridos no sítio eletrônico do projeto MapBiomas (disponíveis gratuitamente em https://mapbiomas.org/) e no estudo de Anjinho *et al*. (2021), ambos desenvolvidos para o ano de 2018. Os dados do MapBiomas são gerados por meio de técnicas de processamento de imagens de satélite, obtidas pelos sensores *Thematic Mapper* (TM), *Enhanced Thematic Mapper Plus* (ETM+) e *Operational Land Imager* (OLI), a bordo dos satélites Landsat 5, Landsat 7 e Landsat 8, respectivamente (Souza *et al*., 2020).

Utilizou-se dados da 5ª coleção do MapBiomas. Os dados apresentados no estudo de Anjinho *et al*. (2021) foram gerados por meio da digitalização manual de imagens obtidas pelo sensor *Multispectral Instrument* (MSI), embarcado no satélite Sentinel-2, com resolução espacial de 10 metros, disponíveis gratuitamente na base de dados do *U.S Geological Survey* (USGS), no sítio eletrônico www.earthexplorer.usgs.gov.

2.3. Avaliação da acurácia dos dados de uso e ocupação do solo

Neste estudo, a acurácia dos dados de uso e ocupação do solo do MapBiomas foi avaliada por meio da comparação com os dados gerados por Anjinho *et al*. (2021). Utilizou-se o estudo desses autores como referência, pois eles digitalizaram manualmente, pixel por pixel, o uso e ocupação do solo da BHRL, além de, mensalmente, realizarem saídas de campo para amostras de qualidade da água (Neves *et al*., 2021, no prelo), proporcionando maior conhecimento das características de uso e ocupação do solo da área estudo.

O índice estatístico Kappa (*Index Kappa of Agreement* – KIA) foi usado para comparar os dados de uso do solo (Landis; Koch,1977). Esse índice permite mensurar a qualidade do mapeamento tendo como referência um outro mapeamento, considerado "real" e representativo da área de estudo. A Tabela 1 apresenta a classificação de desempenho conforme os valores de índice Kappa.

Tabela 1: Classificação do índice Kappa.

Índice Kappa	Classificação
≤ 0,2	Ruim
0,2 – 0,4	Razoável
0,4 – 0,6	Bom
0,6 – 0,8	Muito bom
0,8 – 1,0	Excelente

Fonte: adaptado de Landis e Koch (1977)

Utilizou-se dois níveis de uso do solo para a comparação dos dados. O nível 1 representa as classes gerais de uso do solo, que foram divididas em: i) vegetação natural; ii) silvicultura; iii) pastagem; iv) agricultura; v) infraestrutura urbana; vi) água; e vii) outras áreas não vegetadas. O nível 2 apresenta um detalhamento maior, em que as classes foram divididas em: i) formação florestal, representada por matas ciliares, matas de galeria e cerradão; ii) cerrado, que abrange formações campestres e savânicas; iii) silvicultura, caracterizada por pinus e eucalipto; iv) pastagem; v) cana de açúcar; vi) lavoura perene, vii) soja; viii) outras lavouras temporárias; ix) área urbana; x) ocupação humana de baixa densidade, xi) água; xii) outras áreas não vegetadas; e xii) mineração.

3. RESULTADOS

A representação espacial das classes de uso e ocupação do solo obtidas pelo satélite Sentinel-2 e MapBiomas é apresentada na Figura 2, e a Tabela 1 apresenta o valor e o percentual de cada classe de uso do solo. A classe mais representativa, considerando os dados oriundos do satélite Sentinel-2, é a silvicultura que ocupa 33,26% da BHRL. A vegetação natural é a segunda classe com maior percentual (30,29%), caracterizada por fragmentos de Cerrado, formação florestal e áreas alagadas. As áreas de pastagem ocupam 15,16%; a agricultura (cana de açúcar e lavoura perene) abrange 15,64% da área de estudo; as áreas urbanizadas, incluindo as ocupações humanas de baixa densidade, ocupam 3,66%; e as atividades de mineração representam 0,13% da BHRL.

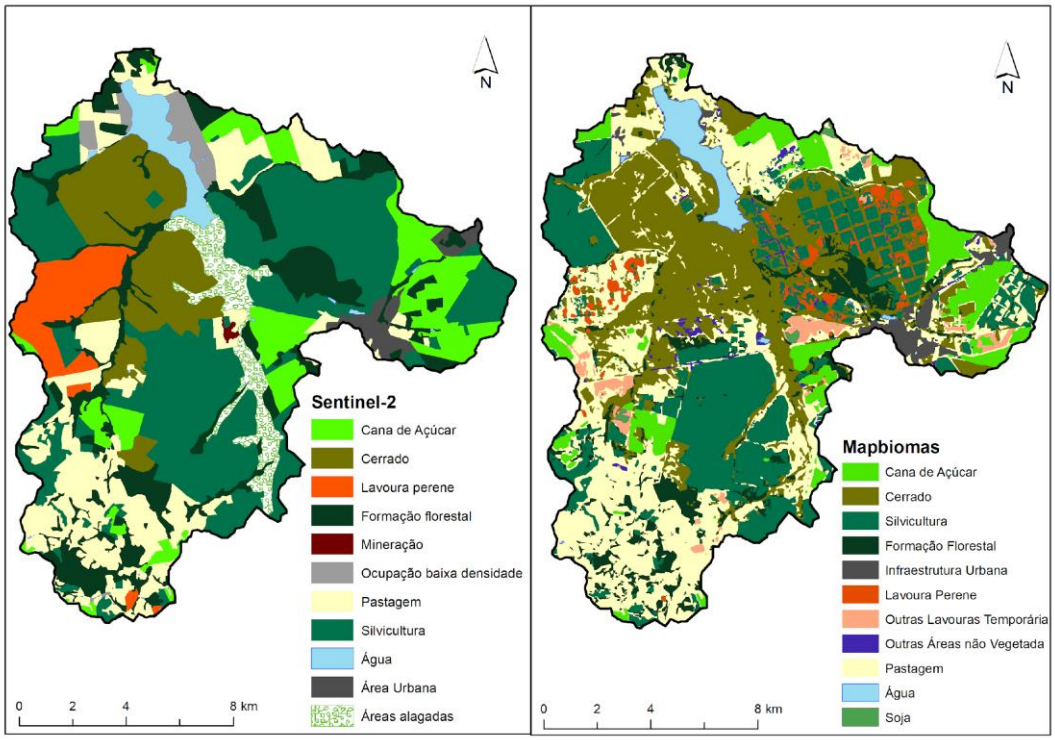
Para os dados obtidos pelo MapBiomas, as classes com maior percentual de cobertura do solo foram: vegetação natural (33,26 %), representada por fragmentos de cerrado e formação florestal; pastagem (26,72%), que inclui as áreas classificadas como “mosaico de agricultura e pastagem”; silvicultura (22,54%); agricultura (12,15%), representada por cana de açúcar, lavoura perene, outras lavouras temporárias e soja; área urbana (1,94%); água (2,54%); e outras áreas não vegetadas (0,85%).

Tabela 2: Área e percentual das classes de uso e ocupação do solo obtidas pelo Sentinel-2 e Mapbiomas.

Classes de uso do solo	Sentinel-2 (km²)	%	MapBiomas (km²)	%
Água	6,07	2,74	5,63	2,54
Área urbana	3,96	1,79	4,30	1,94
Áreas alagadas	10,07	4,55	0	0
Cana de açúcar	23,66	10,70	17,40	7,87
Cerrado	26,04	11,77	58,13	26,28
Lavoura perene	10,93	4,94	4,73	2,14
Formação florestal	31,12	14,07	15,44	6,98

Mineração	0,29	0,13	0	0
Ocupação humana de baixa densidade	4,14	1,87	0	0
Pastagem	33,53	15,16	59,11	26,72
Silvicultura	71,41	32,28	49,86	22,54
Outras áreas não vegetadas	0	0	1,87	0,85
Outras lavouras temporárias	0	0	4,18	1,89
Soja	0	0	0,55	0,25

Figura 2: Mapas de uso e ocupação do solo (nível 2) gerados por meio da digitalização manual de imagens obtidas pelo satélite Sentinel-2 e pelo projeto MapBiomas (2018).



A Figura 3 mostra a distribuição das áreas da BHRL que apresentam classes de uso do solo concordantes e discordantes, e a Figura 4 apresenta o percentual de acertos e erros de cada classe de uso do solo geradas pelo MapBiomas. Os resultados mostram, considerando os dois níveis de detalhamento, que os dados são similares no que se refere à representação do padrão espacial do uso e ocupação do solo da BHRL. Para o nível 1 de detalhamento, o valor do índice Kappa foi considerado muito bom (Kappa = 0,62), com percentual médio de acertos de 71%. Para o nível 2, o índice Kappa foi igual a 0,51, o que indica uma classificação boa, e o percentual médio de acertos foi de 60%.

A maior consistência entre os dados, para os dois níveis de detalhamento, foi observada para a classe água, que representa os corpos hídricos encontrados na bacia. Considerando o nível 1 de detalhamento, praticamente todas as classes de uso do solo apresentaram altos percentuais (> 70%) de concordância, exceto as áreas de pastagem e mineração. Os maiores erros observados na classe de

pastagem foram encontrados na região oeste da bacia e nas áreas lindeiras ao reservatório do Lobo. As culturas perenes localizadas na margem esquerda do ribeirão do Lobo e as áreas de veraneio do balneário Santo Antônio foram classificadas como pastagem nos dados do MapBiomas. Parte das áreas alagadas da BHRL também foram classificadas como pastagem. As atividades de mineração situadas nas planícies de inundação do rio Itaqueri foram identificadas como "outra área não vegetada".

Para o nível 2 de detalhamento, os maiores percentuais de acertos (>70%) foram observados para as classes de silviculturas (93%), cana de açúcar (85%) e áreas urbanas (84%). Os maiores erros encontrados, além dos já mencionados para as áreas de lavoura perene e pastagem, foram observados para a vegetação de cerrado (63%). O MapBiomas classificou como vegetação de cerrado (formações campestres e savânicas), as áreas alagadas do rio Itaqueri e as estradas que cortam a estação experimental de Itirapina, unidade de conservação destinada à produção florestal e pesquisa científica. Outros fragmentos de formação florestal também foram classificados como cerrado pelo MapBiomas.

Figura 3: Distribuição espacial das áreas concordantes e discordantes.

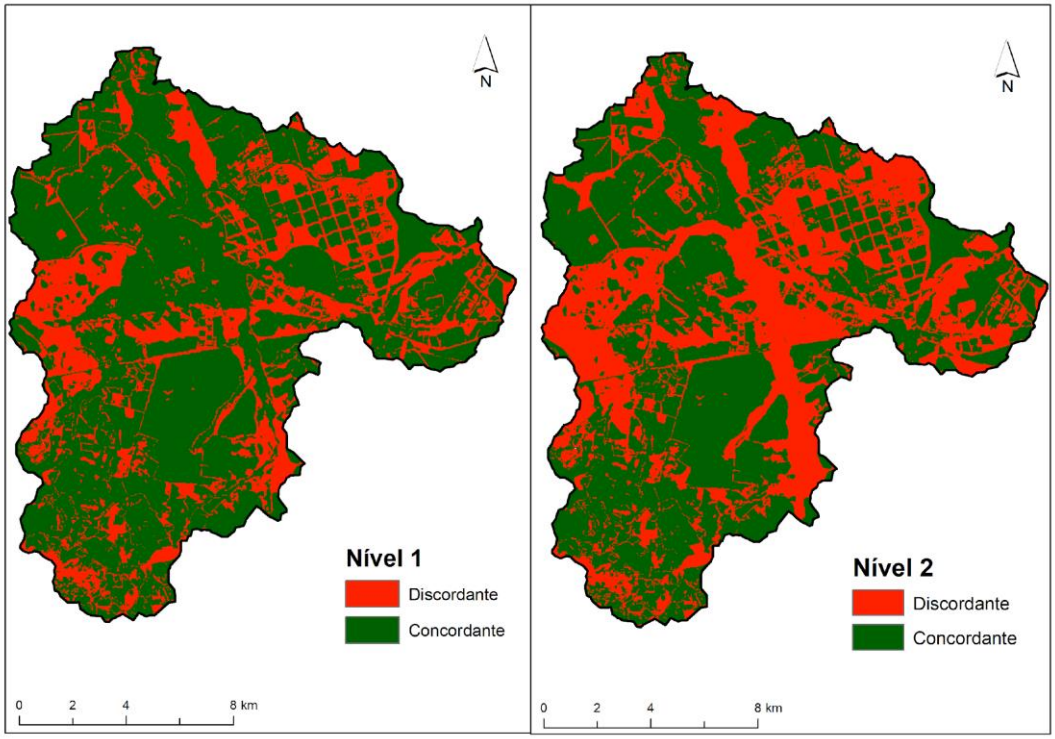
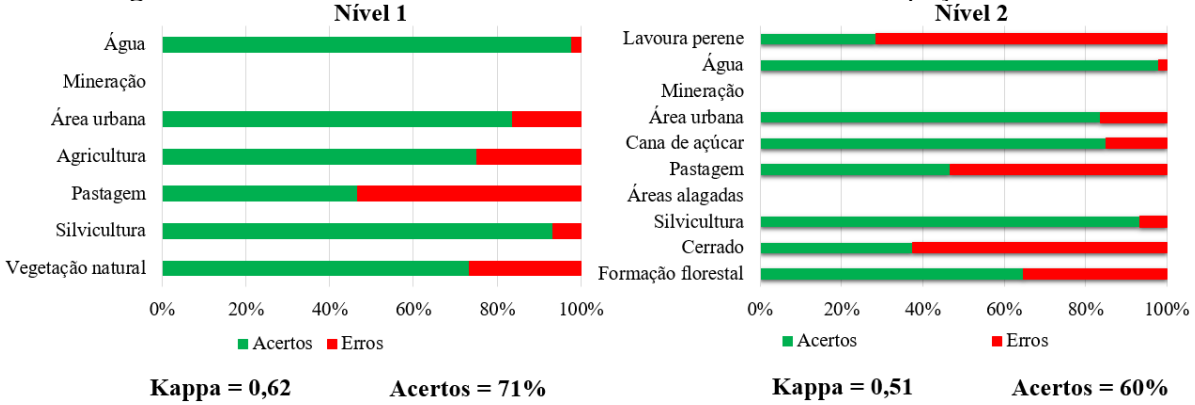


Figura 4: Percentual de acertos e erros das classes de uso e ocupação do solo.



4. DISCUSSÃO

Os resultados deste estudo mostram que os dados do MapBiomas representam bem o padrão espacial de uso do solo da BHRL. A grande maioria dos erros observados são aceitáveis, com exceção das classes lavouras perenes e ocupação humana de baixa densidade que foram classificadas como pastagem pelo MapBiomas. Os demais erros ocorrem, possivelmente, em virtude da similaridade da resposta espectral das classes de uso do solo discordantes, o que gera maior dificuldade para o algoritmo distinguir as classes corretamente. Exemplo disto são as áreas alagadas que foram classificadas como formações savânicas e campestres pelo MapBiomas. Essas classes são similares quando representadas na imagem de satélite, o que dificulta sua distinção.

Embora o método deste estudo tenha permitido avaliar a acurácia dos dados do MapBiomas, ressalta-se que a abordagem usada apresenta algumas incertezas que devem ser consideradas. O mapa de uso do solo de referência utilizado neste trabalho (Anjinho *et al*., 2021) foi gerado apenas por meio de digitalização manual em ambiente SIG, ou seja, fortemente dependente da análise subjetiva e interpretação dos autores. Por mais que eles tenham realizado algumas verificações durante saídas de campo para amostragens de qualidade da água (Neves *et al*., 2021, no prelo), nenhum levantamento específico para validação de uso do solo foi realizado. Isso pode levar a incertezas nos dados de uso do solo usados como referência neste estudo. Sendo assim, para pesquisas futuras recomenda-se realizar levantamentos de campo para validar os dados de uso do solo e reduzir as incertezas.

5. CONCLUSÃO

Este estudo visou avaliar a acurácia dos dados de uso e ocupação do solo do projeto MapBiomas, por meio de análise comparativa com os dados obtidos a partir de digitalização manual de imagens do satélite Sentinel-2.

A análise comparativa indicou que os dados gerados pelo MapBiomas foram consistentes e representaram de forma adequada o padrão espacial de uso do solo da BHRL. A acurácia dos dados foi classificada como muito boa (nível 1) e boa (nível 2). O percentual médio de acertos dos dados do MaBiomas foi de 71% para o nível 1 e 60% para o nível 2.

Os resultados deste estudo mostram que dados disponibilizados pelo MapBiomas são de boa qualidade e podem ser utilizados para a elaboração de outros estudos ambientais que demandam dados dessa natureza, como modelagens hidrossedimentológicas, ecossistêmicas e planejamento ambiental. Entretanto, tendo em vista que o mapeamento de referência utilizado neste estudo não foi devidamente validado por amostragem de campo, ressalta-se a necessidade de mais pesquisas para avaliar com mais consistência os dados gerados pelo MapBiomas.

AGRADECIMENTOS

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelas bolsas concedidas.

REFERÊNCIAS

ANJINHO, P. S.; BARBOSA, M. A. G. A.; NEVES, G. L.; DOS SANTOS.; A. R.; MAUAD, F. F. (2021). *Integrated empirical models to assess nutrient concentration in water resources: case study of a small basin in southeastern Brazil.* Environmental Science and Pollution Research, v. 28, n. 18, p. 23349-23367.

ARRUDA, V. L.; PIONTEKOWSKI, V. J.; ALENCAR, A.; PEREIRA, R. S.; MATRICARDI, E. A. (2021). *An alternative approach for mapping burn scars using Landsat imagery, Google Earth*

Engine, and Deep Learning in the Brazilian Savanna. Remote Sensing Applications: Society and Environment, v. 22, p. 100472, 2021.

BUCAK, T.; TROLLE, D.; TAVŞANOĞLU, Ü. N.; ÇAKIROĞLU, A. İ.; ÖZEN, A.; JEPPESEN, E.; BEKLIOĞLU, M. (2018). *Modeling the effects of climatic and land use changes on phytoplankton and water quality of the largest Turkish freshwater lake: Lake Beyşehir.* Science of the Total Environment, v. 621, p. 802-816, 2018.

ELLISON, D.; MORRIS, C. E.; LOCATELLI, B.; SHEIL, D.; COHEN, J.; MURDIYARSO, D.; SULLIVAN, C. A. (2017). *Trees, forests and water: Cool insights for a hot world.* Global Environmental Change, v. 43, p. 51-61, 2017.

FENDRICH, A. N.; BARRETTO, A.; DE FARIA, V. G.; DE BASTIANI, F.; TENNESON, K.; PINTO, L. F. G.; SPAROVEK, G. (2020). *Disclosing contrasting scenarios for future land cover in Brazil: Results from a high-resolution spatiotemporal model.* Science of the Total Environment, v. 742, p. 140477, 2020.

FERREIRA, P.; VAN SOESBERGEN, A.; MULLIGAN, M.; FREITAS, M.; VALE, M. M. (2019). *Can forests buffer negative impacts of land-use and climate changes on water ecosystem services? The case of a Brazilian megalopolis.* Science of the Total Environment, v. 685, p. 248-258, 2019.

LANDIS, R. J.; KOCH, G. G. *The measurement of observer agreement for categorical data.* biometrics, p. 159-174, 1977.

ROSA, M. R.; BRANCALION, P. H.; CROUZEILLES, R.; TAMBOSI, L. R.; PIFFER, P. R.; LENTI, F. E.; METZGER, J. P. (2021). *Hidden destruction of older forests threatens Brazil's Atlantic Forest and challenges restoration programs.* Science advances, v. 7, n. 4, p. eabc4547, 2021.

SOUZA, C., AZEVEDO, Tasso. *MapBiomas general handbook. MapBiomas: São Paulo, Brazil*, p. 1-23, 2017.

SOUZA, C. M.; Z SHIMBO, J.; ROSA, M. R.; PARENTE, L. L.; A ALENCAR, A.; RUDORFF, B. F.; AZEVEDO, T. (2020). *Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine.* Remote Sensing, v. 12, n. 17, p. 2735, 2020.

SUN, Ge. HALLEMA, D., ASBJORNSEN, H. *Ecohydrological processes and ecosystem services in the Anthropocene: a review.* Ecological Processes, v. 6, n. 1, p. 1-9, 2017.