

USO DE SÉRIES TEMPORAIS E ALGORITMO RANDOM FOREST NA CLASSIFICAÇÃO DE ÁREAS PANTANOSAS NO ESTADO DO RIO DE JANEIRO

Evelyn de Castro Porto Costa ¹
Raul Sanchez Vicens ²

INTRODUÇÃO

As paisagens naturais são constantemente transformadas e modificadas pela ação de diferentes atores que ordenam o espaço, dentro de uma lógica que provoca modificações na estrutura, no funcionamento, na dinâmica e, inclusive, na trajetória evolutiva das paisagens originais. Dentre as muitas mudanças nas coberturas terrestres, destacam-se as mudanças das paisagens pantanosas, que são sistemas terrestres que se caracterizam por altos níveis de umedecimento, com alagamentos dinâmicos, possuindo grande relevância ambiental. Os *wetlands* podem ser compreendidos, de forma mais ampla, como áreas cobertas por águas rasas (RUSSI et al, 2013). Moore (2008) aponta que as os *wetlands* são ecossistemas dinâmicos que estão constantemente em evolução.

A Convenção Ramsar (2013) define como “*wetlands*”: “Áreas de pântano, pântano, turfa ou água, seja natural ou artificial, permanente ou temporário, com água estática ou fluindo, fresco, doce ou salgado, incluindo áreas de água do mar, cuja profundidade na maré baixa não exceda seis metros”. Pode ser destacado que a dinâmica das paisagens alagáveis está diretamente relacionada aos impulsos de água, variação climática, pluviosidade, entre outras variáveis ambientais que impactam diretamente no ciclo da água. A dinâmica dessas paisagens pode ser caracterizada como periódicos, onde ocorrem mudanças em recortes temporais similares; e cíclicos, onde retornam ao seu estado inicial através de intervalos de tempo semelhantes, ou seja, em mesmas épocas do ano.

Do ponto de vista ambiental, Moore (2008) aponta que estas áreas ocupam 6% da superfície total do planeta, estando espalhados, geralmente em pequenas manchas, em uma área muito ampla, e sua fragmentação e ampla dispersão aumentam sua vulnerabilidade. Na perspectiva metodológica, esse comportamento disperso e

¹ Doutoranda do POSGEO da Universidade Federal Fluminense - UFF, evelyncastroporto@gmail.com;

² Docente do POSGEO da Universidade Federal Fluminense - UFF, rsvicens@gmail.com.

fragmentado é um dos aspectos que dificultam a sua devida identificação nos mapeamentos de uso e cobertura da terra, sendo essas áreas pantanosas, muitas das vezes, subestimados por questão de semântica, de generalização, ou ainda em função de interpretações que não levam em consideração a sazonalidade dessas feições. Em sua grande maioria, as paisagens de *wetlands* são classificadas como áreas de pastagem, água, ou cobertura de gramíneas.

A área de estudos desta pesquisa é o estado do Rio de Janeiro. Se comparado com outros territórios, no contexto Brasil, esse Estado dispõe de pequenos e fragmentados *wetlands*, entretanto conta com diversidade de ambientes inundados, compostos por sistemas deltaicos, lagunas costeiras, aluviais, marinhos, palustres, por exemplos, o que os torna de relevante interesse ambiental. Na área de estudos, essas paisagens pantanosas são relevantes ecossistemas de interesse ambiental, econômico e social, que possuem um equilíbrio dinâmico próprio e abrangem riquezas ambientais e sociais. Entretanto, essas áreas sofrem constantes pressões e interferências antrópicas, que desequilibram seus estados dinâmicos e seu poder de auto regulação. Deve ser destacado a importância desses geossistemas, tendo em vista sua relevância ao ciclo hidrológico e controle de inundações, por exemplo, prestando serviços ecossistêmicos importantes, entretanto, essas áreas sofrem grande influência antrópica, estando vulneráveis à poluição, canalizações, obras, edificações e usos agropastoris.

Nesse sentido, o objetivo dessa pesquisa é mapear as áreas pantanosas localizadas no Estado do Rio de Janeiro, utilizando séries temporais de imagens de satélite de sensores óticos e de radar, atrelados a ferramentas de aprendizado de máquina para a classificação de imagens orbitais. Ressalta-se que a grande maioria dos mapeamentos de uso e cobertura da terra, adotam apenas uma única data de imagem para a classificação, o que pode subestimar a classificação de *wetlands* temporariamente inundados. Dessa forma, o uso de séries temporais remete-se a necessidade de identificar multitemporalmente a variação espacial dos *wetlands*, tendo em vista a dinâmica espaço-temporal desses ambientes.

METODOLOGIA

Tendo em vista a demanda por trabalhar com séries temporais em um grande recorte espacial, optou-se pelo processamento em nuvem, a partir da plataforma gratuita Google Earth Engine (GEE), que permite o rápido processamento de imagens, sem

demandar do uso de computadores de grande capacidade de armazenamento e memória. Na plataforma GEE foram selecionadas algumas especificidades da imagem, através do script, utilizando linguagem de programação. Foram adotadas imagens do satélite europeu da ESA (Agência Espacial Europeia), oriundos dos satélites Sentinel 1 (sensor SAR) Sentinel 2 (sensor MSI).

O sensor de radar (Sentinel 1) foi escolhido devido a sua sensibilidade a classificação de corpos d'água, pois as imagens de radares geralmente não são prejudicadas por efeitos atmosféricos e são capazes de gerar imagens através de nuvens tropicais e pancadas de chuva. Entretanto, esse sensor não é capaz de capturar as imagens com teor de umidade e solos encharcados com presença de vegetação, dessa forma, concomitantemente, foi utilizado séries temporais de sensores óticos (Sentinel 2).

O recorte temporal se refere ao total de 36 meses, no qual foram criados *composites* mensais de cada mês dos anos de 2019, 2020 e 2021. Destaca-se que esse recorte temporal foi escolhido por se tratar dos últimos três anos contínuos e de pluviosidade expressiva, que em conjunto, permitem compreender a dinâmica sazonal das áreas alagáveis. A classificação multitemporal das imagens separadamente para o recorte temporal definido, bem como, para as imagens adotadas, tendo em vista as especificidades técnicas do pré processamento de imagens de radar (SAR) e imagens óticas (MSI).

Destaca-se que a classificação realizada no Google Earth Engine (GEE) foi a classificação pixel a pixel e o algoritmo de aprendizagem de máquina utilizado foi o *Random Forest*, que combina classificações provenientes de várias árvores de decisão, treinadas individualmente para obter a classificação desejada. Logo, a classe vencedora é aquela que aparece o maior número de vezes, ou seja, votada pela maioria. Na classificação SAR foram adotadas, com 2.958 amostras de treinamento, enquanto na classificação MSI foram adotadas 4.901 amostras, essa variação se refere a diversidade espectral demandas para o sensor ótico. Em ambas as classificações, foram adotadas 500 árvores de decisão na classificação.

O índice espectral NDWI (Índice de Diferença Normalizada de Água) também foi utilizado no sensor óptico. Esse índice é muito utilizado para ressaltar corpos hídricos e minimizar a resposta dos demais alvos. O NDWI proposto por McFeeters (1996) é calculado usando a Equação: $(Green - NIR) / (Green + NIR)$, no qual, Green é a

refletância na faixa verde do visível e o NIR é a refletância na faixa do infravermelho próximo.

As classificações se deram de modo binário, tendo apenas duas classes a serem classificadas, tais como: “paisagens pantanosas” e “outros usos”. Até que por fim, as classificações foram unidas no software de geostística R studio e posteriormente validadas. O fluxograma metodológico apresenta mais detalhes das etapas metodológicas da pesquisa (Figura 1).

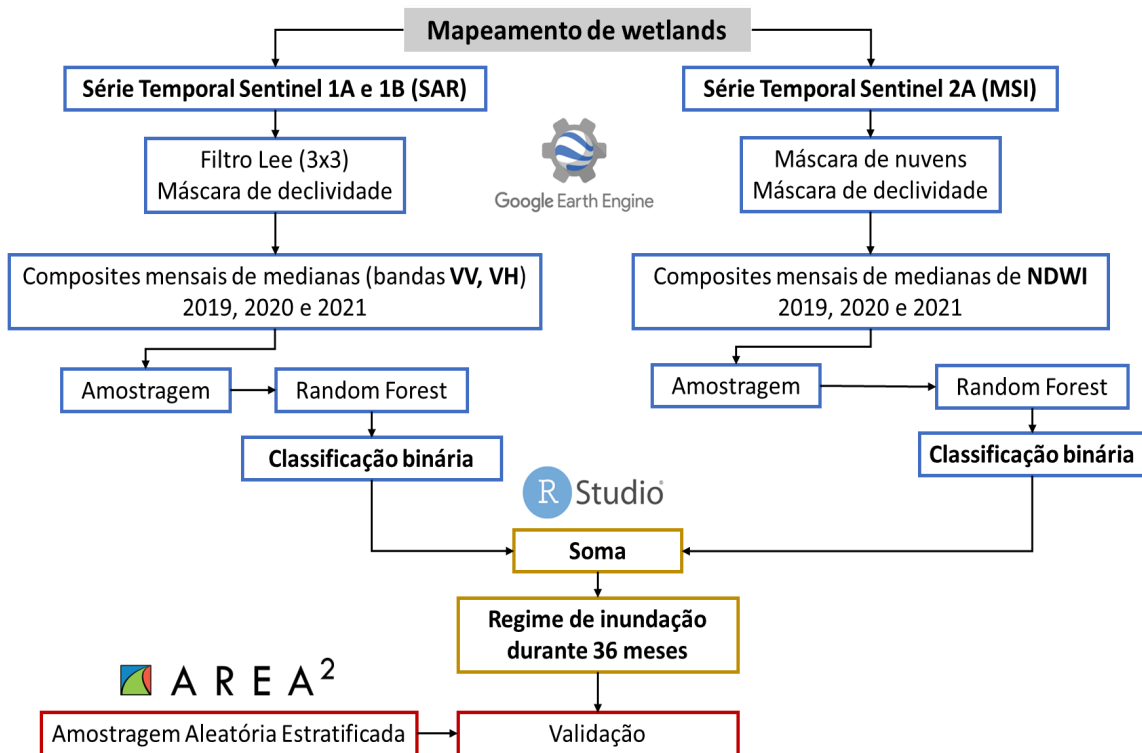


Figura 1. Procedimentos metodológicos adotados na pesquisa. Fonte: os autores (2024).

Para a validação do mapeamento foi adotado a metodologia por amostragem aleatória e estratificada. Para tanto, foi adotado o método AREA² (“área ao quadrado” ou “área dois), que é a abreviação de Área Estimada & Avaliação de Precisão, que é um aplicativo do GEE que fornece suporte abrangente para amostragem e estimativa em uma estrutura de inferência baseada em design (BULLOCK & OLOFSSON, 2018).

Nessa pesquisa, foi adotado do aplicativo do AREA², o desenho amostral do tipo aleatório estratificado, que requer a especificação de um mapa para definir os estratos, um tamanho total da amostra e uma alocação de unidades amostrais nos estratos. Ela oferece a opção de aumentar o tamanho da amostra em classes que ocupam uma pequena proporção da área para reduzir os erros padrão das estimativas de precisão

específicas da classe (OLOFSSON et al, 2014). Dessa forma, o desenho amostral da validação aleatória e estratificada, considerou a quantidade de amostras de acordo com a sua extensão territorial na área de estudos, considerando 140 amostras de paisagens pantanosas e 100 amostras de outras classes.

RESULTADOS E DISCUSSÃO

A união das classificações oriundas dos sensores óticos e de radar, bem como, o uso de filtros para redução de ruídos, trouxe como resultado um mapeamento síntese, considerando o regime de inundação das paisagens pantanosas. Ou seja, estas áreas foram classificadas através do seu regime de inundação, sendo possível analisar a dinâmica espaço-temporal dessas paisagens (Figura 2).

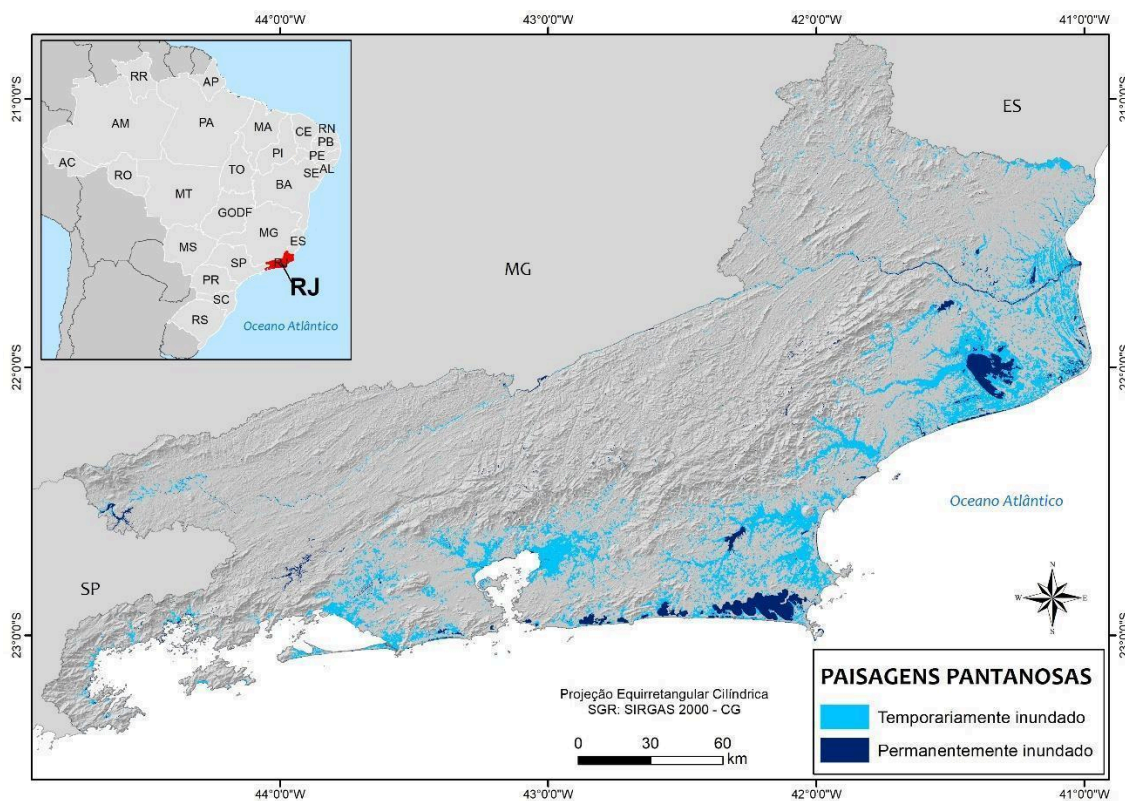


Figura 2. Classificação do regime de inundação das paisagens pantanosas do Estado do Rio de Janeiro. Fonte: os autores (2024)

Como resultado quantitativo, é possível analisar que as paisagens pantanosas correspondem a cerca de 10% da superfície do Estado do Rio de Janeiro, sendo classificados por regimes de inundação temporária ou permanente (Tabela 1).

Classificação	Hectares	%
Temporariamente inundado	409.153,96	8,7
Permanentemente inundado	89.781,25	1,9
TOTAL	4.727.004,09	100

Tabela 1. Regime de inundação. Fonte: os autores (2024).

As paisagens pantanosas que se apresentam de forma permanente se caracterizam por estar presente em 1,9% de todo o território fluminense. Essa classe considera, principalmente, os espelhos d'água rasas da região, tais como as lagunas costeiras, drenagens fluvio-marinhas e demais corpos rasos continentais que permanecem inundados por água por até 36 meses ininterruptamente. Exemplos de corpos d'água classificadas como “permanentemente inundadas” podem ser identificados na Lagoa Feia, localizada em Campos dos Goytacazes, a Lagoa de Jurnaíba, em Silva Jardim e a Lagoa de Araruama, na região dos Lagos, que podem ser caracterizados como *wetlands* possui uma média de 2 a 3 metros de profundidade.

As paisagens pantanosas caracterizadas como temporariamente inundadas, são representadas por áreas inundadas que se apresentam de forma mais dinâmica no território, com comportamentos mais sazonais, variando de acordo com as estações mais secas e mais úmidas, ou até mesmo eventos extremos. Ressalta-se que a classe “temporariamente inundada” é representada por um regime de inundação entre 6 a 35 meses. A partir da metodologia adotada para a validação do mapeamento, destaca-se o excelente desempenho da classificação. Através da matriz de erro gerada pela validação do mapeamento se torna possível analisar que o erro de comissão foi maior que o erro de omissão.

	Referência Terrestre		
	Classes	Paisagens pantanosas	Outras classes
Classificação	Paisagens pantanosas	112	28
	Outras classes	4	96
	Total	116	124

Tabela 2. Matriz de Erro da classificação de imagens. Fonte: os autores (2024)

Os erros de comissão se referem a classificação errônea de água em áreas de sombras. Dentre as 240 amostras totais, apenas 32 foram classificadas equivocadamente, sendo 28 por erros de comissão e 4 por erros de omissão. Esse resultado sinaliza poucos erros de comissão na classe de “Paisagens pantanosas” e de omissão na classe de “Outras classes”.

Os maiores erros estavam presentes nas áreas de sombra e afloramento rochoso, que foram classificados equivocadamente como paisagens pantanosas, configurando como erro de comissão. Já os erros por parte de omissão, se configuram como áreas planas que deveriam ter sido classificadas como Paisagens pantanosas e não foram. Esse ruído pode ter sido reflexo dos padrões utilizados na classificação, bem como, os índices espectrais que apoiaram a classificação. Entretanto, ressalta-se que esses erros de omissão foram considerados baixos e pouco expressivos para o resultado da validação, que obteve bons valores de desempenho.

A validação do mapeamento apontou bons resultados de desempenhos com um índice Kappa alcançando 0,73, que é classificado por Landis e Koch (1977) como uma classificação “Muito Boa”. Além disso, obteve-se uma exatidão global de 0,86, uma acurácia total de 0,867 e um F-score de 0,875. O que representa um bom desempenho de validação, dando confiabilidade ao uso desse mapeamento para diversas finalidades (Tabela 3).

Indicadores	Validação
Exatidão Global	0,87
Índice Kappa	0,74
Acurácia Total (Overall accuracy)	0,87
Acurácia do Produtor (Precision)	0,80
Acurácia do Usuário (Recall)	0,97
F-score	0,88

Tabela 3. Resultados da validação de wetlands. Fonte: os autores (2024)

CONSIDERAÇÕES FINAIS

As áreas classificadas como paisagens pantanosas se localizam em áreas de planície costeira ou fluvial, com predisposição para regimes de inundação devido a sua gênese. Outro aspecto se trata das áreas no entorno de sistemas lagunares e baías,

mantendo sua disposição espacial em áreas próximas ao litoral. Logo, o resultado destes mapeamentos permitiu elaborar um mapeamento síntese, classificando as paisagens pantanosas quanto a frequência de inundação. Esse dado permite compreender as dinâmicas dos geossistemas inundáveis do estado do Rio de Janeiro, bem como, classificar espaço-temporalmente a amplitude dessas paisagens.

Do ponto de vista metodológico, ressalta-se o uso das imagens de radar como importante recurso para este mapeamento, pois permite uma classificação temporal sem a interferência da atmosfera, principalmente nuvens, que afetam diretamente este resultado. Além disso, as imagens utilizadas apresentam um diferencial em relação a sua resolução espacial, que é de 10 metros, apresentando um resultado de maior detalhamento a nível de pixel. Nesse sentido, essa classificação apoia a identificação de objetos menores, como o caso de *wetlands*, que podem ser subestimados nas classificações devido a sua dimensão espacial reduzida.

O uso de imagens multiespectrais, como as imagens óticas Sentinel 2, também foi satisfatório para essa pesquisa, tendo em vista que possibilitou a utilização do índice radiométrico NDWI, auxiliando na identificação dos *wetlands*, que são alvos bastante complexos, tendo em vista que sua resposta espectral se confunde com classes, como áreas agropastoris e afloramentos rochosos. Ressalta-se que o uso de algoritmos computacional Random Forest foi de grande relevância para a classificação de imagens de sistemas ambientais dinâmicos e complexos.

Palavras-chave: *Wetlands*, *Machine Learning*, *Google Earth Engine*, Sentinel 1 e 2.

AGRADECIMENTOS

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

REFERÊNCIAS

BULLOCK, Eric; OLOFSSON, Pontus. **What is AREA2?** Disponível em: <https://area2.readthedocs.io/en/latest/overview.html>.

LANDIS, R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics**, v.33, n.1, p.159-174, Mar. 1977.

MOORE, Peter D. Ecosystem: **Wetlands**. Revised Edition. New York, NY, 2008.

OLOFSSON, Pontus et al. Good practices for estimating area and assessing accuracy of land change. **Remote sensing of Environment**, v. 148, p. 42-57, 2014.



**Simpósio Brasileiro
de Geografia Física Aplicada**

IV Encontro Lusofonamericano de Geografia Física e Ambiente

RAMSAR. Manual de la Convención de Ramsar: Guía a la Convención sobre los Humedales (Ramsar, Irán, 1971), 6a. edición. Secretaría de la Convención de Ramsar, Gland (Suiza), 2013

RUSSI D., TEN BRINK P., FARMER A., BADURA T., COATES D., FÖRSTER J., KUMAR R. AND DAVIDSON N. The Economics of Ecosystems and Biodiversity for Water and Wetlands. IEEP, London and Brussels; Ramsar Secretariat, Gland, 2013.