

Detecção e delimitação automática de corpos hídricos em imagens Sentinel-2: uma proposta de integração do algoritmo *Fmask* aos índices espectrais NDWI e MNDWI

Thales Vaz Penha¹, Mikhaela Aloísia Jéssie Santos Pletsch¹, Celso Henrique Leite Silva Junior², Thales Sehn Körting¹, Leila Maria Garcia Fonseca¹

¹Divisão de Processamento de Imagens – Instituto Nacional de Pesquisas Espaciais
Caixa Postal 12227-010 – São José dos Campos – SP – Brazil

²Divisão de Sensoriamento Remoto – Instituto Nacional de Pesquisas Espaciais
Caixa Postal 12227-010 – São José dos Campos – SP – Brazil

{thales.penha, mikhaela.pletsch, celso.junior, thales.korting,
leila.fonseca}@inpe.br

Abstract. *Precise delimitation of water bodies is essential for several studies including watershed analysis. Its detection is able through techniques, such as spectral water indices, NDWI and MNDWI, and Fmask algorithm. Considering that individually those approaches are not accurate in different environments, we aimed in this work to evaluate the integration of Fmask to the NDWI and MNDWI for automatic detection of water bodies in Sentinel-2 images. The results indicate that the combination of these methods allow to reduce omission errors in an automatic process of water bodies detection.*

Resumo. *A delimitação e detecção precisa de corpos hídricos é essencial para diversos estudos, incluindo análises de bacias hidrográficas. A sua detecção é possível por meio de técnicas como os índices espectrais de água, NDWI e MNDWI, e o algoritmo Fmask. Considerando que esses métodos isoladamente podem não ser precisos em diferentes ambientes, o objetivo deste trabalho foi avaliar a integração do Fmask aos índices NDWI e MNDWI para a detecção automática de corpos hídricos em imagens Sentinel-2. Os resultados obtidos indicam que o uso combinado destes métodos ajuda a diminuir os erros de omissão em um processo automático de detecção de corpos hídricos.*

1. Introdução

A delimitação e o monitoramento de corpos hídricos são aplicações fundamentais presentes no sensoriamento remoto [McFeeeters, 1996], uma vez que seus dados podem subsidiar a avaliação e a análise de recursos hídricos, incluindo inventários e mapeamentos de áreas úmidas e enchentes, dando suporte assim ao gerenciamento de águas superficiais [Rokni et al., 2014]. Para o mapeamento de corpos hídricos, diversos métodos foram desenvolvidos tendo como base imagens multiespectrais [Du et al., 2016]. Dentre os métodos mais utilizados, Jiang et al. (2014) destacam os classificadores supervisionados, não supervisionados e os índices espectrais de água. Os índices espectrais de água têm sido amplamente utilizados devido à precisão na detecção de corpos hídricos a baixo custo computacional [Jiang et al., 2014; Du et al., 2016]. Os principais índices utilizados são o NDWI (*Normalized Difference Water Index*) e o MNDWI (*Modified Normalized Difference Water Index*), os quais realçam feições de

água de forma eficaz na maioria dos casos (verificar seção 2.3). No entanto, dependendo da dinâmica de uso e cobertura da terra da área de estudo, esses índices podem apresentar erros de comissão e omissão, devido à confusão espectral com alvos escuros na imagem, como cicatrizes de queimadas e sombras [Bochow et al., 2012]. Assim, a simples aplicação de um índice, por vezes, não é suficiente para uma delimitação precisa, uma vez que é influenciado por outros tipos de coberturas da terra [Ji et al., 2009].

O algoritmo *Fmask* (*Function of Mask*) (Zhu et al. 2015) é uma ferramenta capaz de detectar feições em imagens da Série Landsat (4-8) e Sentinel-2. O produto gerado pelo *Fmask* consiste em uma imagem de saída com os alvos nuvem, sombra de nuvem, neve, água e terreno que podem ser utilizados como máscaras. Nesse contexto, os produtos gerados pelo *Fmask* podem tanto contribuir para refinar a delimitação de corpos hídricos, ao ser integrado aos índices espectrais, quanto para eliminar alvos que causam confusão espectral, como sombra de nuvem. Levando em consideração que esse algoritmo se apresenta como uma potencial ferramenta de auxílio na detecção automática de corpos hídricos em imagens multiespectrais, o presente trabalho teve como objetivo avaliar o uso integrado do algoritmo *Fmask* aos índices espectrais de água NDWI e MNDWI em duas cenas do sensor MSI Sentinel-2 para a detecção e delimitação automática de corpos hídricos no bioma Cerrado.

2. Materiais e Métodos

2.1. Áreas de Estudo

As áreas de estudos estão situadas nos estados brasileiros de Goiás, Mato Grosso do Sul e Minas Gerais, dentro dos limites do bioma Cerrado (Figura 1). As áreas de estudo correspondem aos limites das cenas 23KNV (Área Piloto 1 - AP1) e 22KCE (Área Piloto 2 – AP2) do sensor MSI do Sentinel-2A. Estas regiões foram selecionadas devido à proximidade de alguns afluentes a importantes bacias hidrográficas do Cerrado brasileiro, incluindo o Rio São Francisco (AP1) e o Rio Paraná (AP2), os quais são fontes de recursos hídricos e abastecimento para as regiões sul e sudeste do Brasil. Nessas regiões, as principais coberturas da terra, para o ano de 2013, na AP1 foram áreas Natural/Natural Não Vegetado (47,96%), Pastagens (39,87%) e Silvicultura (9,72%), enquanto na AP2 as coberturas predominantes foram Agricultura Perene (41,55%), áreas Natural/Natural Não Vegetado (26,77%) e Agricultura (22,01%) [MMA, 2015].

2.2. Base de dados

As duas cenas MSI Sentinel-2 foram adquiridas, respectivamente, para as datas de passagem do satélite de 01 de julho de 2017 (AP1) e 09 de agosto de 2017 (AP2) no *website Copernicus Scientific Data Hub* (ESA - Agência Espacial Europeia, disponível em: <http://www.scihub.copernicus.eu/>) como produto *Top-Of-Atmosphere* (TOA) *Level-1C* (L1C) com correção radiométrica e geométrica no sistema de projeção UTM/WGS84, fuso 23 sul (AP1) e fuso 22 sul (AP2). As imagens Sentinel-2 apresentam potencial para o mapeamento em escalas regionais de corpos hídricos devido às propriedades do sensor MSI, resolução espacial de 10-60m, resolução temporal de 10 dias e distribuição gratuita das cenas [Du et al., 2016]. Para o desenvolvimento do presente trabalho, foi preciso reamostrar a banda 11 (SWIR) do sensor MSI, uma vez que esta possui resolução espacial de 20m enquanto as bandas 3 (verde) e 8 (vermelho) possuem 10m.

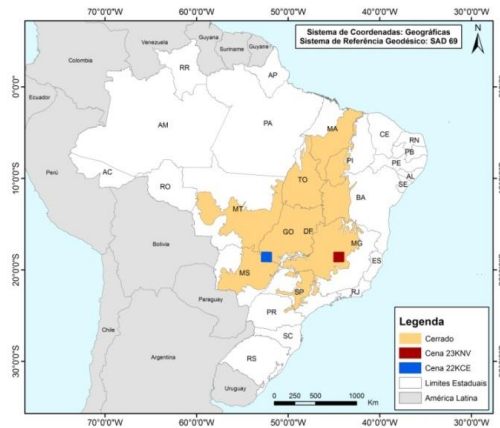


Figura 1. Mapa de localização das áreas de estudo.

2.3. Índices Espectrais de Água

O NDWI (Equação 1), proposto por McFeeters (1996), foi concebido visando maximizar a reflectância da água na banda do verde e minimizar a reflectância na banda NIR (Infravermelho próximo) [Du et al., 2016].

$$NDWI = \frac{(\rho_3 - \rho_8)}{(\rho_3 + \rho_8)} \quad (1)$$

onde: ρ_3 é a reflectância TOA da banda 3 (banda verde) e ρ_8 é a reflectância TOA da banda 8 (banda NIR) da imagem Sentinel-2 MSI.

O MNDWI (Equação 2) foi concebido por Xu (2006) e teve como objetivo minimizar a principal limitação do NDWI, a ineficiência em suprimir o ruído proveniente das características das áreas construídas [Xu, 2006]. Assim, esse índice foi desenvolvido considerando que um corpo hídrico apresenta maior absorção na banda SWIR (Infravermelho de ondas curtas) se comparado ao da banda NIR, diferentemente das áreas construídas. Neste trabalho, valores positivos (>0.0) em ambos os índices foram considerados como corpos hídricos [McFeeters, 1996; Xu, 2006].

$$MNDWI = \frac{(\rho_3 - \rho_{11})}{(\rho_3 + \rho_{11})} \quad (2)$$

onde, ρ_3 é a reflectância TOA da banda 3 (banda verde) e ρ_{11} é a reflectância TOA da banda 11 (banda SWIR) da imagem Sentinel-2 MSI.

3. Metodologia

A metodologia deste trabalho consistiu em três etapas principais (Figura 2). Primeiramente, o algoritmo *Fmask* foi aplicado nas imagens Sentinel-2 por meio da linguagem *open source Python* para a obtenção dos dados relativos à detecção de água, os quais foram utilizados posteriormente como dados complementares aos índices de água. As demais feições identificadas pelo *Fmask*, como nuvem e sombra de nuvem, foram utilizadas na etapa de filtragem, visando eliminar eventuais ruídos da detecção de corpos hídricos. Em uma segunda etapa, realizou-se a compatibilização das resoluções espaciais

das bandas do Sentinel-2 por processo de reamostragem (vizinho mais próximo) para o cálculo dos índices de água em cada uma das APs. Posteriormente, os índices NDWI e MNDWI, e as áreas identificadas como água pelo *Fmask* foram integrados e combinados por meio de uma álgebra de mapas (soma). Os resultados foram filtrados com base nas feições de nuvem e sombra de nuvens geradas pelo algoritmo *Fmask*. Por fim, realizou-se uma validação dos resultados utilizando um dado de referência para a geração de matrizes de confusão para cada método empregado nas duas APs.

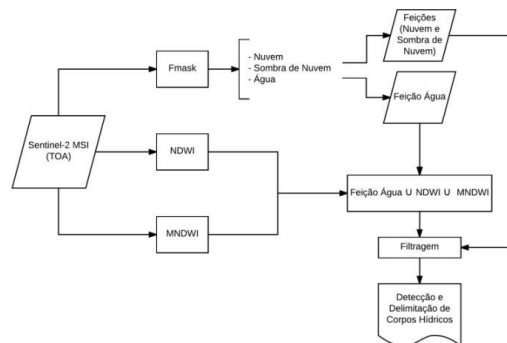


Figura 2. Fluxograma do trabalho.

O dado de referência, utilizado na etapa de validação, foi obtido a partir do Mapa brasileiro de lâmina de água (Namikawa et al., 2016). Com o intuito de verificar a eficiência da metodologia proposta, foram geradas quatro matrizes de confusão para cada área de estudo, considerando: a) os resultados dos dois índices espectrais de água; b) a detecção de água realizada pelo algoritmo *Fmask*; c) a integração dos dados de água dos índices espectrais e do *Fmask*. Para avaliar a exatidão da delimitação de corpos hídricos, considerando o alvo *Água* frente aos demais elementos presentes na imagem, nomeados como *Não Água*, foram calculadas a acurácia do produtor (relativa aos erros de omissão), a acurácia do usuário (relativa aos erros de comissão), a exatidão global e o índice kappa (Congalton, 2001).

4. Resultados e Discussão

A Tabela 1 apresenta os resultados das matrizes de confusão para as áreas AP1 e AP2. De modo geral, observou-se que a exatidão dos métodos propostos foi alta (exatidão global acima de 99% e índice kappa entre 0,64 e 0,71), entretanto, ao analisar especificamente o alvo *Água*, nota-se que as acurácias do produtor e usuário foram relativamente baixas e variaram significativamente de acordo com o método utilizado.

Para a AP1, a detecção de corpos hídricos do *Fmask* (Tabela 1A) apresentou melhor acurácia do usuário (90,43%), enquanto o índice MNDWI (Tabela 1E) apresentou melhor acurácia do produtor (59,10%). Os erros de omissão foram elevados (em torno de 41- 46%). Já na AP2, a detecção de corpos hídricos pelo MNDWI (Tabela 1F) apresentou melhor acurácia do usuário (61,54%) e acurácia do produtor (83,78%). No entanto, este índice não foi consideravelmente superior aos outros dois métodos (NDWI e *Fmask*), indicando que quando utilizados de forma isolada os três métodos omitem muita informação de corpos hídricos. Além disso, a acurácia do usuário foi baixa (em torno de 52 - 61%), apresentando mais erros de comissão. Assim como na AP1 nenhum dos três métodos utilizados isoladamente foi consideravelmente superior ao outro. A dificuldade na detecção e delimitação de corpos hídricos nas APs pode ser explicada pela presença de

outras feições na imagem apresentarem potencial de confusão, como sombras geradas pela mata ciliar dos rios, e pelo fato do tamanho das feições hídricas serem pequenas.

Tabela 1. Matrizes de confusão para AP1 e AP2.

A		AP1 - Detecção Fmask			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
Fmask	Água	45.75	4.84	50.59	90.43
	Não Água	38.78	12100.53	12139.31	99.68
Total (Km ²)		84.53	12105.37	12189.90	
Acurácia do Produtor (%)		54.12	99.96	Exatidão Global	99.64
				Kappa	0.68
B		AP2 - Detecção Fmask			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
Fmask	Água	42.50	29.34	71.84	59.16
	Não Água	12.69	11972.06	11984.75	99.89
Total (Km ²)		55.19	12001.40	12056.59	
Acurácia do Produtor (%)		77.01	99.76	Exatidão Global	99.65
				Kappa	0.67
C		AP1 - Detecção NDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
NDWI	Água	49.10	12.99	62.09	79.08
	Não Água	35.43	12092.38	12127.81	99.71
Total (Km ²)		84.53	12105.37	12189.90	
Acurácia do Produtor (%)		58.09	99.89	Exatidão Global	99.60
				Kappa	0.67
D		AP2 - Detecção NDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
NDWI	Água	45.85	41.14	86.99	52.71
	Não Água	9.34	11960.26	11969.60	99.92
Total (Km ²)		55.19	12001.40	12056.59	
Acurácia do Produtor (%)		83.08	99.66	Exatidão Global	99.58
				Kappa	0.64
E		AP1 - Detecção MNDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
MNDWI	Água	49.96	17.94	67.90	73.58
	Não Água	34.57	12087.43	12122.00	99.71
Total (Km ²)		84.53	12105.37	12189.90	
Acurácia do Produtor (%)		59.10	99.85	Exatidão Global	99.57
				Kappa	0.65
F		AP2 - Detecção MNDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
MNDWI	Água	46.24	28.90	75.14	61.54
	Não Água	8.95	11972.50	11981.45	99.93
Total (Km ²)		55.19	12001.40	12056.59	
Acurácia do Produtor (%)		83.78	99.76	Exatidão Global	99.69
				Kappa	0.71
G		AP1 - Detecção Fmask-NDWI-MNDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
Método proposto	Água	52.18	13.91	66.09	78.95
	Não Água	32.34	12028.31	12060.65	99.73
Total (Km ²)		84.52	12042.22	12126.74	
Acurácia do Produtor (%)		61.74	99.88	Exatidão Global	99.62
				Kappa	0.69
H		AP2 - Detecção Fmask-NDWI-MNDWI			
		Referência (Km ²)		Total (Km ²)	Acurácia Usuário (%)
Método proposto	Água	48.36	38.88	87.24	55.43
	Não Água	6.82	11908.17	11914.99	99.94
Total (Km ²)		55.18	11947.05	12002.23	
Acurácia do Produtor (%)		87.64	99.67	Exatidão Global	99.62
				Kappa	0.68

Nas Tabelas 1G e 1H, é possível observar os resultados do método proposto (integração do *Fmask* com os índices) com a aplicação do filtro de nuvem e sombra de nuvens. No caso da AP1 (Tabela 1G), houve maior acerto de áreas de corpos hídricos em relação à referência (4,63% de ganho em relação à média dos três métodos isolados), ou seja, acurácia do produtor foi superior. Porém, do ponto de vista dos erros de comissão houve uma perda da ordem de 2% em relação à média dos três métodos isolados. Já na AP2 (Tabela 1H), também houve maior acerto de áreas de corpos hídricos e menores erros de omissão (6,65% de ganho em relação à média dos três métodos isolados). Já a acurácia do usuário praticamente se manteve estável (perda de 0,33% em relação à média dos três métodos isolados). A integração do *Fmask* com os índices apresentou maior equilíbrio dos erros de comissão e diminuiu os erros de omissão, o que representa uma melhora do ponto de vista de exatidão do alvo *Água*.

5. Conclusões

Após a análise das matrizes de confusão, foi possível concluir que a metodologia proposta para a detecção de corpos hídricos de forma integrada mostrou-se eficiente. Apesar de não apresentar uma melhora significativa na exatidão de todo o mapeamento, a metodologia proposta contribuiu para diminuir os erros de omissão na detecção e delimitação de corpos hídricos no Cerrado. No entanto, os erros de comissão persistem,

o que pode ser minimizado quando se elimina os ruídos, como as feições de nuvem e sombra de nuvens na etapa de filtragem.

Do ponto de vista da automatização do processo de detecção de corpos hídricos, o uso da linguagem *Python* permitiu uma rápida estruturação dos procedimentos metodológicos, bem como sua execução com baixo custo computacional. Possibilitando assim a replicação da metodologia para outras áreas de estudo, além de apresentar potencial para a delimitação de corpos hídricos em qualquer imagem da série Sentinel-2. Como trabalhos futuros, novas combinações entre o *Fmask* e os índices de água devem ser exauridas, assim como, outras formas de refinamento podem ser investidas, como a aplicação de filtros, a utilização de classificadores automáticos ou mesmo a utilização de um mapa de uso e cobertura da terra, a fim de evitar falsos positivos.

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

Referências

- Bochow, M., Heim, B., Küster, T., Rogäß, C., Bartsch, I., Segl, K., Sandra Reigber, & Kaufmann, H. (2012). On the use of airborne imaging spectroscopy data for the automatic detection and delineation of surface water bodies, *Remote Sensing of Planet Earth*, 3–23.
- Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, 10(10), 321–328. <https://doi.org/10.1071/WF01031>
- Du, Y., Zhang, Y., Ling, F., Wang, Q., Li, W., & Li, X. (2016). Water bodies' mapping from Sentinel-2 imagery with Modified Normalized Difference Water Index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sensing*, 8(4). <https://doi.org/10.3390/rs8040354>
- Ji, L., Zhang, L., & Wylie, B. (2009). Analysis of dynamic thresholds for the Normalized Difference Water Index. *Photogrammetric Engineering & Remote Sensing*, 75(11), 1307–1317. <https://doi.org/10.14358/PERS.75.11.1307>
- Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., & Xiao, T. (2014). An automated method for extracting rivers and lakes from Landsat imagery. *Remote Sensing*, 6(6), 5067–5089. <https://doi.org/10.3390/rs6065067>
- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- MMA, Ministério do Meio Ambiente. (2015). Mapeamento do Uso e Cobertura do Cerrado: Projeto TerraClass Cerrado 2013. Disponível em: < <http://www.dpi.inpe.br/tccerrado/index.php?mais=1> > Acessado em set/2017.
- Namikawa, L. M., Körting, T. S., & Castejon, E. F. (2016). Water body extraction from Rapideye images: an automated methodology based on hue component of color transformation from RGB to HSV model. *Revista Brasileira de Cartografia*, 68(6), 1097–1111. Disponível em: <<http://www.rbc.lsie.unb.br/index.php/rbc/article/view/1662>>. Acessado em set/2017.
- Rokni, K., Ahmad, A., Selamat, A., & Hazini, S. (2014). Water feature extraction and change detection using multitemporal landsat imagery. *Remote Sensing*, 6(5), 4173–4189. <https://doi.org/10.3390/rs6054173>
- Xu, H. (2006). Modification of Normalized Difference Water Index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. <https://doi.org/10.1080/01431160600589179>
- Zhu, Z., Wang, S., & Woodcock, C. E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment*, v. 159, p. 269-277. <https://doi.org/10.1016/j.rse.2014.12.014>