



Evolution of MapBiomias Collections' Accuracy for the Highly Fragmented São Paulo landscape

Evolution of MapBiomias Collections' Accuracy for the Highly Fragmented São Paulo landscape

Nadinne Fernandes de Oliveira ¹, Eduardo Moraes Arraut ².

¹ Technological Institute of Aeronautics (ITA), São José dos Campos, Brazil. nadinne@ita.br.
ORCID: <https://orcid.org/0000-0001-5683-9161>

² Technological Institute of Aeronautics (ITA), São José dos Campos, Brazil. emarraut@ita.br.
ORCID: <https://orcid.org/0000-0001-5323-4431>

Recebido: 06.2023 | Aceito: 06.2024

Resumo: A avaliação da acurácia de mapeamentos da cobertura da terra é essencial para os usos científico, prático e político dos mapas. No Brasil, o projeto MapBiomias vem mapeando anualmente a cobertura da terra em todo o território via classificação automática de imagens Landsat de média resolução espacial (30 m) desde 1985. Cada nova versão do algoritmo de classificação gera uma nova coleção de mapas que são sujeitos a uma avaliação da acurácia em nível nacional. Entretanto, é cada vez mais frequente o uso do MapBiomias para estudos regionais, municipais ou locais para os quais a avaliação da acurácia em nível nacional não é adequada. Aqui avaliamos a acurácia e a evolução da acurácia das principais categorias de cobertura do MapBiomias para o estado de São Paulo (SP), o mais urbanizado do país e objeto de muitos estudos e políticas públicas relacionadas à cobertura da terra. Analisamos as coleções 3.1, 4.1, 5.0, 6.0 e 7.0 para o ano de 2017, o mais recente com coincidência de classes em todas as coleções, considerando as classes: Formação Florestal, Floresta Plantada, Pastagem, Cana de Açúcar, Infraestrutura Urbana e Rio, Lago e Oceano. A acurácia global (AG) foi menor na coleção 4.1 (91%) e maior na 7.0 (96%). As acurácias do produtor (AP) e do usuário (AU) melhoraram da coleção 3.0 para a 7.0, com exceção das AP para Floresta Plantada que praticamente permaneceu inalterada, e de Infraestrutura Urbana que vem piorando gradativamente até atingir o menor valor na 7.0 (0,87). Mesmo assim, o fato da AG, e particularmente da AP e AU da coleção 7.0 estarem acima de 0,84 indica que em São Paulo o MapBiomias é acurado o suficiente para, por exemplo, muitas análises comuns na escala da paisagem, em ecologia (ex. modelagem da distribuição de espécie de mamífero ou ave), agricultura (ex. estimativa de safra) ou engenharia (ex. escolha de sítio aeroportuário).

Palavras-chave: Avaliação da acurácia; Mapeamento da cobertura da terra; QGIS; GRASS.

Abstract: The assessing the accuracy of land cover mappings is essential for the scientific, practical and policy uses of maps. In Brazil, the MapBiomias project has been annually mapping the land cover across the entire territory via automatic classification of Landsat images of medium spatial resolution (30 m) since 1985. Each new version of the classification algorithm generates a new collection of maps that are subject to an accuracy assessment at the national level. However, MapBiomias is increasingly used for regional, municipal or local studies for which the assessment of accuracy at the national level is not adequate. Here we evaluate the accuracy and evolution of accuracy of the main categories of MapBiomias coverage for the state of São Paulo (SP), the most urbanized in the country and the object of many studies and public policies related to land cover. We analyzed collections 3.1, 4.1, 5.0, 6.0 and 7.0 for the year 2017, the most recent with coincidence of classes in all collections, considering the classes: Forest Formation, Planted Forest, Pasture, Sugar Cane, Urban Infrastructure and River, lake and ocean. The global accuracy (GA) had its lowest value in collection 4.1 (91%) and highest in 7.0 (96%). Producer (PA) and user accuracies (UA) improved from collection 3.0 to 7.0, with the exception of PA for Planted Forest, which remained virtually unchanged, and for Urban Infrastructure, which has been showing a tendency to worsen over the course of the collections, reaching its lowest value at 7.0 (0,87). Even so, the fact that the GA, and particularly the PA and UA of collection 7.0 are above 0.84 indicates that in São Paulo MapBiomias is accurate enough for, for example, many common analyzes at the landscape scale, in ecology (e.g. distribution modeling of species of mammal or bird), agriculture (e.g. harvest estimation), or engineering (e.g. choice of airport site).

Keywords: Accuracy assessment; Land-cover mapping; QGIS; GRASS.

1. INTRODUÇÃO

O uso e cobertura da terra estão relacionados a clima e biodiversidade, disponibilidade hídrica e agricultura, engenharia, economia e vários outros fatores cruciais para a vida humana. Por exemplo, uma avaliação da relação entre o desmatamento na Amazônia e a emissão de carbono mostrou que até 2017 o ganho em estoque total de carbono pelo crescimento da floresta compensou em torno de 10% das emissões resultantes da perda de florestas antigas (SMITH et al, 2020). Por sua vez, na Inglaterra a substituição de pradarias por silvicultura para reduzir alagamentos e construir casas de madeira que estocam carbono está reduzindo a disponibilidade de recursos cruciais para o gavião-comum *Buteo buteo*, o que nas próximas décadas deve induzir o terceiro grande declínio dessa população (ARRAUT et al, 2021; KENWARD et al, 2018). No âmbito do processo de seleção de novo sítio aeroportuário regional em município Brasileiro, o mapeamento da cobertura da terra tem sido essencial para, por exemplo, posicionar novo sítio com, simultaneamente, menor custo via redução de volume de terraplenagem, e menor impacto ambiental e social, como desmatamento ou drenagem de curso d'água, evitação de terras indígenas, quilombolas e assentamentos, e redução do número de propriedades rurais a serem desapropriadas (ALVES et al, 2020).

Com a evolução do sensoriamento remoto e o aumento na quantidade e diversidade de mapeamentos de cobertura da terra, juntamente com o aumento no uso dessas informações em análises que envolvem outros dados geográficos, tornou-se fundamental avaliar a acurácia desses mapeamentos (CONGALTON, 1991). Uma avaliação da acurácia de mapeamento, reportada essencialmente em uma matriz de confusão (ou matriz de erro), quantifica a coincidência entre o mapa de forma geral e por categoria e a realidade. Quando essa acurácia é feita seguindo um desenho amostral probabilístico bem executado, tanto as estimativas de acurácia quanto os erros padrão ou intervalos de confiança associados são válidos, garantindo a quantificação crucial da incerteza associada ao mapeamento (STEHMAN, FOODY, 2019 e CONGALTON, 1991).

O projeto MapBiomias, nascido em 2015 via colaboração entre universidades, ONGs e startups de tecnologia, mapeia anualmente a cobertura e o uso da terra em todo o território brasileiro e disponibiliza seus mapas gratuitamente. Esse mapeamento é feito classificando-se um histórico de imagens multiespectrais com resolução espacial de 30m de 1985 até o presente, baseadas nos sensores dos satélites da missão Landsat (MAPBIOMAS, 2022). Como o algoritmo de classificação vem sendo aprimorado, a cada nova versão dele é gerado um novo mapa por ano da série. A avaliação da acurácia do MapBiomias é baseada em uma amostra de ~75.000 píxeis criada por meio de interpretação visual de imagens Landsat, chamada de base de dados de referência. Tanto o tamanho quanto a distribuição dessa amostra de referência são estipulados considerando todo o território brasileiro.

Entretanto, é cada vez mais comum o uso do MapBiomias para estudos em subáreas bem menores do país, incluindo estado, município ou mesmo localidades específicas para as quais a avaliação da acurácia em extensão nacional pode não ser adequada. Isto porque, apesar da metodologia robusta de avaliação da acurácia usada pelo MapBiomias, e em particular dos pequenos erros-padrão geral e por bioma obtidos (SOUZA Jr et al, 2020), em uma matriz de confusão não há informação sobre os padrões espaciais dos erros de mapeamento, significando que eventuais mudanças nas taxas de erro derivadas de variabilidade em certa classe ao longo do país não são detectáveis (WICKHAM et al, 2017). Não é possível, por isso, saber o quanto os resultados sobre a mudança no uso da terra para a Microrregião de Pirapora, norte de Minas Gerais, que representa < 0,003 % do território brasileiro, ou para o município de Bragança-PA, que representa < 0,0003 da área nacional (SENA-SOUZA et al, 2022; RIBEIRO, 2022), derivam de alterações reais da paisagem ou de erro localizado de mapeamento induzido por especificidades das paisagens locais (ex. variação na vegetação, solo, relevo). Mais ainda, dado o grande investimento da equipe MapBiomias no aprimoramento do algoritmo de classificação, é importante também quantificar a mudança na acurácia entre as coleções. Neste artigo, avalia-se a acurácia e a evolução da acurácia de mapeamento das coleções 3.1 a 7.0 do MapBiomias para todo o estado de São Paulo, o mais populoso do país e objeto de estudo de diversos estudos. Por exemplo, em São Paulo o MapBiomias vem sendo usado para avaliar a dinâmica da paisagem nas pequenas (comparativamente à extensão do Estado de SP) zonas de amortecimento de unidades de conservação (UC), Áreas de Preservação Permanente (APP) ou áreas úmidas periurbanas, ou embasar planos

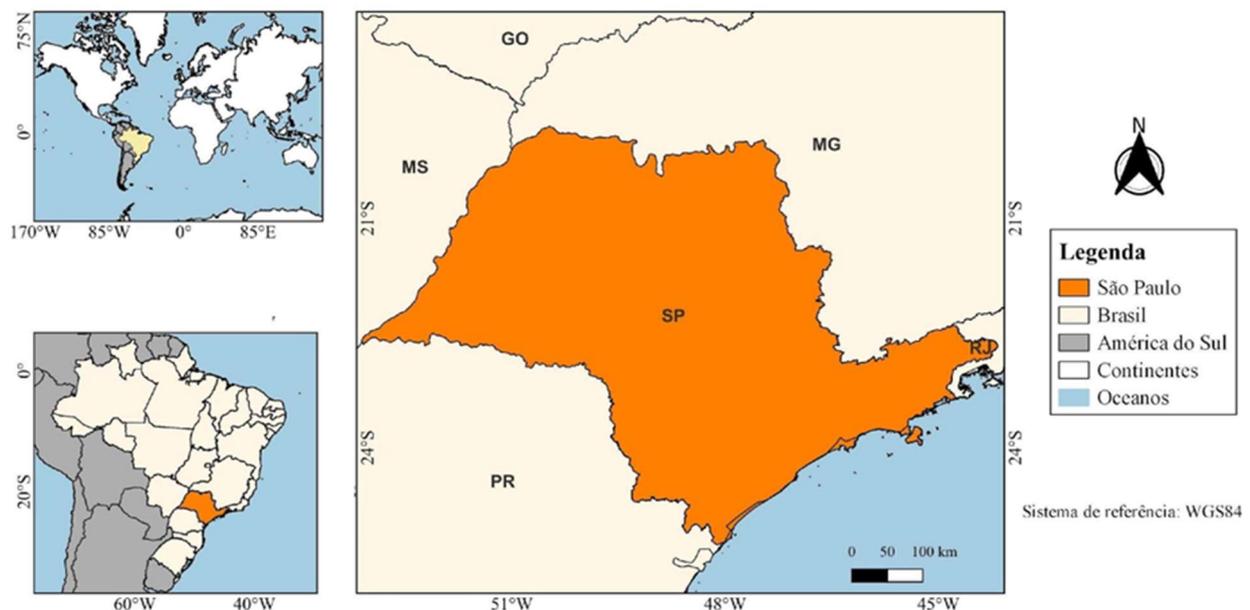
de manejo de UC do estado (CARMO et al, 2023; SANTOS JUNIOR, 2023; CASTOJO, JESUS, 2022; COSTA, 2022). A expectativa é que este estudo possa colaborar para o aprimoramento da qualidade do mapeamento do MapBiomias, além de oferecer uma base referencial da incerteza do mapeamento para investigações direcionadas ao estado de SP.

2. MATERIAIS E MÉTODOS

2.1 Área de estudo

A área de estudo tem como foco o estado de São Paulo, com extensão territorial de 248.219,485 km² (IBGE, 2022) dentro dos biomas Cerrado e Mata Atlântica, população estimada em 46,3 milhões de habitantes (estado mais populoso do país, com 21,9% da população total) (IBGE, 2020), e paisagem altamente heterogênea e fragmentada. De acordo com o ranking de 2020 do Instituto Brasileiro de Geografia e Estatística (IBGE), o estado de São Paulo abriga três das 17 cidades com mais de 1 milhão de habitantes no país, sendo a cidade de São Paulo a mais populosa, seguida de Guarulhos e Campinas, em 13º e 14º lugares, respectivamente. A região metropolitana de São Paulo é a mais populosa do Brasil, com uma população de 21,9 milhões de habitantes (IBGE, 2020).

Figura 1 - Mapa de localização do estado de São Paulo.

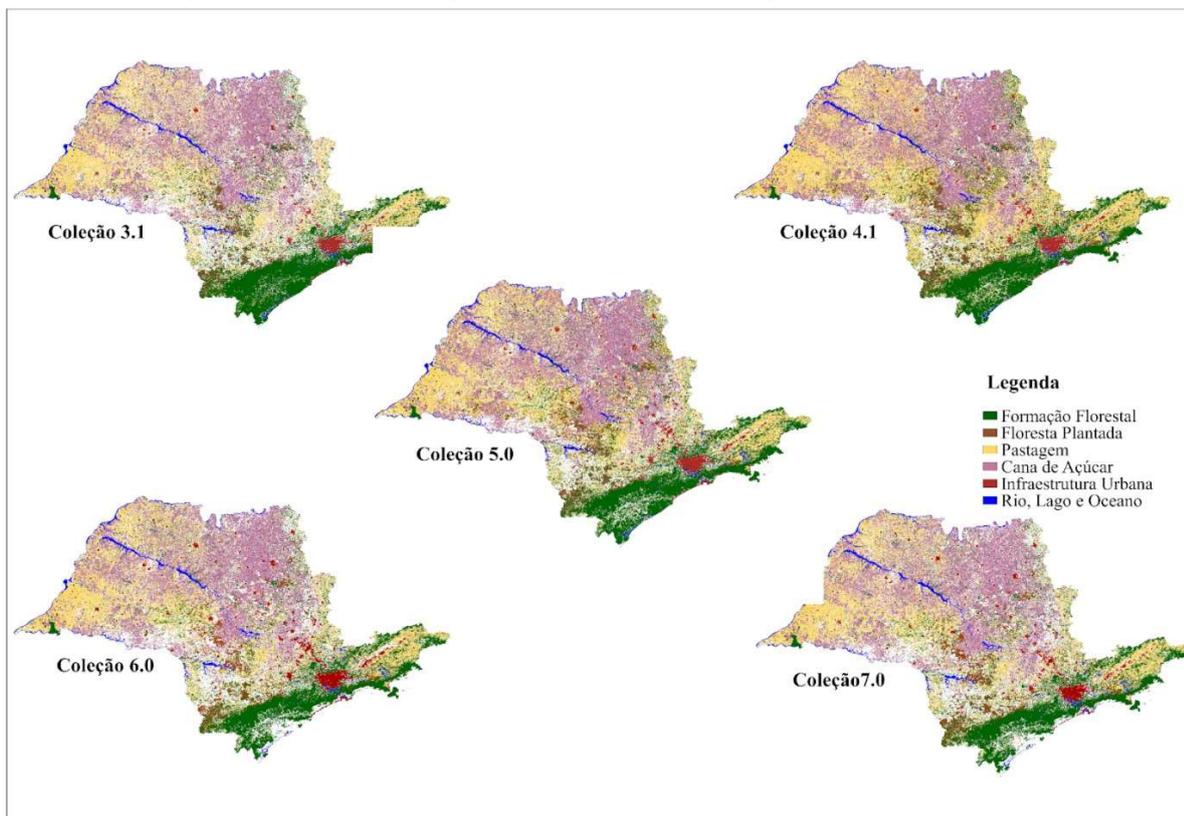


Fonte - Autores (2022)

2.2 Metodologia

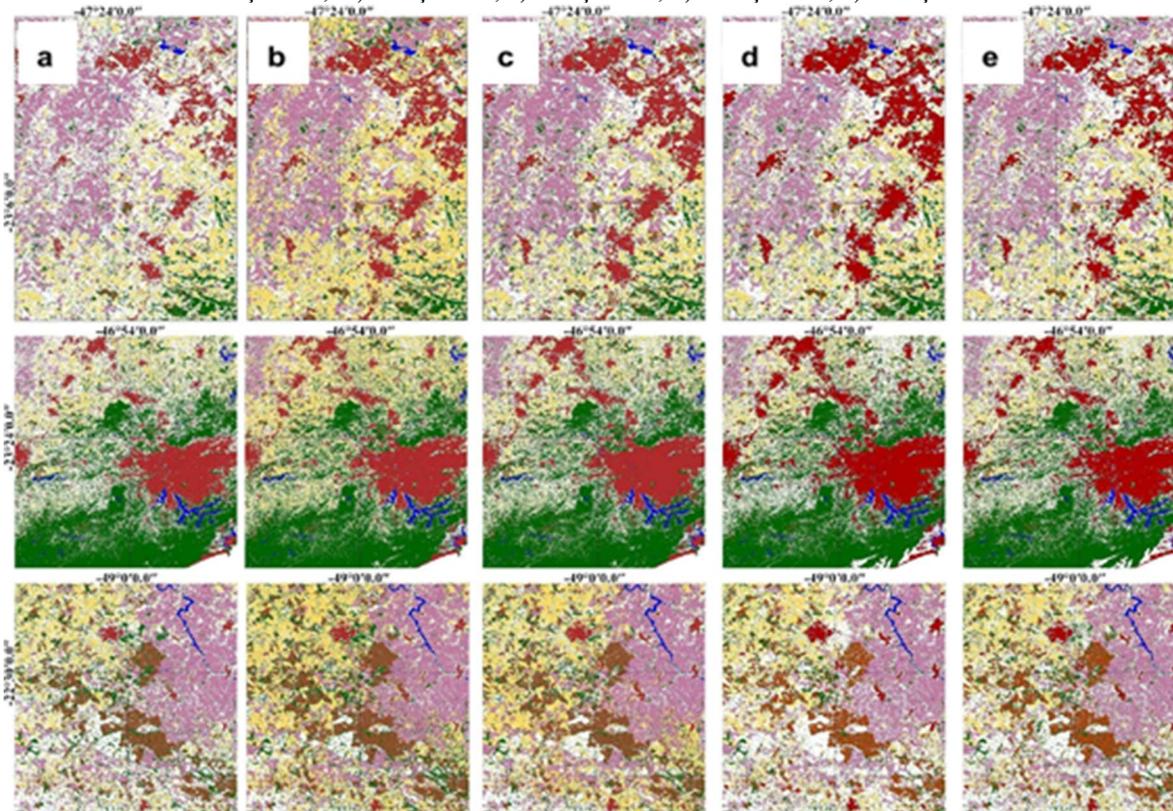
Para esta análise, que envolveu as coleções 3.1, 4.1, 5.0, 6.0 e 7.0, o ano de 2017 foi escolhido por ser o mais recente que constava em todas as coleções em questão (Figs. 2 e 3). A fim de permitir a comparação entre as coleções, apenas as classes comuns a todas as coleções do MapBiomias avaliadas foram consideradas na análise da acurácia. Assim, por exemplo, as classes Café, Citrus e Restinga Arborizada presentes somente nas coleções 6 e 7, a classe Lavoura Perene encontrada apenas na coleção 5, bem como a classe Lavoura Temporária exclusiva das coleções 3.1 e 4.1 não foram contempladas na análise. As classes finais consideradas foram: Formação Florestal, Floresta Plantada, Pastagem, Cana de Açúcar, Infraestrutura Urbana e Rio, Lago e Oceano. A classe 34, denominada de classe “erros”, foi utilizada quando a interpretação visual do mapa-base de altíssima resolução espacial indicava que o MapBiomias havia classificado erroneamente uma determinada classe. Por exemplo, se o MapBiomias classificou uma área como "Pastagem" mas a interpretação visual indica claramente que é "Infraestrutura Urbana", é atribuída a classe "erros".

Figura 2 – Coleções do MapBiomias e classes analisadas para o estado de São Paulo.



Fonte: Autores (2022).

Figura 3 – Detalhes das coleções do MapBiomias e classes analisadas para três subregiões do estado de São Paulo com diferentes características em termos de cobertura da terra: linha 1 = região de Campinas e arredores linha, com alta cobertura urbana distribuída, amplas áreas de cana e pastagem e pouca floresta nativa, 2 = região metropolitana da cidade de São Paulo e arredores, com densa e extensa cobertura urbana rodeada principalmente por pastagem e floresta nativa, e linha 3 = região de Bauru e arredores, com pequena cobertura urbana, e bastante pasto e cana. Colunas: a) Coleção 3.1; b) coleção 4.1; c) coleção 5.0; d) Coleção 6.0; e) Coleção 7.0.



Fonte: Autores (2023).

Para realizar as análises, foi utilizado o software QGIS 3.22 com GRASS 7.2.2 para calcular a área das classes, o plugin AcATama para gerar os pontos aleatórios e produzir a matriz de confusão. Para estimar a área das classes, as imagens foram reprojatadas para coordenadas UTM e, em seguida, foi utilizada a função "r.report" para obter a área de cada classe de acordo com a unidade de medida escolhida (Tabela 1).

Tabela 1 - Proporção das classes por coleção. Área em km² da coleção 7 para referência: Formação F. (46.837,20), Floresta P. (8.752,95), Pastagem (46.652,05), Cana de açúcar (56.299,28), Infraestrutura U.(7.310,91) e Rio, Lago e Oceano (6.144,56).

Classe	Nome	Área (%) 3.1	Área (%) 4.1	Área (%) 5.0	Área (%) 6.0	Área (%) 7.0
3	Formação Florestal	20,15	20,14	20,38	19,09	18,87
9	Floresta Plantada	3,34	4,34	4,06	3,40	3,53
15	Pastagem	18,54	26,90	21,49	18,45	18,79
20	Cana de Açúcar	22,64	24,80	24,84	23,00	22,68
24	Infraestrutura Urbana	2,42	2,76	3,01	3,33	2,95
33	Rio, Lago e Oceano	2,29	2,28	2,35	2,39	2,48

Elaboração: Os autores (2022).

Para garantir uma distribuição proporcional dos pontos amostrais de acordo com o tamanho de cada classe, utilizando o plugin AcATama 3166 pontos foram distribuídos seguindo uma amostragem estratificada (STEHMAN e FOODY, 2019). O tamanho amostral por classe foi determinado utilizando a seguinte fórmula (STEHMAN e FOODY, 2019) e considerando um intervalo de confiança de 90% e uma margem de erro de 5%:

$$N = (Z^2p(1-p))/d^2 \tag{1}$$

onde N é o número de amostras, Z é o intervalo de confiança, p é a porcentagem da área da classe e d é a margem de erro. Na Tabela 1 é apresentada a quantidade de amostras necessárias para cada classe considerada na análise, por coleção.

Tabela 2 - Quantitativo de amostras por coleção.

Classe	Nome	3.1	4.1	5.0	6.0	7.0
3	Formação Florestal	173	173	174	166	164
9	Floresta Plantada	35	45	42	35	36
15	Pastagem	162	211	181	162	164
20	Cana de Açúcar	188	200	201	190	188
24	Infraestrutura Urbana	25	29	31	35	31
33	Rio, Lago e Oceano	24	24	25	25	26
		608	682	654	613	609
Total: 3166						

Elaboração: Os autores (2022).

O desenho resposta, que é o protocolo utilizado para associar a classe de referência a cada unidade amostral, consistiu em considerar como classe da amostra de referência aquela que predominasse no *buffer* com raio de 10 metros em torno de cada ponto (Fig. 4). O tamanho desse *buffer* foi escolhido de modo a caber na unidade mínima classificada (um pixel) e ao mesmo tempo reduzir eventuais erros na determinação visual da classe decorrentes de diferenças no registro entre mapa e amostra de referência. A determinação da classe de referência por ponto foi feita com base na interpretação visual de até quatro imagens ou mapas diferentes, incluindo o Google Satellite, Yandex, ESRI e o projeto Canasat (CANASAT, 2013).

Ao comparar os mapas-base com a função *time lapse* do Google Earth Pro (GEP), foi observado que estes eram de datas diferentes: o Yandex com datas mais antigas, a maioria em 2016, e o Google Satellite mais atualizado, principalmente de 2023..Para melhorar a distinção da cultura perene cana-de-açúcar, que nessas imagens pode ser confundida com outras categorias como pastagem (após a colheita da cana-de-

açúcar) ou vegetação plantada (quando está no seu estágio de crescimento alto e denso), foi usado o mapa de 2013 para o estado de São Paulo do Projeto Canasat. Foram calculadas as acurácias global, e do produtor (AP, ou 100% - erro de omissão) e do usuário (AU, ou 100% - erro de comissão) para cada classe de cada coleção.

Figura 4 - Exemplos de amostras para a validação por classe e por coleção. Imagem de fundo do Google Satellite. a) Formação Florestal (3.1); b) Floresta Plantada (4.1); c) Pastagem (5.0); d) Cana de Açúcar (6.0); e) Infraestrutura Urbana (7.0); f) Rios, Lagos e Oceanos (7.0).



Elaboração: Os autores (2023).

3. RESULTADOS E DISCUSSÃO

A acurácia global de um mapa representa a proporção de acerto desse mapa relativa a uma amostra de referência de menor incerteza (Stehman, Foody 2019). Considerando as sete categorias do MapBiomas avaliadas aqui, que juntas compõem 69,29% da área de São Paulo mapeada na coleção 7.0, a houve declínio da acurácia global das coleções 3.1 (92%) para a 4.1 (91%), seguido de subida gradual até a coleção 7.0

(96%) (Tabelas 3-7). Isso indica que ao longo das coleções a qualidade média do MapBiomias para o Estado de São Paulo vem melhorando.

Tabela 3 - Matriz de confusão - coleção 3.1. FF = Formação Florestal; FP = Floresta Plantada; PA = Pastagem; CA = Cana de Açúcar; IU = Infraestrutura Urbana; RLO = Rios, Lagos e Oceanos; 34 = Classe erros; AU = Acurácia do Usuário e AP= Acurácia do Produtor.

		MapBiomias 3.1							Total	AU
		FF	FP	PA	CA	IU	RLO	34		
Mapa Base	FF	154	1	2	0	0	0	0	157	0,98
	FP	8	32	0	0	0	0	0	40	0,80
	PA	8	1	147	7	0	0	0	163	0,90
	CA	0	0	10	177	0	0	0	187	0,95
	IU	0	0	1	0	25	0	0	26	0,96
	RLO	3	0	0	0	0	24	0	27	0,89
	34	0	1	2	4	0	0	0	7	
	Total	173	35	162	188	25	24	0		
	AP	0,89	0,91	0,91	0,94	1	1			92%

Elaboração: Os autores (2022).

Tabela 4 - Matriz de confusão - coleção 4.1. FF = Formação Florestal; FP = Floresta Plantada; PA = Pastagem; CA = Cana de Açúcar; IU = Infraestrutura Urbana; RLO = Rios, Lagos e Oceanos; 34 = Classe erros; AU = Acurácia do Usuário e AP= Acurácia do Produtor.

		MapBiomias 4.1							Total	AU
		FF	FP	PA	CA	IU	RLO	34		
Mapa Base	FF	157	3	8	2	0	0	0	170	0,92
	FP	13	38	6	0	0	0	0	57	0,67
	PA	2	2	182	8	1	0	0	195	0,93
	CA	0	1	10	189	0	0	0	200	0,95
	IU	0	0	1	0	28	0	0	29	0,97
	RLO	1	1	4	1	0	24	0	31	0,77
	34	0	0	0	0	0	0	0	0	
	Total	173	45	211	200	29	24	0		
	AP	0,91	0,84	0,86	0,97	0,97	1	0		91%

Elaboração: Os autores (2022).

Tabela 5 - Matriz de confusão - coleção 5.0. FF = Formação Florestal; FP = Floresta Plantada; PA = Pastagem; CA = Cana de Açúcar; IU = Infraestrutura Urbana; RLO = Rios, Lagos e Oceanos; 34 = Classe erros; AU = Acurácia do Usuário e AP= Acurácia do Produtor.

		MapBiomias 5.0							Total	AU
		FF	FP	PA	CA	IU	RLO	34		
Mapa Base	FF	161	1	5	1	2	1	0	171	0,94
	FP	9	37	1	0	0	0	0	47	0,79
	PA	1	3	156	6	1	0	0	167	0,93
	CA	1	1	12	193	0	0	0	207	0,93
	IU	0	0	1	0	28	0	0	29	0,97
	RLO	1	0	1	0	0	24	0	26	0,92
	34	1	0	5	1	0	0	0	7	
	Total	174	42	181	201	31	25	0		
	AP	0,93	0,88	0,86	0,96	0,90	0,96	0		92%

Elaboração: Os autores (2022).

Tabela 6 - Matriz de confusão - coleção 6.0. FF = Formação Florestal; FP = Floresta Plantada; PA = Pastagem; CA = Cana de Açúcar; IU = Infraestrutura Urbana; RLO = Rios, Lagos e Oceanos; 34 = Classe erros; AU = Acurácia do Usuário e AP= Acurácia do Produtor.

		MapBiomias 6.0								AU
		FF	FP	PA	CA	IU	RLO	34	Total	
Mapa Base	FF	155	2	2	1	1	0	0	161	0,96
	FP	6	33	0	1	0	0	0	40	0,83
	PA	5	0	149	4	0	0	0	158	0,94
	CA	0	0	9	184	0	0	0	193	0,95
	IU	0	0	0	0	34	0	0	34	1
	RLO	0	0	0	0	0	25	0	25	1
	34	0	0	2	0	0	0	0	2	
	Total	166	35	162	190	35	25	0		
	AP	0,93	0,94	0,92	0,97	0,97	1,00	0		95%

Elaboração: Os autores (2022).

Tabela 7 - Matriz de confusão - coleção 7.0. FF = Formação Florestal; FP = Floresta Plantada; PA = Pastagem; CA = Cana de Açúcar; IU = Infraestrutura Urbana; RLO = Rios, Lagos e Oceanos; 34 = Classe erros; AU = Acurácia do Usuário e AP= Acurácia do Produtor.

		MapBiomias 7.0								AU
		FF	FP	PA	CA	IU	RLO	34	Total	
Mapa Base	FF	157	3	1	0	2	0	0	163	0,96
	FP	6	32	0	0	0	0	0	38	0,84
	PA	0	0	155	2	1	0	0	158	0,98
	CA	0	0	5	185	0	0	0	190	0,97
	IU	0	0	1	0	27	0	0	28	0,96
	RLO	0	0	0	0	1	26	0	27	0,96
	34	1	1	2	1	0	0	0	5	
	Total	164	36	164	188	31	26	0		
	AP	0,96	0,89	0,95	0,98	0,87	1,00	0		96%

Elaboração: Os autores (2022).

Quando um mapa é usado na prática, entretanto, é comum haver interesse em categorias específicas, cujas acurácias por isso também precisam ser conhecidas. A acurácia do produtor (AP), que informa a fração dos píxeis de referência de uma certa categoria que foram corretamente mapeados, aumentou progressivamente até coleção 7. As exceções foram as classes Pastagem (15), com menor precisão na coleção 4.1 e maior precisão na coleção 6.0, e Silvicultura, cujo pico deu-se na coleção 6.. As classes Formação Florestal (30) e Cana de Açúcar (20) melhoraram a cada coleção. Infraestrutura Urbana (24), por sua vez, piorou consistentemente da coleção 3.1 para a 7.0.

As principais fontes de confusão ocorreram entre Formação Florestal e Floresta Plantada, e entre Pastagem e Cana de Açúcar. No primeiro caso, a confusão derivou da predominância de vegetação arbórea em ambas as classes, diferenciadas predominantemente pelo maior sombreamento da Formação Florestal derivado da maior irregularidade do dossel. Entre Pastagem e Cana de Açúcar, a confusão foi principalmente por ambas serem dominadas por vegetação arbustiva, sendo bastante semelhantes quando a cana ainda é jovem, além de a cana de açúcar ser uma cultura perene que após a colheita se assemelha espectralmente à pastagem.

A acurácia do usuário, que indica a probabilidade de um pixel classificado dentro de uma certa classe realmente pertencer àquela classe no campo, para Formação Florestal (3) foi maior na coleção 3.1 (98%), menor na coleção 4.1 (92%), e gradualmente melhorou até a coleção 7.0 (0,96). A classe Pastagem (15) apresentou um aumento na precisão ao longo das coleções, enquanto Floresta Plantada (9) teve uma queda considerável na coleção 4.1 (67%) e atingiu seu valor máximo (87%) na coleção 7. A Cana de Açúcar teve o pior resultado na coleção 5 (93%), enquanto Infraestrutura Urbana (24) e Rios, Lagos e Oceanos (33) obtiveram o máximo de precisão (100%) na coleção 6.0.

Em suma, a análise das acurácias global, do produtor e usuário das cinco coleções indicou melhora

da qualidade do mapeamento para a maioria das classes analisadas. A única exceção foi a Infraestrutura Urbana, que atingiu os valores mais baixos de AP e AU na coleção 7.0 (respectivamente, 0,87 e 0,96). Esta leve piora na acurácia da classificação da Infraestrutura Urbana está relacionada a pelo menos dois fatores: i) a maior diversidade de materiais presentes nas zonas urbanas (ex. concreto, telhas de barro, vegetação arbórea e vegetação rasteira), alguns dos quais presentes nas outras categorias mapeadas e por isso passíveis de confusão pelo classificador (ex. solo exposto e telhado de barro), e ii) à taxa mais rápida de mudanças nas composições dos alvos ao longo do tempo, como por exemplo o asfaltamento de ruas que antes eram de terra. Essa variabilidade de alvos na área urbana comumente gera confusão de mapeamento com classes não urbanas, como campo agrícola (GUINDON et al, 2004).

Quando o MapBiomias foi comparado ao outro principal mapeamento da cobertura da terra de São Paulo, realizado pela Secretaria de Infraestrutura e Meio Ambiente do estado de São Paulo (SIMA-SP) com imagens Landsat 5 do ano de 2010, observaram-se quatro classes coincidentes. A acurácia do mapa da SIMA-SP foi estimada considerando-se uma amostra de referência de 100 pontos por classe (total de sete classes e 700 pontos), tendo sido obtida uma acurácia global de 97,14% e número de acertos por classe superiores a 90% (SÃO PAULO, 2013). Esses resultados são comparáveis aos obtidos aqui para a última coleção do MapBiomias, tanto em termos de acurácia global quanto por classe considerada aqui. Não foi encontrado outro mapeamento submetido a uma avaliação sistemática de acurácia para comparação dos resultados.

4. CONSIDERAÇÕES FINAIS

Este estudo apresenta a primeira avaliação sistemática da acurácia e sua evolução ao longo das coleções do MapBiomias em nível estadual. Os resultados indicaram que as coleções 4.1 a 7.0 apresentaram valores de acurácia global superiores a 90%, demonstrando boa qualidade de mapeamento do Estado de São Paulo desde o início do projeto. Notavelmente, a coleção 7.0 obteve o melhor desempenho quando comparado aos valores calculados para todo o Brasil, sugerindo uma melhoria gradual do algoritmo de classificação do MapBiomias para a paisagem de São Paulo. Entretanto, quando o interesse é especificamente por Infraestrutura Urbana, por exemplo para fins de estimativa da expansão dessas áreas, os resultados indicam que a acurácia do MapBiomias vem piorando gradualmente. Quando consideradas as coberturas presentes nas zonas rurais, o fato das AP e AU da coleção 7.0 estarem todas acima de 0,84 indica que o MapBiomias é acurado o suficiente para, por exemplo, muitas análises comuns na escala da paisagem, em ecologia (ex. modelagem da distribuição de espécie de mamíferos ou aves), agricultura (ex. estimativa de safra) ou engenharia (ex. escolha de sítio aeroportuário). Esperamos que a quantificação da acurácia apresentada aqui contribua para a evolução deste importante projeto de mapeamento, e proveja uma base de referência da acurácia para pesquisas baseadas no MapBiomias que tenham o estado de São Paulo como área de estudo.

5. REFERÊNCIAS

- Alves, C.J.P.; Da Silva, E.J.; Müller, C.; Borille, G.M.R.; Guterres, M.X.; Arraut, E.M.; Dos Santos, R.J. (2020). Towards an objective decision-making framework for regional airport site selection. *Journal of Air Transport Management*, 89, 101888.
- Arraut, E. M. et al. (2021). Anticipation of common buzzard population patterns in the changing UK landscape. *Proceedings of the Royal Society B-biological Sciences*, v. 288, p. 20210993. <https://royalsocietypublishing.org/doi/10.1098/rspb.2021.0993>.
- Castojo, A. M. Z.; Jesus, S. C. de. (2022). As Unidades de Conservação do Estado de São Paulo - Planos de Manejo e Representatividade. *Revista Brasileira de Geografia Física* v.15, n.06.
- CANASAT. *Monitoramento da Cana-de-açúcar*. <http://www.dsr.inpe.br/laf/canasat/>
- Carmo, A. B. G.; Gavioli, F. R.; Molin, P. G.(2023). Dinâmica da Paisagem e Classificação Qualitativa das Zonas de Amortecimento de Unidades de Conservação do estado de São Paulo. *Anais do XX Simpósio*

Brasileiro de Sensoriamento Remoto. INPE - Florianópolis-SC. 2023.

- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, v.49, n.12, p.1671-1678.
- Costa, V. M. (2022). *Avaliação das Situações das Áreas de Preservação Permanente no Estado de São Paulo*. 2022. [Trabalho de Conclusão de Curso Engenharia Ambiental]. Universidade Federal de São Carlos, Buri.
- Gindon, B.; Zhang, Y.; Dillabaugh, C. (2004) Landsat Urban Mapping Based on a Combined Spectral–Spatial Methodology. *Remote Sensing of Environment*. v 92. Issue 2, p 218-232, ISSN 0034-4257. <https://doi.org/10.1016/j.rse.2004.06.015>.
- IBGE, Instituto Brasileiro de Geografia e Estatística. (2022, jun 07). *Cidades e Estados*. <https://www.ibge.gov.br/cidades-e-estados/sp.html>.
- IBGE, Instituto Brasileiro de Geografia e Estatística. (2022, out 13). *Estatísticas Sociais*. 27 ago 2020. <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/28668-ibge-divulga-estimativa-dapopulacao-dos-municipios-para-2020>.
- Kenward, R. E.; Arraut, E. M.; Robertson, P. A.; Paredes, S. S.; Casey, N. M.; Aebischer, N. J. Resource-Area-Dependence Analysis: inferring animal resource needs from home-range and mapping data. *PLOS ONE. Article-journal*. <https://doi.org/10.1371/journal.pone.0206354>.
- MAPBIOMAS. (2022, jun 01). *Análise de Acurácia*. <https://mapbiomas.org/analise-de-acuracia>.
- MAPBIOMAS. (2024, abril 04). *Urban Area - Appendix*. 2023. https://brasil.mapbiomas.org/wp-content/uploads/sites/4/2023/08/Urban-Area-Appendix-ATBD-Col.8-v1_revisado.pdf.
- QGIS Development Team e GRASS Development Team. (2022). QGIS 3.22.0 "Białowieża" com GRASS GIS 7.2.2 [Software]. <https://qgis.org> e <https://grass.osgeo.org>.
- Ribeiro, F.O. (2022). O uso do MapBiomias na análise de perda de vegetação natural e apoio a Legislação Florestal atual em Bragança (Pará). *Revista Brasileira de Meio Ambiente*. v.10, n3, 150-167. <https://revistabrasileirademeioambiente.com/index.php/RVBMA/article/view/1252>.
- Santos Junior, E. R. dos. (2023) *Análise dos efeitos da mudança de cobertura da terra sobre a oferta de serviços ecossistêmicos em áreas úmidas periurbanas no estado de São Paulo*. 150p. [Dissertação de Mestrado, Escola de Engenharia de São Carlos]. Universidade de São Paulo, São Carlos.
- São Paulo. Secretaria de Infraestrutura e Meio Ambiente do estado de São Paulo (2023, jun 16). *Mapeamento Temático de Cobertura da Terra do Estado de São Paulo na escala 1:100.000*. <https://www.infraestruturameioambiente.sp.gov.br/cpla/mapa-de-cobertura-da-terra-do-estado-de-sao-paulo/>.
- Sena-Souza, J.P. et al. (2022). Influência do relevo e na dinâmica temporal do uso e cobertura da terra no norte de Minas Gerais, Brasil. *Revista Brasileira de Geografia Física*. v.15, n.05. 2475-2485. <https://periodicos.ufpe.br/revistas/rbgfe/article/view/252707>.
- Smith, C.C , Espírito-Santo, FDB , Healey, Jr , et al. (2020). As florestas secundárias compensam menos de 10% das emissões de carbono mediadas pelo desmatamento na Amazônia brasileira. *Globo. Troque Biol.*; 26 : 7006-7020. <https://doi.org/10.1111/gcb.15352>.
- Souza Jr, C. M. et al. (2020). Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing*. 12 10.3390/rs12172735.
- Stephen V. Stehman, Giles M. Foody. (2019) Key issues in rigorous accuracy assessment of land cover products. *Remote Sensing of Environment*, v 231, 111199, ISSN 0034-4257. <https://doi.org/10.1016/j.rse.2019.05.018>.
- Wickham, J.; Stehman, S. V.; Homer, C. G. (2017). Spatial patterns of the United States National Land Cover Dataset (NLCD) land-cover change thematic accuracy (2001–2011). *International Journal of Remote*

Sensing, 39(6). 1729–1743. <https://doi.org/10.1080/01431161.2017.1410298>.

Contribuição dos Autores

Nadinne Fernandes de Oliveira: Análise formal; Aquisição de financiamento; Investigação; Metodologia; Administração do projeto; Recursos; Software; Validação; Visualização e Redação -minuta inicial.
Eduardo Moraes Arraut: Conceptualização; Supervisão e Redação -revisão e edição.

Conflitos de Interesse

Os autores declaram que não há nenhum conflito de interesse.



Nadinne Fernandes de Oliveira was born in São Gabriel-RS in 1992. She holds a technician's and a Technologist's degree in geoprocessing and a master's degree in Geography from the Federal University of Santa Maria (UFSM). Currently, she is in charge of the Geomatics Laboratory at the Technological Institute of Aeronautics (ITA) in São José dos Campos, where she teaches Geoprocessing and Remote Sensing in ITA's undergraduate and postgraduate programs in Aeronautical Infrastructure Engineering and is involved in various LabGeo research projects.



Esta obra está licenciada com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.